

A meta-evaluative synthesis of the effects of custodial and community-based offender rehabilitation

European Journal of Criminology

1–27

© The Author(s) 2024



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14773708241256501

journals.sagepub.com/home/euc**Johann Koehler** 

London School of Economics and Political Science, UK

Friedrich Lösel

University of Cambridge, UK; University of Erlangen-Nuremberg, Germany

Abstract

We synthesize 53 meta-analyses on the effectiveness of correctional treatment applied to a wide variety of offender groups delivered in either custodial or community-based settings. Those meta-analyses revealed positive overall effects on reoffending of correctional treatment delivered in both settings. However, the treatment setting is also associated with complex moderator effects. With respect to *effect size*, for most groups, community-based correctional treatment is associated with statistically significant larger reductions in reoffending than treatments delivered in custodial settings. With respect to *effect precision*, custodial treatments report more consistent effects on reoffending than community-based treatments. The findings extend and develop the insight that treatment flexibility, such as is found among community-based treatments, can optimize program effectiveness. Likewise, the opportunities for monitoring and treatment fidelity that custodial settings enable can homogenize outcomes. Nonetheless, the promising results observed among treatments delivered both inside and outside institutional settings implicate a complex policy tradeoff between prioritizing strong performance and consistent effects.

Keywords

Correctional treatment, meta-analysis, offender rehabilitation, reoffending, review of reviews

Corresponding author:

Johann Koehler, Department of Social Policy, London School of Economics and Political Science, Houghton St., London WC2A 2AE, UK.

Email: j.koehler@lse.ac.uk

Introduction

Criminologists have long sought to pry open punishment's black box to understand the prospects of meaningfully rehabilitative criminal justice. To that end, meta-analytic evidence on the effects of correctional treatment consistently shows mean positive effects in reducing reoffending. Yet those meta-analytic insights often skirt thorny questions about the role of custodial and community settings in rehabilitative criminal justice. Consequently, hard battle lines have developed over the propriety of offender rehabilitation programs within correctional institutions and outside them. In particular, prior correctional research leaves few cues to inform policy-makers who are balancing complex tradeoffs between policy imperatives extending from strong performance to consistent effects. We therefore compare meta-analytic findings on the effectiveness of correctional treatment applied to a wide variety of offender groups delivered in either the community or institutional settings: if meaningful rehabilitation is among criminal justice's goals, then *does the treatment setting matter?*

Proponents of community-based rehabilitation doubt that sound correctional treatment can overcome the criminogenic influences that custodial settings foster. These qualities include, but are not limited to, "contagion" effects within criminal subculture (e.g., Bayer et al., 2009; Ouss, 2011; Stevenson, 2017); instruction in skills conducive toward offending (e.g., Damm and Gorinas, 2020); adoption of antisocial values and identities (e.g., Pyrooz and Decker, 2019; Skarbek, 2014); defiance and resentment (e.g., Beijersbergen et al., 2016; Bierie, 2013); desensitization to deterrent effects through continued exposure to severe punishment (e.g., Kleiman, 2009; Robinson and Darley, 2004); social isolation and stress (e.g., Liebling and Arnold, 2004; Phillips, 2012); loss of supportive social bonds (e.g., Crewe, 2012); stigmatization in the community (e.g., Goffman, 2014; Jones, 2018); individual-level obstacles to secure employment, financial and housing problems (e.g., Chesney-Lind and Mauer, 2003; De Giorgi, 2017; Western, 2018); and structural-level cumulative disadvantage (e.g., Clear, 2009; Kirk and Wakefield, 2018).

On the other hand, research suggests that some custodial conditions may even be conducive to desistance or rehabilitation. Examples of the potential benefits include, but are not limited to, specific deterrence due to the experience of harsh punishment (at least for some offender groups; e.g., Sherman and Berk, 1984; Weinrath and Gartrell, 2001); temporary incapacitation (e.g., Blumstein, 2005; Lee and McCrary, 2017); "growing out" of crime due to aging (e.g., Piquero et al., 2007; Sampson and Laub, 1993); reduced access to alcohol and drugs and custodial measures against addiction (e.g., Mitchell et al., 2007; Shapland et al., 2012); separation from previous gangs and deviant cliques (e.g., LaFree et al., 2018); new life chances due to education and vocational training (e.g., Bozick et al., 2018; Wilson et al., 2000); opportunities to cultivate prosocial skills in correctional treatment (e.g., Bonta and Andrews, 2024; Lipsey and Cullen, 2007); and support for resettlement and aftercare in the community (e.g., Maguire and Raynor, 2006; Markson et al., 2015).

Two practical problems beset policymakers seeking to balance these competing claims. The first practical problem is that both sets of claims are partially confounded—and besides, they may be relevant for only some groups under specific conditions

(for differentiated reviews, see Nagin et al., 2018). For example, research that directly pits the two sets of claims against one another is legally and practically difficult because the individuals in both settings are different, especially in terms of risk and seriousness of offending. Accordingly, there is limited research on direct comparisons, and this research shows mixed results. Some studies that compared community sanctions with incarceration observe a criminogenic rather than deterrent effect of custody (e.g., Abramovaite et al., 2019; Cid, 2009; Jolliffe and Hedderman, 2015). Thorough reviews of the literature encompassing various research designs affirm incarceration's null or mildly criminogenic effect (Durlauf and Nagin, 2011; Nagin et al., 2009; Wermink et al., 2010). There is, however, some apprehension about the methodological foundations on which those conclusions rest. Most of the empirical studies relied on quasi-experimental designs of varying levels of quality. Randomized controlled trials are rare. A meta-analysis of studies on custodial versus noncustodial sanctions showed that a majority favored noncustodial sanctions (Villettaz et al., 2015) but there was no significant difference in mean recidivism rates in the few experiments contained in the sample.

True experiments that compare the effects of custodial versus noncustodial sanctions are impracticable. It is rarely possible to allocate people convicted of serious and violent offenses randomly to one condition or the other. People sentenced to prison are not comparable to those who receive a community sanction in their risk of reoffending and harm. Therefore, most experimental and quasi-experimental studies focus on less serious offenders or first-time prison sentences. Studies that use propensity score matching or other matching procedures aim for the equivalence of custodial and noncustodial groups by eliminating cases at the upper end of offense seriousness. This is only one of the various problems of these matching procedures (e.g., Jann, 2017; King and Nielsen, 2019; Lösel et al., 2020; Luellen et al., 2005). A focus on shorter prison sentences in the respective comparisons is likely to inflate reoffending rates. International data show that recidivism rates for shorter sentences are typically higher and often less harmful than for longer sentences (e.g., Jehle et al., 2016; Mews et al., 2015; Nagin et al., 2009; Uhrig and Atherton, 2020). Notwithstanding the foregoing concerns, the brevity of short sentences often forestalls rehabilitative programming in practice.

Policymakers weighing competing claims from prior correctional research also confront a second practical problem, namely, in relating those insights to certain pressing policy imperatives. For instance, effect heterogeneity bedevils the policy-maker entrusted with balancing a policy's *performance* (such as by calling for correctional treatments that maximize effect size) with its *precision* (such as by calling for correctional treatments that maximize consistent and predictable effects). Where effect magnitude and variance point in different policy directions, as we find below, the stakes of striking such a balance press even more acutely. It is therefore important to extend research on those characteristics known to reduce reoffending best, such as treatment modality, offender type, fidelity, and evaluation design (Lösel, 1995, 2012), to generate policy-informing insights about which treatment settings promise which outcomes.

Criminologists are aware of how some characteristics known to reduce reoffending may sit uneasily alongside one another. We extend and develop that preoccupation by concentrating on the moderator effects of the treatment setting. For example, correctional research consistently indicates that positive outcomes follow from maintaining treatment

fidelity, tailoring treatment to participants' specific needs, and mimicking contexts where (especially cognitive-behavioral) lessons will ultimately be put to use. But although all three can promote reductions in reoffending, nonetheless different treatment settings may pit those characteristics at cross purposes. On one hand, community settings can accommodate flexible treatment delivery in ways that promote sensitivity and responsiveness; on the other hand, custodial settings can contain and control variation in ways that promote monitoring and fidelity. A strong program of correctional research should distinguish between these policy desiderata and should also clarify for policy-makers how treatment characteristics and settings relate to either one.

A sharper focus is therefore warranted on the specific association between correctional treatment's effects and the setting in which that treatment is delivered. As mentioned above, for the same reasons that preclude a direct comparison of penal sanctions in different settings, so too a direct comparison of a given treatment's effectiveness delivered in different settings is by extension impracticable as well. Treatments delivered inside a custodial facility tend to be delivered to different people presenting different profiles of risk and need than in the community; the content of treatment between these settings differs; the staff administering those treatments in custody are differently positioned and trained than those in the community; and besides, the very character and quality of treatment itself carries different meaning inside a closed facility than outside one. Methodological wisdom therefore counsels against comparing treatment's effects between settings directly.

But it would be unwise to dismiss outright prior meta-analytic research on offender rehabilitation, which can contribute to an understanding of correctional treatment's promise in different settings. Doing so requires proceeding with clear eyes about what the data can and cannot say. That is to say that prior meta-analytic research *has* imagined treatments delivered in both settings as susceptible to coherent and meaningful summaries. Likewise, moderator analysis can usefully tease apart custodial treatment's effect *compared to treatment-as-usual in custody* on one hand, and community-based treatment's effect *compared to treatment-as-usual in the community* on the other. Although *treatment-as-usual* is a complex concept that suggests an "absence" of intervention which is hard to square with the realities of criminal justice contact, nonetheless investigating the relationship between treatment setting and effectiveness in this way promises valuable insights into the conditions under which settings might relate to furthering rehabilitative ends.

To ensure a broad database, we extend a rich criminological precedent that leverages prior meta-analyses of correctional research in the form of a "meta-synthesis" (e.g., Farrington et al., 2017; Lipsey and Cullen, 2007; Lösel, 1995; Weisburd et al., 2017; Wilson, 2016). To do so, we integrate meta-analytic findings from moderator analyses on correctional treatment's effects across custodial and noncustodial settings. Moreover, we explore the extent to which findings varied across four different groups, namely young people convicted of general (often violent) crimes, adults, individuals specifically convicted of sex offenses, and people presenting drug misuse or mental health problems. Synthesizing meta-analyses carries tradeoffs: the advantage is that a "meta-meta-analysis" accounts for outcomes that differ not only between primary studies but also between meta-analyses on the same or a similar topic. This approach also bestows the advantage of drawing on visible and accessible meta-analytic research

that is less vulnerable to threats associated with searching for unpublished and “gray” scholarship. The disadvantage of “meta-meta-analysis,” however, analogizes at a higher level of abstraction to the challenge that researchers confront when conducting meta-analysis; namely, that low descriptive validity in the sampled studies can obscure important features such as overlap between studies or information about causal processes that might further open evaluation’s black box.

Method

Search of studies: We conducted an electronic search of meta-analyses on correctional treatment, correctional rehabilitation, and offender treatment. In addition, we carried out a bibliographic search of previous reviews of systematic reviews (e.g., Caudy et al., 2013; Gill, 2016; Grietens and Hellinckx, 2004; Lipsey and Cullen, 2007; Lösel, 1995; McGuire, 2002; Wilson, 2016). Although low descriptive validity in most meta-analyses impeded thorough scrutiny of their primary study samples, we excluded meta-analyses if the sample of primary studies was comprehensively incorporated into another more recent meta-analysis. For example, Andrews et al. (1990) analyzed a sample of studies that incorporated and expanded upon the sample used in Whitehead and Lab (1989); we therefore excluded the latter to minimize double-counting primary studies. However, where studies analyzed overlapping but distinct samples, we included both (e.g., Koehler et al., 2013; Redondo et al., 1999). While this necessarily introduces some issues of “double-counting” studies, the differences in eligibility criteria, coding, national background, study dates and other variations between meta-analyses justified an examination of the different findings observed in each meta-analysis. For example, close scrutiny of a selection of meta-analyses that most likely threatened to introduce double-counting revealed limited overlap: there was an overlap of 6 evaluations out of 58 between Redondo et al. (1999) and Koehler et al. (2013); there was an overlap of 21 evaluations out of 73 between Lösel and Schmucker (2005) and Schmucker and Lösel (2017); and there was an overlap of 0 evaluations out of 49 between Tong and Farrington (2006) and Ferguson and Wormith (2013).

Types of programs: We focused on meta-analyses that had synthesized the effects of evaluations of correctional treatment on recidivism. The primary comparison of interest was the effectiveness of correctional treatment delivered in community settings as opposed to that delivered in custodial settings. Therefore, the study had to report comparative effects, typically as an outcome of a moderator analysis, between community and institutional settings. Study authors operationalized these categories differently, with varying terminology. We coded as “community” any treatments that were delivered exclusively or primarily in nonsecure settings, such as when the treatment context was labeled “ambulatory,” “probation,” or “parole.” We coded as “institutional” any treatments that were delivered exclusively or primarily in secure settings, such as when the treatment context was labeled “custodial,” “prison,” or “jail.” The terms “residential” and “forensic clinic” were used ambiguously, as were outdated terms like “intramural” and “extramural,” so we collaboratively coded those only after thoroughly examining how authors applied them to the respective meta-analysis. Meta-analysis study authors rarely described in detail coding practices concerning the setting of treatments in their

primary study samples. Our use of terms such as “institutional” or “community” may therefore be necessarily somewhat constrained by a presumption of homogeneity between study authors’ uses of those terms.

To maximize generalizability, we included meta-analyses that reported no moderator analysis of the treatment setting but the sample of primary studies comprised evaluations of treatments that were delivered exclusively in the community (e.g., Visher et al., 2005) or exclusively in an institutional setting (e.g., Wilson et al., 2000). These studies, however, were included only if clear analogs existed for the assessed treatment in the opposite setting. Therefore, we excluded meta-analyses if the sample included exclusively evaluations of drug courts, boot camps, or intensive probation supervision. We found no eligible meta-analyses or systematic reviews that compared the effectiveness of domestic violence perpetrator programs between custodial and community settings. The results of the meta-analyses that contained only custodial or noncustodial programs appear separately.

We grouped studies according to five broad categories. These included meta-analyses of treatments applied to juvenile and young offenders below the age of 25, adult offenders, sex offenders, drug-involved offenders, and offenders with mental illness. Meta-analytic outcomes for studies dealing with the first four sets of offenders appear in the tables; we present results from the few meta-analyses on mentally disordered offenders in the text alone. To enhance comparability of effects, outcomes of studies for young and adult offenders were disaggregated by treatment approach when possible. Most study authors presented overall data for comparisons between community and custody-based treatment, aggregated across all treatment modalities. Those studies appear as “mixed treatments.” Results were analyzed separately when study authors provided disaggregated information for “cognitive-behavioral and behavioral treatments,” or “nonbehavioral treatments.” Cognitive-behavioral and behavioral treatments were understood to include thinking skills programs and treatments based on reinforcement of behavioral change, and nonbehavioral treatments were understood to include counseling and psychodynamic treatments, in addition to vocational training and educational curricula. Because we synthesized meta-analyses, we had to rely on the categories and coding contained in the selected studies and—also for reasons of parsimony—could not address other moderators. For example, a comparison between mandatory and voluntary treatment may also have been relevant, but this was neither our focus nor reliably assessable across different meta-analyses.

Methodological quality: Prior correctional research indicates a non-linear relationship between research design and effect size (e.g., Wilson, 2016). We therefore proceeded cautiously with a two-step analytic strategy: first, we were lenient with respect to the methodological rigor of primary studies that a meta-analysis could include in its sample. Since standard reporting practices in this area of research rely on the Maryland Scale of Methodological Rigor (Farrington et al., 2003), we included syntheses that contained primary studies at Level 2 (i.e., doubtful equivalence of comparison groups) and above. To ensure transparency we report the Maryland Scale level for each meta-analysis in our study, while also being careful not to take such scores at face value. Nonetheless, to mitigate risks the first step might have been introduced, as a second step we compared meta-analytic results from the full sample to results from a

further meta-analysis of a subsample that included meta-analyses with primary studies only at Level 3 or above. Those sensitivity checks did not alter the results we describe below. We therefore report the results from the full analysis.

Reoffending outcome: We placed no restrictions on the measure of reoffending: We included meta-analyses that reported outcomes relating to rearrest, reconviction, reincarceration, or self-reported reoffending. Study authors reported effect sizes differentiated by outcome type too rarely to support further sensitivity analysis on this measure; we therefore relied on the integrated summary effects available to us. Further to this point, since descriptive validity is an important correlate of effect size and reproducibility (Farrington, 2003; Lösel, 2018; Weisburd et al., 2017; Wong and Bouchard, 2022), we coded the information presented in the meta-analyses as a trichotomous variable:

1. *High descriptive validity:* Studies reported an effect size and corresponding measure of precision (e.g., standard deviation, standard error, or confidence interval) for both community-based and custodial treatment outcomes. Alternatively, studies reported a coefficient comparing the effectiveness of treatment in different settings, accompanied by both a measure of precision and a means of assessing the baseline level of effectiveness.
2. *Medium descriptive validity:* Studies reported an effect size but no measure of precision, or they reported a measure of comparative effectiveness but no baseline measure of effectiveness against which such comparison could be assessed (e.g., a regression coefficient without an intercept). In rare instances (e.g., Wilson et al., 2005), secondary study authors did not conduct the moderator analysis of interest in this study, but they provided sufficient information for us to perform those calculations ourselves. Such meta-analyses were coded as having “high” descriptive validity.
3. *Low descriptive validity:* Studies provided a narrative report of the comparative effectiveness of treatments in different settings, but did not provide a discernible metric of comparative treatment success. This criterion thus meant that we included studies that may have performed the moderator analysis of interest, but the secondary study authors chose not to provide information about the calculation.

Only those secondary studies that fell within the first level of this variable permitted the meta-analysis that is the core of our study. However, we also report the outcomes of studies falling within the other two categories. Outcomes of studies with high and medium levels of descriptive validity appear in the tables; for the studies with high levels of descriptive validity, outcomes were statistically synthesized. Outcomes for those studies with low levels of descriptive validity appear in the text alone.

Effect size computation: Effect sizes presented in any format were eligible. When study authors reported both unadjusted (raw) and adjusted (e.g., for sample size bias or methodological quality) meta-analytic outcome effects we collected the adjusted outcomes. To increase comparability, we presented outcomes as odds ratio (OR) effect sizes. When studies reported a standardized mean difference (Cohen’s d or Hedges’ g)

those values were converted to a logged OR using conversion formulae found in Borenstein et al. (2009). Correlation coefficients (Pearson's r or ϕ) were converted to a standardized mean difference and then into an $\ln OR$. These were then exponentiated to yield the desired OR.

Because most meta-analyses had calculated moderator effects based on heterogeneous primary studies, fixed-effects models yield a more defensible estimate of our meta-meta-analytic effects; although fixed-effects models fail to account for the overlap between the primary study samples across meta-analyses, the summary outcomes and associated Q statistics convey the best crude estimate of the overall impression.

We convert OR into a Pearson's r to gain some purchase on the effect sizes' practical, policy-informing significance. Assuming a base rate of reoffending of 50%, for example, an r of .20 equates to a recidivism rate of 60% in the control groups and 40% in the treatment groups, which represents a reduction of 20 percentage points or 33%. We report the latter statistic. Although we follow a standard criminological convention in doing so, the practice requires rather tall assumptions about true base rates of reoffending, many of which can disperse widely beyond the assumed 50% (Prins and Reich, 2021: 589). Nonetheless, when meta-analysis authors reported the baseline rate of reoffending in primary studies, we calculated the percent reduction reoffending based on that value; otherwise, we deferred to the conventional assumption of a base rate of 50% (see Lipsey and Cullen, 2007). Although calculating the percent reduction based on this assumption is a common practice in effect size comparisons, it understates the true gains associated with the intervention whenever the base rate of specific outcome measures is low—with that said, the heuristic's policy-informing value is neither more nor less than an imperfect, if convenient, shorthand.

Results

We retrieved 53 meta-analyses and reviews of the effectiveness of correctional treatment.

Programs for juvenile and young offenders

Twenty-one meta-analyses reported outcomes on the effectiveness of programs applied to juvenile and young offenders on reoffending. The 19 studies that provided information on the effectiveness of treatment appear in Table 1.

Aggregated effects across all meta-analyses: Ten meta-analyses provided sufficient descriptive detail to allow the computation of a combined effect size across all the studies. In both community-based and custodial settings, correctional treatments reported statistically significant positive effects ($OR_{\text{Community}} = 1.35$ [$CI_{95\%}$ 1.24–1.47]; $OR_{\text{Custody}} = 1.31$ [$CI_{95\%}$ 1.24–1.38]), which equates to reductions in reoffending amounting to 15.2% for community-based treatment and 13.8% for custodial treatment (see Figure 1). The difference in treatment effectiveness between settings was not statistically significant.

Behavioral and cognitive-behavioral treatments: Nine meta-analyses comprising primary studies with some level of between-group equivalence provided information with which to compute subset analyses of the effectiveness of cognitive-behavioral

Table 1. Meta-analytic outcomes of correctional treatment on recidivism among young offenders.

Study	Community		Custody		<i>r</i> ^a [95% CI]
	SMS (<i>Q</i> _{Between})	OR [95% CI]	<i>k</i> (<i>n</i>)	OR [95% CI]	
Meta-analysis	(0.48)	1.35 [1.24–1.47]		1.31 [1.24–1.38]	
Aggregated effect across all meta-analyses					
Meta-analysis	(8.68**)	2.11 [1.66–2.70]		1.42 [1.29–1.57]	
Andrews et al. (1990) ^b	3	4.67 [2.70–8.06]	21	2.87 [1.82–4.50]	14
Lipsey and Wilson (1998)	3	2.14 (behav) 2.22 (interpers)		1.82 (behav) 2.03 (interpers)	
Tong and Farrington (2006) ^b	4	1.33 [0.33–5.93]	1	1.30 [0.85–1.98]	4
Garrido Genovés et al. (2006)	3			1.31 [1.1–1.54]	18
Armeliuss and Andraessen (2007)	3			1.47 [1.1–1.92]	12 (4441)
De Swart et al. (2012)	3 + 4	1.48 [0.85–2.57]	7 (1558)	1.84 [1.28–2.65]	17 (5420)
Ferguson and Wormith (2013) ^b	3	1.60 [0.66–3.86]	2	1.57 [0.91–2.7]	3
Koehler et al. (2013)	3	2.39 [1.48–3.87]	8	1.30 [1.06–1.60]	4
WSIPP (2014a)	3	1.71 [0.65–4.51] (ART)	4	1.71 [0.65–4.51] (ART)	4

(Continued)

Table 1. (Continued)

Study	SMS (Q_{Between})	Community		Custody		r^a [95% CI]
		OR [95% CI]	k (n)	OR [95% CI]	k (n)	
Mixed programs						
Meta-analysis	(0.32)	1.23 [1.13–1.33]		1.26 [1.19–1.34]		
Garrett (1984)	2	3.19	21	1.82	90	
Andrews et al. (1990) ^b	3	1.19 [1.01–1.40]	79	0.97 [0.65–1.45]	15	0.36
Izzo and Ross (1990)	3					
Roberts and Camasso (1991)	2	2.03	22	1.52	8	
Lipsey (1992)						
Scherrer (1994)	2	1.00	11	1.00	18	0.20
Lipsey and Wilson (1998)	3	1.99 (group couns.) 2.30 (indiv. couns.) 1.49 (education)		1.09 (group couns.) 1.31 (indiv. couns.) 1.31 (education)		
Redondo et al. (1999) ^b	2	2.14 [0.49–9.39]	1	1.9 [1.49–2.42]	5	0.08 [0.05–0.11]
Smith et al. (2002) ^b						
Latimer et al. (2003)	2	1.20	126	1.20	39	
Garrido Genovés et al. (2006)	3			1.22 [0.74–2.00]	11	
Lipsey (2009)						0.08
De Swart et al. (2012)	3 + 4	1.34 [0.58–3.08]	3 (10,060)			
James et al. (2013)	3	1.27 [1.14–1.39]	4 (1245)	1.24 [1.16–1.34]	16 (2902)	
Koehler et al. (2013)	3	1.02 [0.79–1.31]	9	0.83 [0.59–1.15]	4	

Note: ART: aggression replacement training; MTFC: multidimensional treatment foster care.

^aDenotes that all positive coefficients correspond to higher effects among community-based treatments than custodial treatments.

^bDenotes subset analysis of treatments applied exclusively to juvenile and young offenders.

† $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

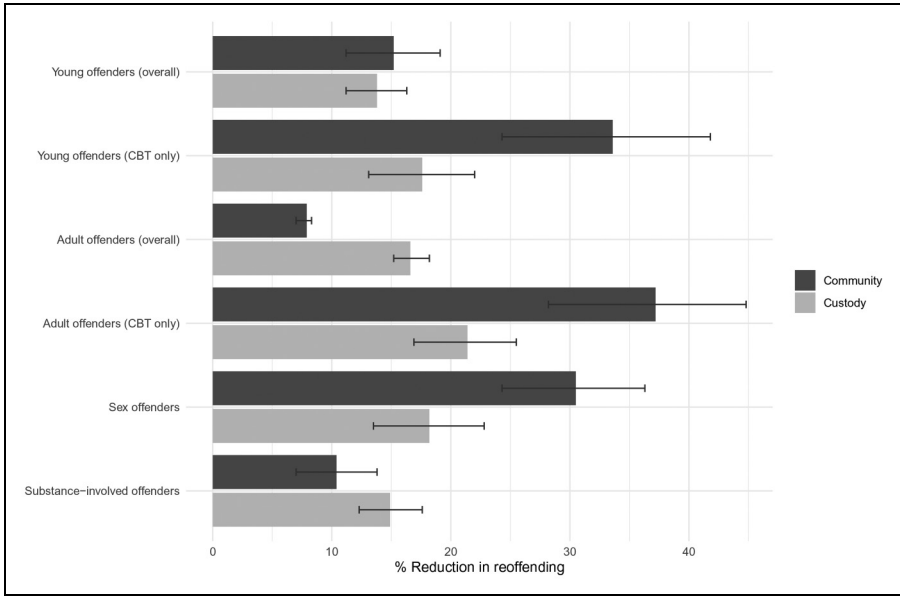


Figure 1. Synthesized meta-analytic outcomes of correctional treatment on recidivism among offender groups, expressed as percentage reduction in reoffending assuming a 50% base rate of reoffending in control groups.

and behavioral treatments. With one exception (De Swart et al., 2012), meta-analyses reported that community-based treatments performed moderately to substantially better than custodial treatments. More recent meta-analyses reported a decline, compared to older meta-analyses, in both the size of overall effects and in the observation of stronger effects for correctional treatment delivered in the community than in custody. Indeed, our subset analysis of the relevant primary studies by Andrews et al. (1990) and Koehler et al. (2013) significantly drove the observation that community-based treatment effect sizes were greater than those observed among evaluations of custodial treatment. The mean effect for cognitive-behavioral and behavioral treatments delivered in the community ($OR = 2.11 [CI_{95\%} 1.66-2.70]$) was higher than what was observed among treatments delivered in custody ($OR = 1.42 [CI_{95\%} 1.29-1.57]$). These effect sizes equate to a 33.6% reduction in reoffending among community-based treatments and a 17.6% reduction in reoffending among custodial treatments. This difference was statistically significant ($Q_{Between} = 8.68, p < .01$).

Mixed treatments: Seventeen studies provided data for the effectiveness of mixed treatment types integrating multiple modalities. Among those meta-analyses, fifteen contained sufficient descriptive validity to calculate a summary effect. With two exceptions (Latimer et al., 2003; Scherrer, 1994), all meta-analyses reported stronger effects for treatment delivered in the community than those delivered in custody. The mean effect for mixed treatments delivered in the community ($OR = 1.23 [CI_{95\%} 1.13-1.33]$) was approximately similar to what we observed among treatments delivered in custody ($OR = 1.26 [CI_{95\%}$

1.19–1.34]). These effect sizes equate to a 10.8% reduction in reoffending among community-based treatments and a 12.0% reduction in reoffending among custodial treatments. This difference was not statistically significant ($Q_{\text{Between}} = .32, p > .05$).

Evaluations of mixed program types were heterogeneous both with regard to treatment modality and methodological rigor. That heterogeneity calls for further caution in interpreting treatment effects in different settings. For example, in a subset analysis of only those evaluations with higher levels of methodological quality, Garrett (1984) observed that positive mean effects among community-based treatments disappeared (dropping from $OR = 3.19$ to a null effect of $.98$), while modest positive effects for institutional treatments persisted (dropping from $OR = 1.82$ to 1.57). We observed the same decline in treatment effects in our subset analysis of behavioral and cognitive-behavioral treatments in Andrews et al. (1990), which had reported the highest mean effect for CBT treatments in both settings. There, the difference in effects between the two settings was attenuated when the analysis was limited to only those evaluations with high methodological quality ($OR_{\text{Community}}$ dropped from 4.67 to 2.63 [$CI_{95\%}$ 1.37 – 5.26]; OR_{Custody} dropped from 2.87 to 2.86 [$CI_{95\%}$ 1.69 – 4.54]).

The heterogeneity of treatment types is also substantively meaningful when interpreting the findings. For example, James et al. (2013) analyzed the effectiveness of aftercare programs applied to offenders released from institutions. It therefore does not provide a pure comparison between community-based and custodial treatments, as even the participants who completed community-based treatments may have experienced the effects of institutional confinement. They attribute the similar outcomes they observed between settings to the possibility that youths negatively associate aftercare with the contaminated environment of a correctional facility.

Two meta-analyses (Antonowicz and Ross, 1994; Knorth et al., 2008) provided insufficient information with which to report the relative effectiveness of treatments delivered in different settings; they therefore do not appear in Table 1. Knorth et al. (2008) analyzed four quasi-experimental studies that evaluated treatments delivered in residential settings, and observed ORs ranging from 1.00 to 2.92 , although the authors highlighted the heterogeneity among the comparison groups in their sample. Antonowicz and Ross (1994) observed positive treatment effects in both community and custodial settings among 44 “rigorously controlled” primary studies, and remarked that treatments delivered in an institution ‘can be effective ... if they somehow escape from or diminish the usual prison ambience and create an “alternative community” within the institution.’ (Antonowicz and Ross, 1994: 101).

Programs for adult offenders

Twenty-two meta-analyses reported the effectiveness of correctional treatment applied to adult offenders. Twenty-one of those studies contained sufficient descriptive validity to appear in Table 2.

Aggregated effects across all meta-analyses: Seventeen meta-analyses provided sufficient descriptive detail to allow the computation of a combined effect size across all the studies. In both community-based and custodial settings, correctional treatments reported statistically significant positive effects ($OR_{\text{Community}} = 1.16$ [$CI_{95\%}$ 1.14 – 1.17];

Table 2. Meta-analytic outcomes of correctional treatment on recidivism among adult offenders.

Study	SMS (Q_{Between})	Community		Custody		r^a [95% CI]
		OR [95% CI]	k (n)	OR [95% CI]	k (n)	
Meta-analysis	(103.13***)	Aggregated effect across all meta-analyses		1.39 [1.35–1.44]		
		1.16 [1.14–1.17]				
Meta-analysis	(9.50**)	Behavioral and cognitive-behavioral programs		1.55 [1.40–1.71]		
Andrews et al. (1990) ^b	3	2.34 [1.83–2.99]	5	1.55 [1.40–1.71]		
Dowden et al. (2003)		4.59 [2.12–9.92]	11	1.31	29	
Wilson et al. (2005)	2	1.34	7	1.87 [1.48–2.35]	10	
Tong and Farrington (2006) ^b	4	1.95 [1.10–3.45]	9	1.27 [1.10–1.46]	12	
Lipsey et al. (2007)	3	1.56 [1.01–2.41]	31		27	0.20
Ferguson and Wormith (2013) ^b	3	2.97 [2.02–4.39]	11	1.84 [1.56–2.18]	21	

(Continued)

Table 2. (Continued)

Study	SMS (Q_{Between})	Community		Custody		r^a [95% CI]
		OR [95% CI]	k (n)	OR [95% CI]	k (n)	
Meta-analysis						
Mixed programs						
Lösel and Köferl (1989)	(46.47***)	1.16 [1.14–1.17]		1.38 [1.33–1.43]		
Andrews et al. (1990) ^b	3	1.42 [0.80–2.51]	12	1.49 [1.31–1.70]	16	
Redondo et al. (1999) ^b	2	1.56 [1.30–1.86]	12	0.67 [0.45–1.02]	5	
Wilson et al. (2000)	3			1.44 [1.30–1.59]	11	
Smith et al. (2002) ^b	4	1.06 [1.02–1.14]	10	1.52 [1.37–1.69]	53	.03 [0.02–0.04]
Visher et al. (2005)	5					
Villetaz et al. (2006)						
Parhar et al. (2008)	3	1.47 [1.36–1.59]	68 (28,987)	1.15 [1.08–1.23]	52 (59,930)	.02 [–.02–.06]
Davis et al. (2013)	3	1.25 [0.86–1.81]	42	1.47 [1.28–1.67]	27	
Wartna et al. (2013)	3			0.89 [0.57–1.39]	24	
Usher and Stewart (2014)				1.54 [1.34–1.78]	21	
Ndrecka (2014)		1.16 [1.14–1.17]	30 (273,204)	1.55 [1.34–1.80]	5 (2388)	
WSIPP (2014b)	3	1.14 [1.02–1.28] (training)	16	1.51 [1.30–1.75] (vocational education)	3	
				1.54 [1.24–1.92] (education)	11	
Gobeil et al. (2016)	3	1.25 [1.01–1.55]	8	1.19 [0.75–1.86]	8	
Gannon et al. (2019)	2	1.54 [1.27–1.89]	16	1.47 [1.23–1.79]	20	
Papalia et al. (2019)	3	1.28 [1.02–1.61]	3	1.70 [1.43–2.00]	12	

Note: ^aDenotes that all positive coefficients correspond to higher effects among community-based treatments than custodial treatments.

^bDenotes subset analysis of treatments applied exclusively to adult offenders.

[†] $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

$OR_{\text{Custody}} = 1.39$ [$CI_{95\%}$ 1.35–1.44]), which equates to reductions in reoffending amounting to 7.9% for community-based treatment and 16.6% for custodial treatment. The difference in treatment effectiveness between settings was statistically significant ($Q_{\text{Between}} = 103.13$, $p < .001$).

Behavioral and cognitive-behavioral treatments: Four meta-analyses comprising primary studies with some level of between-group equivalence provided information with which to compute subset analyses of the effectiveness of cognitive-behavioral and behavioral treatments. All reported that community-based treatments performed moderately to substantially better than custodial treatments. The mean effect for cognitive-behavioral and behavioral treatments delivered in the community ($OR = 2.34$ [$CI_{95\%}$ 1.83–2.99]) was higher than what was observed among treatments delivered in custody ($OR = 1.55$ [$CI_{95\%}$ 1.40–1.71]). These effect sizes are equivalent to a 37.2% reduction in reoffending among community-based treatments and a 21.4% reduction in reoffending among custodial treatments. This difference was statistically significant ($Q_{\text{Between}} = 9.50$, $p < .01$).

WSIPP (2014a) did not report outcomes in a manner that could be tabulated in Table 2. The authors noted that CBT programs delivered in an institutional setting performed better than those delivered in the community, although this relationship was not statistically significant ($p = .57$).

Mixed treatments: The mean effect for mixed treatments delivered in the community ($OR = 1.16$ [$CI_{95\%}$ 1.14–1.17]) was lower than what was observed among treatments delivered in custody ($OR = 1.38$ [$CI_{95\%}$ 1.33–1.43]). These effect sizes are equivalent to a 7.9% reduction in reoffending among community-based treatments and a 16.3% reduction in reoffending among custodial treatments. This difference was statistically significant ($Q_{\text{Between}} = 46.47$, $p < .001$).

Redondo et al. (1999) observed smaller effects in adult prisons ($OR = 1.35$ [$CI_{95\%}$ 1.21–1.50]; $k = 7$) than in psychiatric units ($OR = 2.48$ [$CI_{95\%}$ 1.81–3.40]; $k = 4$), but the combined effect for both of those institutional settings remained lower than the mean effect for community-based treatments. Parhar et al. (2008) analyzed the effect of different types of participant recruitment to treatment, and observed that treatment effectiveness was significantly higher in community-based settings than in custodial settings when comparisons were isolated to mandated, coerced, and voluntary settings. They observed significant main effects between community and institutional treatments, and between voluntary and mandated treatments, but did not observe an interaction of setting by level of coercion.

Five meta-analyses provided the outcomes of mixed treatment types in different settings. Both Smith et al. (2002) and Villetaz et al. (2006) compared the effects on adults of an institutional sentence and a community-based sanction. The effects were small in Smith et al. (2002) and Villetaz et al. (2006) failed to reject the null hypothesis.

Programs for sex offenders

We located ten meta-analyses that provided outcomes of treatment applied to sex offenders. Table 3 shows that with few exceptions (Grønnerød et al., 2015; Hanson et al.,

Table 3. Meta-analytic outcomes of correctional treatment on recidivism among sex offenders.

Study	SMS (Q_{Between})	Community		Custody	
		OR [95% CI]	k (n)	OR [95% CI]	k (n)
Meta-analysis					
(9.47**)		Aggregated effect across all meta-analyses		1.44 [1.30–1.60]	
1.94 [1.66–2.28]					
Hanson et al. (2002)	2	2.99 [1.84–4.88]	8 (1187)	1.44 [1.10–1.90]	7 (1829)
Lösel and Schmucker (2005)	2	1.93 [1.35–2.77]	27	1.16 [0.84–1.60]	21
Reitzel and Carbonell (2006)	2	4.13 [1.34–12.75]	3	2.37 [1.80–3.12]	3
Hanson et al. (2009)	2	1.87 [0.92–3.81]	6 (1270)	1.31 [0.88–1.96]	7 (3531)
WSIPP (2014b)	3	1.63 [0.77–3.45]	7	1.33 [1.02–1.72]	8
Grønnerød et al. (2015)	2	3.95 [1.58–9.90]	1 (126)	1.00 [0.64–1.56]	11 (1576)
Schmucker and Lösel (2017)	3	1.83 [1.12–2.98]	12	1.18 [0.72–1.94]	7
Soldino and Carbonell-Vaya (2017)	2	1.89 [1.01–3.54]	6	1.39 [1.07–1.81]	8
Gannon et al. (2019)	2	1.64 [1.27–2.13]	18 (11,296)	1.49 [1.18–1.92]	25 (14,224)

Note: † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

2009; WSIPP, 2014b), all meta-analytic summary effects were statistically significant, positive, and reported superior effects for correctional treatment delivered in the community than in custody. All nine studies with high levels of descriptive validity reported higher mean effects for treatments delivered in the community ($OR = 1.94$ [$CI_{95\%}$ 1.66–2.28]) than for treatments delivered in custodial settings ($OR = 1.44$ [$CI_{95\%}$ 1.30–1.60]). This difference was statistically significant ($Q_{\text{Between}} = 9.47$, $p < .01$). These effect sizes are equivalent to a 30.5% reduction in reoffending among community-based treatments and an 18.2% reduction in reoffending among custodial treatments, assuming 50% baseline reoffending in the control groups. However, the baseline rate of reoffending among sex offenders is typically far lower; consequently, true reduction in reoffending attributable to sex offender treatment is likely higher.

Three studies (Lösel and Schmucker, 2005; Reitzel and Carbonell, 2006; Schmucker and Lösel, 2015) provided mean effects for treatments delivered in mixed settings such as hospitals; in all three studies the OR for treatments delivered in hospitals fell in between what was observed for treatments delivered in community and custodial settings ($OR = 1.37$ [$CI_{95\%}$.78–2.41], $k = 10$ in Lösel and Schmucker, 2005; $OR = 3.28$ [$CI_{95\%}$ 1.13–9.57], $k = 3$ in Reitzel and Carbonell, 2006; $OR = 1.77$ [$CI_{95\%}$ 1.00–3.14], $k = 5$ in Schmucker and Lösel, 2015). Only Polizzi et al. (1999) had low descriptive validity and therefore does not appear in Table 3. While they did not provide the outcomes of moderator analyses in their report, they observed generally higher effects among evaluations of nonprison-based treatments than for prison-based treatments; however, the low methodological quality of the prison-based treatment evaluations attenuated the authors' confidence in those outcomes.

Programs for substance-involved offenders

Six meta-analyses provided information on the effectiveness of treatments applied to substance-involved offenders. Table 4 shows that the mean effects across all the meta-analyses showed a statistically significant positive effect for treatment delivered in the community ($OR = 1.22$ [$CI_{95\%}$ 1.14–1.31]) that was lower than the effect for treatment delivered in custody ($OR = 1.34$ [$CI_{95\%}$ 1.27–1.42]), which equates to reductions in reoffending amounting to 10.4% for community-based treatment and 14.9% for custodial treatment. This difference was statistically significant ($Q_{\text{Between}} = 4.14$, $p < .05$).

The Q_{Between} analysis' sensitivity to model specification attests both to imprecision and inconsistency in the relative effectiveness of correctional treatments between settings. On one hand, two meta-analyses reported superior effects in the community. Koehler et al. (2014) observed greater effects for treatments delivered in the community than among those delivered in custody; however, they noted that the effectiveness of custodial treatment was driven in significant part by an evaluation of a program delivered with low levels of fidelity whose interpretation warranted caution. WSIPP (2014b) compared the effectiveness of therapeutic communities on substance-involved adults in custodial and community-based settings and observed slightly larger effects in the community. The community-based evaluations principally comprised studies that measured outcomes after varying periods of release from the institutional environment of the therapeutic

Table 4. Meta-analytic outcomes of correctional treatment on recidivism among substance-involved offenders.

Study	SMS (Q_{Between})	Community	Custody		
		OR [95% CI]	k	OR [95% CI]	k
Aggregated effect across all meta-analyses					
Meta-analysis	(4.14*)	1.22 [1.14–1.31]		1.34 [1.27–1.42]	
Perry et al. (2006)	5	1.22 [0.65–2.26]	4	1.52 [0.87–2.63]	2
Mitchell et al. (2007)	3			1.37 [1.24–1.51]	66
Holloway et al. (2008)	2	1.28 [1.11–1.47]	18	1.44 [1.24–1.66]	19
Koehler et al. (2014)	2	2.53 [1.75–3.67]	13	1.72 [1.29–2.29]	2
WSIPP (2014a)	3	1.09 [0.83–1.42] (Non-TC)	6	1.12 [0.85–1.46] (TC)	4
WSIPP (2014b)	3	1.15 [0.98–1.35] (nonintensive)	4	1.37 [1.16–1.62] (nonintensive)	8
		1.09 [0.95–1.25] (intensive)	5	1.37 [1.13–1.66] (intensive)	6
		1.31 [1.11–1.53] (TC)	8	1.24 [1.12–1.38] (TC)	18

Note: Non-TC: treatments other than therapeutic communities; TC: therapeutic communities.

† $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

community. WSIPP (2014b) also compared the effectiveness of nonintensive outpatient treatments and intensive in-patient treatments applied to substance-involved offenders in both settings. For treatments of both levels of intensity, they reported slightly higher effects for custodial settings than for community-based settings, although neither of the community-based treatments was statistically significant.

On the other hand, the remaining meta-analyses reported superior effects when treatment was delivered in custody. Perry et al. (2006) included the most rigorously controlled studies in their meta-analysis, and observed a statistically nonsignificant effect for treatment in both settings. However, there was likely a confounding effect of treatment type with the setting, as the studies in their custody-based sample both evaluated therapeutic communities, whereas the studies in their community-based sample evaluated the effectiveness of intensive supervision-based programs. WSIPP (2014a) also compared therapeutic communities in custodial settings to other community-based treatments applied to juveniles alone. They observed statistically non-significant effects in both settings and a smaller difference between settings than was found in Perry et al. (2006).

Programs for offenders with mental illness

Two meta-analyses reported the effectiveness of treatments for offenders with mental illness and were therefore not tabulated. Martin et al. (2012) synthesized 25 studies

with minimum levels of internal validity equivalent to SMS Level 2, with a total sample size of 15,678. The authors reported roughly equivalent mean effects for evaluations of treatments delivered in the community ($OR = 1.39$ [$CI_{95\%}$ 1.27–1.52]; $k = 22$) as among those delivered in custody ($OR = 1.41$ [$CI_{95\%}$ 1.14–1.75]; $k = 6$). They observed substantially higher effects for treatments delivered in settings that mixed both institutional and community components ($OR = 1.82$ [$CI_{95\%}$ 1.49–2.18]; $k = 9$). WSIPP (2014a) located an additional study that was published after Martin et al. (2012), which found slightly higher effects in the community ($OR = 1.52$ [$CI_{95\%}$.88–2.60]) than in custody ($OR = 1.46$ [$CI_{95\%}$.85–2.50]); adding this to Martin et al. (2012) had negligible effect in distinguishing the effectiveness of treatment between different settings ($Q_{\text{Between}} = .02$; $p > .05$).

Discussion

We synthesized the effects of correctional treatment on reoffending and further conducted moderator analysis differentiating outcomes between treatments delivered in custody or in the community. In doing so, we drew together meta-analytic evidence on the effects of correctional treatments delivered to a wide variety of offender groups. We find that correctional treatments in both settings display moderate to substantial positive effects on reducing reoffending, and that those positive effects hold when treatments are delivered to different groups. We also find that correctional treatments delivered in community settings display stronger effects on reoffending than treatments delivered in custodial settings. However, only rarely did meta-analytic evidence point to a criminogenic effect of correctional treatment in either setting, and we observed no statistically significant evidence of criminogenic effects from custodial treatments.

Our review of meta-analyses points to two further differentiations. First, differences in treatment content and participant type moderated the *magnitude* of correctional treatment's relative effectiveness between settings. In particular, treatments delivered in the community outperformed those delivered in custody to the greatest extent when the treatment was narrowly tailored to a specific population, and when the treatment applied behavioral and cognitive-behavioral modalities. The corollary of this point is that between-setting moderator effects are most obscure at the highest levels of generality and heterogeneity. Second, differences in treatment content and participant type moderated the *precision* of correctional treatment's effectiveness in either setting. In particular, although effects in custodial settings were consistent, effects for community-based treatment varied widely. Taken together, those differentiations yield the following insight: *community-based correctional treatments outperform custodial correctional treatments in a general sense, but custodial settings may provide opportunities to monitor and regiment treatment delivery in ways that can both (i) improve upon and (ii) homogenize treatment outcomes.*

Low descriptive validity in primary evaluations of correctional treatment resulted in meta-analyses containing limited information that might support inferences about what explains effect heterogeneity. By extension, so too the low descriptive validity that prevailed in those meta-analyses limited our efforts to theorize mechanisms as well. Of course, the theoretical expectations we outline warrant further research that can test those theoretical expectations more directly. With that said prior correctional research

buttresses the key insight we extract about effects between settings. For instance, correctional experts and practitioners commonly advise that treatment conditions in the community are more heterogeneous and difficult to monitor than in closed institutions. Those features may lead to more frequent dropouts that are, in turn, associated with more recidivism (e.g., Carl et al., 2019; McMurrin and Theodosi, 2007; Olver et al., 2011). The heterogeneous outcomes we observe in our review lend further credence to the challenges that heterogeneous treatment settings present. It follows that the heterogeneity of treatment settings in the community—by design, more so than in custodial facilities—not only introduces heterogeneity of treatment effects, but moreover that this heterogeneity in turn may relate to the flexibility of treatment delivery that is associated with large effects.

Comparison of effects between offender populations in Figure 1 bears out this theoretical expectation. The results support the interpretation that tight monitoring conditions in custody may allow for more consistent treatment delivery administered to people with substance-involvement offenses, much more so than the relatively heterogeneous treatment settings common in community-based treatments delivered to analogous populations. For example, whereas people with substance-involvement problems often drop out early from treatment in the community, treatment participants in custody who may present motivation problems—even if only temporarily so—cannot simply withdraw or disappear. In contrast, in the context of treatments delivered to people convicted of sex offenses, the tradeoffs of different settings may instead point in a rather different direction. There, the highly regimented and tightly controlled contexts in which custody-based treatments are delivered are, among other benefits, likely to enhance treatment fidelity and may also minimize participant attrition. However, those benefits are likely to be outweighed by the costs of delivering treatment in a context that underprepares participants for the skills that effective treatments teach. In particular, correctional treatments in custody least approximate those risky and stressful situations—such as those with potential child or adult victims—in which participants can best develop and refine effective self-control strategies and skills. Consequently, well-monitored community-based treatment seems to hold particular promise for relapse prevention in the context of people convicted of sex offenses.

Two further patterns emerge from the data that merit attention, although they likewise warrant further research to test more directly the theoretical expectations we generate below. First, effect sizes in more recent meta-analyses were noticeably smaller than in older meta-analyses, and second, the size of the difference in effect sizes between settings was smaller in more recent meta-analyses than in older meta-analyses. A pessimistic interpretation might hold that the effectiveness of correctional treatment may be eroding. Indeed, there may be good reason to believe that hardships that beset many justice systems, from government austerity to corroded capacity to arbitrary control, could contribute to a depletion of the care that justice systems can accommodate. This depletion would be consistent with criminological research that decries the restriction of investment in services from criminal justice to welfare and beyond (e.g., Braithwaite, 2022; Reiner, 2020: Ch. 3; Wacquant, 2001; but see Phelps, 2011). We can also speculate about why such hardships might asymmetrically befall the delivery of correctional treatments delivered in one setting as opposed to another.

Although both patterns warrant attention in future research, we also suggest another possible explanation here. Declining effect sizes might not be attributable to a

deterioration in either the effectiveness of correctional treatments as a whole or in the relative features of correctional treatment settings in particular. Rather, the slight gradual decline in the observed effectiveness of correctional treatments may instead be a function of an improvement in the treatment-as-usual against which correctional programs are measured. If so, then more recent declines in effect sizes may instead be attributable to *improvements in the control conditions*. Although the control group conditions in program evaluations are rarely investigated in detail, they seem to be as important for the outcomes as the treatment content.

Conclusion

The analysis in this article affirms the promise of correctional treatments delivered in both custodial and community-based settings, while at the same time, it complicates simplistic interpretations of “superiority” of treatment delivered in one setting as opposed to the other. We instead find that a moderator analysis of prior meta-analytic research surfaces an important and policy-relevant finding that is sensitive to *both* effect size *and* precision. A tradeoff runs through political questions about which of those desiderata to prioritize on one hand, and scientific questions about their respective determinants on the other. Prior research hints at clues for what might determine especially large effect sizes in one setting and especially homogeneous effects in another; however, a thorough appraisal of those determinants pushes criminologists beyond the study of isolated programs to differentiate the effects of treatment content, implementation context, offender characteristics, and methodological features of evaluation designs (Lösel, 2012). Within this framework, our study underlines the need for differentiated answers to the question of what works for whom, under what conditions, with regard to what outcomes, and why.


Acknowledgments

We thank Xenia Below, Renan Araújo, and Lan Yin for excellent research assistance, Keith Humphreys and Tobias Smith for helpful comments and feedback, and the librarians at the Cambridge University Library for their assistance in securing necessary materials. Support from the London School of Economics and Political Science and the University of Erlangen-Nuremberg made this research possible.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Johann Koehler  <https://orcid.org/0000-0002-1305-891X>

References

- Abramovaite J, Bandyopadhyay S, Bhattacharya S, et al. (2019) Alternatives to custody: Evidence from police force areas in England and Wales. *British Journal of Criminology* 59(4): 800–822.

- Andrews DA, Zinger I, Hoge RD, et al. (1990) Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology* 28(3): 369–404.
- Antonowicz DH and Ross RR (1994) Essential components of successful rehabilitation programs for offenders. *International Journal of Offender Therapy & Comparative Criminology* 38(2): 97–104.
- Armeliuŝ BÅ and Andreassen TH (2007) Cognitive-behavioral treatment for antisocial behavior in youth in residential treatment. *Cochrane Database of Systematic Reviews* 2007(4): CD005650.
- Bayer P, Hjalmarsson R and Pozen D (2009) Building criminal capital behind bars: Peer effects in juvenile corrections. *The Quarterly Journal of Economics* 124(1): 105–147.
- Beijersbergen KA, Dirkzwager AJ and Nieuwbeerta P (2016) Reoffending after release: Does procedural justice during imprisonment matter? *Criminal Justice and Behavior* 43(1): 63–82.
- Bierie D (2013) Procedural justice and prison violence: Examining complaints among federal inmates (2000–2007). *Psychology, Public Policy, and Law* 19(1): 15–29.
- Blumstein A (2005) An overview of the symposium and some next steps. *Annals of the American Academy of Political and Social Science* 602: 242–258.
- Bonta J and Andrews DA (2024) *The Psychology of Criminal Conduct* (7th ed.). New York: Routledge.
- Borenstein M, Cooper H, Hedges L, et al. (2009) Effect sizes for continuous data. *The Handbook of Research Synthesis and Meta-Analysis* 2: 221–235.
- Bozick R, Steele J, Davis L, et al. (2018) Does providing inmates with education improve post-release outcomes? A meta-analysis of correctional education programs in the United States. *Journal of Experimental Criminology* 14: 389–428.
- Braithwaite J (2022) *Macrocriminology and Freedom*. Canberra, Australia: ANU Press.
- Carl L, Schmucker M and Lösel F (2019) Predicting attrition and engagement in the treatment of young offenders. *International Journal of Offender Therapy and Comparative Criminology* 13: 371–379.
- Caudy MS, Tang L, Ainsworth SA, et al. (2013) Reducing recidivism through correctional programming: Using meta-analysis to inform the RNR simulation tool. In: *Simulation Strategies to Reduce Recidivism: Risk Need Responsivity (RNR) Modeling for the Criminal Justice System*. New York: Springer, 167–193.
- Chesney-Lind M and Mauer M (eds.) (2003) *Invisible Punishment: The Collateral Consequences of Mass Imprisonment*. New York: New Press.
- Cid J (2009) Is imprisonment criminogenic? A comparative study of recidivism rates between prison and suspended prison sanctions. *European Journal of Criminology* 6(6): 459–480.
- Clear TR (2009) *Imprisoning Communities: How Mass Incarceration Makes Disadvantaged Neighborhoods Worse*. New York: Oxford University Press.
- Crewe B (2012) *The Prisoner Society: Power, Adaptation and Social Life in an English Prison*. Oxford, UK: Oxford University Press.
- Damm AP and Gorinas C (2020) Prison as a criminal school: Peer effects and criminal learning behind bars. *The Journal of Law and Economics* 63(1): 149–180.
- Davis LM, Bozick R, Steele JL, et al. (2013) *Evaluating the Effectiveness of Correctional Education: A Meta-analysis of Programs That Provide Education to Incarcerated Adults*. Santa Monica, CA: RAND Corporation.
- De Giorgi A (2017) Back to nothing: Prisoner reentry and neoliberal neglect. *Social Justice* 44(1 (147)): 83–120.
- De Swart JJ, Van den Broek H, Stams GJ, et al. (2012) The effectiveness of institutional youth care over the past three decades: A meta-analysis. *Children & Youth Services Review* 34(9): 1818–1824.

- Dowden C, Antonowicz D and Andrews DA (2003) The effectiveness of relapse prevention with offenders: A meta-analysis. *International Journal of Offender Therapy & Comparative Criminology* 47(5): 516–528.
- Durlauf SN and Nagin DS (2011) Imprisonment and crime: Can both be reduced? *Criminology & Public Policy* 10(1): 13–54.
- Farrington DP (2003) Methodological quality standards for evaluation research. *The Annals of the American Academy of Political and Social Science* 587(1): 49–68.
- Farrington DP, Gaffney H and Ttofi MM (2017) Systematic reviews of explanatory risk factors for violence, offending, and delinquency. *Aggression and Violent Behavior* 33: 24–36.
- Farrington DP, Gottfredson DC, Sherman LW, et al. (2002) The maryland scientific methods scale. In: Sherman LW, Farrington DP, Welsh BC, et al. (eds) *Evidence-Based Crime Prevention*. London: Routledge, 13–21.
- Ferguson LM and Wormith JS (2013) A meta-analysis of moral reconnection therapy. *International Journal of Offender Therapy and Comparative Criminology* 57(9): 1076–1106.
- Gannon TA, Olver ME, Mallion JS, et al. (2019) Does specialized psychological treatment for offending reduce recidivism? A meta-analysis examining staff and program variables as predictors of treatment effectiveness. *Clinical Psychology Review* 73: 101752.
- Garrett CJ (1984) *Meta-analysis of the effects of institutional and community residential treatment on adjudicated delinquents*. Unpublished dissertation, University of Colorado, Boulder.
- Garrido Genovés V, Anyela Morales L and Sánchez-Meca J (2006) What works for serious juvenile offenders? A systematic review. *Psicothema* 18(3): 611–619.
- Gill C (2016) *Community interventions. What works in crime prevention and rehabilitation: Lessons from systematic reviews: 77–109*.
- Gobeil R, Blanchette K and Stewart L (2016) A meta-analytic review of correctional interventions for women offenders: Gender-neutral versus gender-informed approaches. *Criminal Justice and Behavior* 43(3): 301–322.
- Goffman A (2014) *On the Run: Fugitive Life in an American City*. Chicago: University of Chicago Press.
- Grietens H and Hellinckx W (2004) Evaluating effects of residential treatment for juvenile offenders by statistical metaanalysis: A review. *Aggression and Violent Behavior* 9(4): 401–415.
- Grønnerød C, Grønnerød JS and Grøndahl P (2015) Psychological treatment of sexual offenders against children: A meta-analytic review of treatment outcome studies. *Trauma, Violence, & Abuse* 16(3): 280–290.
- Hanson RK, Gordon A, Harris AJR, et al. (2002) First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research & Treatment* 14(2): 169–194.
- Hanson RK, Bourgon G, Helmus L, et al. (2009) The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice & Behavior* 36(9): 865–891.
- Holloway K, Bennett TH and Farrington DP (2008) *The effectiveness of treatment in reducing drug-related crime*. Stockholm: Swedish National Council for Crime Prevention (Brottsförebyggande rådet – Brå).
- Izzo R and Ross R (1990) Meta-analysis of rehabilitation programs for juvenile delinquents: A brief report. *Criminal Justice & Behavior* 17(1): 134–142.
- James C, Stams GJ, Asscher JJ, et al. (2013) Aftercare programs for reducing recidivism among juvenile and young adult offenders: A meta-analytic review. *Clinical Psychology Review* 33(2): 263–274.
- Jann B (2017) Why propensity scores should be used for matching. Paper presented at the German Stata Users Group Meeting, Berlin, Germany. https://www.stata.com/meeting/germany17/slides/Germany17_Jann.pdf.

- Jehle J-M, Albrecht H-J, Hohmann, et al. (2016) *Legalbewährung nach strafrechtlichen Sanktionen: Eine bundesweite Rückfalluntersuchung 2010 bis 2013*. Berlin: German Federal Ministry of Justice.
- Jolliffe D and Hedderman C (2015) Investigating the impact of custody on reoffending using propensity score matching. *Crime & Delinquency* 61(8): 1051–1077.
- Jones N (2018) *The Chosen Ones: Black Men and the Politics of Redemption*. Berkeley, CA: University of California Press.
- King G and Nielsen R (2019) Why propensity scores should not be used for matching. *Political Analysis* 27(4): 435–454.
- Kirk DS and Wakefield S (2018) Collateral consequences of punishment: A critical review and path forward. *Annual Review of Criminology* 1: 171–194.
- Kleiman MA (2009) *When Brute Force Fails: How to Have Less Crime and Less Punishment*. Princeton, NJ: Princeton University Press.
- Knorth EJ, Harder AT, Zandberg T, et al. (2008) Under one roof: A review and selective meta-analysis on the outcomes of residential child and youth care. *Children & Youth Services Review* 30(2): 123–140.
- Koehler JA, Lösel F, Akoensi TD, et al. (2013) A systematic review and meta-analysis on the effects of young offender treatment programs in Europe. *Journal of Experimental Criminology* 9(1): 19–43.
- Koehler JA, Humphreys DK and Akoensi TD (2014) A systematic review and meta-analysis on the effects of European drug treatment programmes on reoffending. *Psychology, Crime & Law* 20(6): 584–602.
- LaFree G, Jensen MA, James PA, et al. (2018) Correlates of violent political extremism in the United States. *Criminology; An Interdisciplinary Journal* 56(2): 233–268.
- Latimer J, Dowden C and Morton-Bourgon KE (2003) *Treating youth in conflict with the law: A new meta-analysis*. Ottawa, Canada: Department of Justice, Youth Justice Research.
- Lee DS and McCrary J (2017) The deterrence effect of prison: Dynamic theory and evidence. *Advances in Econometrics* 38: 73–146.
- Liebling A and Arnold H (2004) *Prisons and Their Moral Performance: A Study of Values, Quality, and Prison life*. Clarendon Studies in Criminology. Oxford, UK: Oxford University Press.
- Lipsey MW (1992) Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In: Cook TD, Cooper H and Cordray DS, et al. (eds), *Meta-Analysis for Explanation: A Casebook*. New York, NY: Russell Sage, 83–128.
- Lipsey M (2009) The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders* 4: 124–147.
- Lipsey MW and Cullen FT (2007) The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science* 3: 297–320.
- Lipsey MW, Landenberger NA and Wilson SJ (2007) Effects of cognitive-behavioral programs for criminal offenders. *Campbell Systematic Reviews* 2007: 6.
- Lipsey MW and Wilson DB (1998) Effective intervention for serious juvenile offenders. In: Loeber R and Farrington DP (eds) *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions*. Thousand Oaks, CA: Sage, 313–345.
- Lösel F (1995) The efficacy of correctional treatment: A review and synthesis of meta-evaluations.
- Lösel F (2012) Offender treatment and rehabilitation: What works. In: *The Oxford Handbook of Criminology* (5th ed.). Oxford, UK: Oxford University Press, pp. 986–1016.
- Lösel F (2018) Evidence comes by replication, but needs differentiation: The reproducibility issue in science and its relevance for criminology. *Journal of Experimental Criminology* 14(3): 257–278.

- Lösel F and Köferl P (1989) Evaluation research on correctional treatment in West Germany: A meta-analysis. In: Wegener H, Lösel F and Haisch J (eds) *Criminal Behaviour and The Justice System: Psychological Perspectives*. New York, NY: Springer, 334–355.
- Lösel F and Schmucker M (2005) The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology* 1(1): 117–146.
- Lösel F, Link E, Schmucker M, et al. (2020) On the effectiveness of sexual offender treatment in prisons: A comparison of two different evaluation designs in routine practice. *Sexual Abuse* 32(4): 452–475.
- Luellen JK, Shadish WR and Clark MH (2005) Propensity scores: An introduction and experimental test. *Evaluation Review* 29: 530–558.
- Maguire M and Raynor P (2006) How the resettlement of prisoners promotes desistance from crime: Or does it? *Criminology and Criminal Justice* 6: 19–38.
- Markson L, Lösel F, Souza K, et al. (2015) Male prisoners' family relationships and resilience in resettlement. *Criminology and Criminal Justice* 15: 423–441.
- Martin MS, Dorken SK, Wamboldt AD, et al. (2012) Stopping the revolving door: A meta-analysis on the effectiveness of interventions for criminally involved individuals with major mental disorders. *Law & Human Behavior*. DOI: 10.1007/s10979-011-9274-4.
- McGuire J (2002) Integrating findings from research reviews. In: McGuire J (ed) *Offender Rehabilitation and Treatment*. Chichester: Wiley, 3–38.
- McMurrin M and Theodosi E (2007) Is treatment non-completion associated with increased reconviction over no treatment? *Psychology, Crime & Law* 13: 333–343.
- Mews A, Hillier J, McHugh M, et al. (2015) *The Impact of Short Custodial Sentences, Community Orders and Suspended Sentence Orders on Re-offending*. London: Ministry of Justice Analytical Series.
- Mitchell O, Wilson DB and Mackenzie DL (2007) Does incarceration-based drug treatment reduce recidivism? A meta-analytic synthesis of the research. *Journal of Experimental Criminology* 3(4): 353–375.
- Nagin DS, Cullen FT and Jonson CL (2009) Imprisonment and reoffending. *Crime and Justice* 38(1): 115–200.
- Nagin DS, Cullen FT and Lero-Jonson C (eds.) (2018) *Deterrence, Choice, and Crime: Contemporary Perspectives*. New York: Routledge.
- Ndrecka M (2014) *The impact of reentry programs on recidivism: A meta-analysis*. Doctoral dissertation, University of Cincinnati.
- Olver ME, Stockdale KC and Wormith JS (2011) A meta-analysis of predictors of offender treatment attrition and its relationship to recidivism. *Journal of Consulting and Clinical Psychology* 79: 6–21.
- Ouss A (2011) Prison as a school of crime: Evidence from cell-level interactions. Available at SSRN 1989803.
- Papalia N, Spivak B, Daffern M, et al. (2019) A meta-analytic review of the efficacy of psychological treatments for violent offenders in correctional and forensic mental health settings. *Clinical Psychology: Science and Practice* 26(2): e12282.
- Parhar KK, Wormith JS, Derksen DM, et al. (2008) Offender coercion in treatment: A meta-analysis of effectiveness. *Criminal Justice & Behavior* 35(9): 1109–1135.
- Perry A, Coulton S, Glanville J, et al. (2006) Interventions for drug-using offenders in the courts, secure establishments and the community. *Cochrane Database of Systematic Reviews*: 3. DOI: 10.1002/14651858.CD005193.pub2.
- Phelps MS (2011) Rehabilitation in the punitive era: The gap between rhetoric and reality in US prison programs. *Law & Society Review* 45(1): 33–68.
- Phillips C (2012) *The Multicultural Prison: Ethnicity, Masculinity, and Social Relations among Prisoners*. Oxford, UK: Oxford University Press.

- Piquero AR, Farrington D and Blumstein A (2007) *Key Issues in Criminal Career Research: New Analyses of the Cambridge Study in Delinquent Development*. New York: Cambridge University Press.
- Polizzi DM, MacKenzie DL and Hickman LJ (1999) What works in adult sex offender treatment? A review of prison- and non-prison-based treatment programs. *International Journal of Offender Therapy & Comparative Criminology* 43(3): 357–374.
- Prins SJ and Reich A (2021) Criminogenic risk assessment: A meta-review and critical analysis. *Punishment & Society* 23(4): 578–604.
- Pyrooz DC and Decker SH (2019) *Competing for Control: Gangs and the Social Order of Prisons*. New York: Cambridge University Press.
- Redondo S, Sánchez-Meca J and Garrido V (1999) The influence of treatment programmes on the recidivism of juvenile and adult offenders: An European meta-analytic review. *Psychology, Crime & Law* 5(3): 251–278.
- Reiner R (2020) *Social Democratic Criminology*. Abingdon, UK: Routledge.
- Reitzel LR and Carbonell JL (2006) The effectiveness of sexual offender treatment for juveniles as measured by recidivism: A meta-analysis. *Sexual Abuse: A Journal of Research and Treatment* 18(4): 401–421.
- Roberts AR and Camasso MJ (1991) The effect of juvenile offender treatment programs on recidivism: A meta-analysis of 46 studies. *Notre Dame Journal of Law, Ethics, & Public Policy* 5: 421–442.
- Robinson PH and Darley JM (2004) Does criminal law deter? A behavioural science investigation. *Oxford Journal of Legal Studies* 24(2): 173–205.
- Sampson RJ and Laub JH (1993) *Crime in the Making: Pathways and Turning Points through Life*. Cambridge, MA: Harvard University Press.
- Scherrer JL (1994) *A meta-analysis of the effectiveness of residential treatment programs for children and adolescents*. Unpublished dissertation, University of Illinois, Chicago.
- Schmucker M and Lösel F (2015) The effects of sexual offender treatment on recidivism: An international meta-analysis of sound quality evaluations. *Journal of Experimental Criminology* 11(4): 597–630.
- Schmucker M and Lösel F (2017) Sexual offender treatment for reducing recidivism among convicted sex offenders: A systematic review and meta-analysis. *Campbell Systematic Reviews* 13(1): 1–75.
- Shapland J, Bottoms A and Muir G (2012) Perceptions of the criminal justice system among young adult would-be desisters. In: Lösel F, Bottoms AE and Farrington DP (eds) *Young Adult Offenders: Lost in Transition?* Abingdon, UK: Routledge, 128–145.
- Sherman L and Berk R (1984) The specific deterrent effects of arrest for domestic assault. *American Sociological Review* 49(2): 261–272.
- Skarbak D (2014) *The Social Order of the Underworld: How Prison Gangs Govern the American Penal System*. New York: Oxford University Press.
- Smith P, Goggin C and Gendreau P (2002) *The effects of prison sentences and intermediate sanctions on recidivism: General effects and individual differences*. Rep. JS42–103/2002. Ottawa, CA: Solicitor General.
- Soldino V and Carbonell-Vayá EJ (2017) Effect of treatment on sex offenders' recidivism: A meta-analysis. *Anales De Psicología/Annals of Psychology* 33(3): 578–588.
- Stevenson M (2017) Breaking bad: Mechanisms of social influence and the path to criminality in juvenile jails. *Review of Economics and Statistics* 99(5): 824–838.
- Tong LSJ and Farrington DP (2006) How effective is the Reasoning and Rehabilitation programme in reducing offending? A meta-analysis of evaluations in four countries. *Psychology, Crime & Law* 12(1): 3–24.

- Uhrig SCN and Atherton K (2020) *Reoffending Following Custodial Sentences or Community Orders, by Offence Seriousness and Offender Characteristics, 2000–2018*. London: UK Ministry of Justice Analytical Series.
- Usher AM and Stewart LA (2014) Effectiveness of correctional programs with ethnically diverse offenders: A meta-analytic study. *International Journal of Offender Therapy and Comparative Criminology* 58(2): 209–230.
- Villetaz P, Killias M and Zoder I (2006) The effects of custodial vs. non-custodial sentences on re-offending: A systematic review of the state of knowledge. *Campbell Systematic Reviews* 2006: 13.
- Villetaz P, Gillieron G and Killias M (2015) The effects on re-offending of custodial vs. non-custodial sanctions: An updated systematic review of the state of knowledge. *Campbell Systematic Reviews* 11(1): 1–92.
- Visher CA, Winterfield L and Coggeshall MB (2005) Ex-offender employment programs and recidivism: A meta-analysis. *Journal of Experimental Criminology* 1(3): 295–316.
- Wacquant L (2001) Deadly symbiosis: When ghetto and prison meet and mesh. *Punishment & Society* 3(1): 95–133.
- Wartna B, Alberda D and Verweij S (2013) Een meta-analyse van Nederlands recidiveonderzoek naar de effecten van strafrechtelijke interventies. *Tijdschrift voor Criminologie* 55(1): 3–23.
- Weinrath M and Gartrell J (2001) Specific deterrence and sentence lengths: The case of drunk drivers. *Journal of Contemporary Criminal Justice* 17: 105–122.
- Weisburd D, Farrington DP and Gill C (2017) What works in crime prevention and rehabilitation: An assessment of systematic reviews. *Criminology & Public Policy* 16(2): 415–449.
- Wermink H, Blokland A, Nieuwbeerta P, et al. (2010) Comparing the effects of community service and short-term imprisonment on recidivism: A matched samples approach. *Journal of Experimental Criminology* 6(3): 325–349.
- Western B (2018) *Homeward: Life in the Year after Prison*. New York: Russell Sage Foundation.
- Whitehead JT and Lab SP (1989) A meta-analysis of juvenile correctional treatment. *Journal of Research in Crime & Delinquency* 26(3): 276–295.
- Wilson DB (2016) Correctional programs. In: Weisburd D, Farrington DP and Gill C (eds) *What Works in Crime Prevention and Rehabilitation: Lessons from Systematic Reviews*. New York: Springer, 193–217.
- Wilson DB, Bouffard LA and MacKenzie DL (2005) A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. *Criminal Justice & Behavior* 32(2): 172–204.
- Wilson DB, Gallagher CA and MacKenzie DL (2000) A meta-analysis of corrections-based education, vocation, and work programs for adult offenders. *Journal of Research in Crime & Delinquency* 37(4): 347–368.
- Wong JS and Bouchard J (2022) Do meta-analyses of intervention/prevention programs in the field of criminology meet the tests of transparency and reproducibility? *Trauma, Violence, & Abuse*: 24(3): 1522–1542.
- WSIPP (Washington State Institute for Public Policy) (2014a) *Benefit-Cost Results: Juvenile Justice*. Olympia, WA: WSIPP.
- WSIPP (Washington State Institute for Public Policy) (2014b) *Benefit-Cost Results: Adult Criminal Justice*. Olympia, WA: WSIPP.