



Article

Still Asking ‘What Works’: A Punishment Question for the Ages or an Aging Punishment Question?

Karol Lucken

Department of Criminal Justice, University of Central Florida, Orlando, FL 32816, USA; karol.lucken@ucf.edu

Abstract: The expectation that punishment be effective at controlling crime is a longstanding convention in the U.S., and no doubt elsewhere. While the history of American punishment has not been shaped entirely by the question of efficacy, it has played a predominant role in justifying penal policy for over 200 years. The question has become even more salient in policy decision-making of late, as research has begun to certify and consolidate findings on what is effective at reducing recidivism. What is lacking in this ongoing conversation, however, is a critique of this penal policy question and the answers it generates in the form of recidivism rates. The current paper fills this void by interrogating the claims of the evaluation literature, namely that better proof of what is effective is available and that more research is still needed. The questions and findings of 19th, 20th, and 21st Century seekers of what is effective in the American adult penal system are recounted and analyzed using several data sources. They include government reports, professional association meeting minutes, legislative documents, scholarly reports, individual studies, research reviews, and statistical analysis of systematic reviews. Ultimately, an overarching narrative is provided that deepens and challenges our understanding of what is known about what is effective.

Keywords: recidivism; punishment; history; evaluation; corrections; effectiveness



Citation: Lucken, K. Still Asking ‘What Works’: A Punishment Question for the Ages or an Aging Punishment Question? *Histories* **2021**, *1*, 53–68. <https://doi.org/10.3390/histories1020010>

Academic Editor: Jon Mathieu

Received: 24 February 2021

Accepted: 19 May 2021

Published: 28 May 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

A mainstay of the American mindset in criminal justice, and no doubt elsewhere, is that sanctions imposed by the state should and can be effective at controlling crime. Whether sanctions are rationalized through the logic of deterrence, rehabilitation, or incapacitation, their value has been chiefly assessed by their capacity to reduce crime. A predictable outgrowth of this crime reduction expectation has been a longstanding and vigorous pursuit of what is effective by criminal justice professionals, criminological scholars, and lawmakers alike.

While the history of American punishment has not been driven solely by the question of effectiveness, its role in justifying penal reform and existing programs and policies has been indispensable [1]. For example, under the weight of fiscal constraint brought on by the 2008 economic recession, nearly every state in the U.S. was forced to adopt or seriously consider adopting new policies and alternatives to incarceration [2,3]. Helping to guide and defend this policy transition was research identifying penal interventions that worked. That is, the interventions reduced or at least did not increase offender recidivism.

Research in this vein has received heightened attention of late under what has been called the evidence-based movement or evidenced-based corrections [4–6]. An extension of this movement is the promotion and dissemination of findings from evaluation studies on the effectiveness of certain interventions. For example, national and international platforms (e.g., The National Registry of Evidence-Based Programs and Practices, CrimeSolutions.gov, and The Campbell Collaboration) distill information for government officials and other interested parties on interventions that are working, not working, or promising based on their impact on crime and recidivism.

What is missing from the ongoing conversation about what is effective in the domain of offender sanctioning is a critical exposition of this penal policy question and the answers

it generates in the form of recidivism rates. Though some dialogue on this front has emerged in the academic literature, challenges of this kind have been more intermittent and exhortatory than empirical and systematic. Nevertheless, a common theme in these challenges is apparent, namely that a program or policy's worth should be judged on more than just crime reduction [7–11]. For example, Lee and Stohr [8] contend that certain correctional programs have been unfairly denounced as “quackery” due to the overemphasis on recidivism as a measure of penal success. Klingele [12] has opined that the pragmatic focus of the evidence-based literature has suppressed a broader discussion about correctional values apart from utility (e.g., fairness, dignity, or proportionality). Notably, Brown [1] characterized the “what works” pursuit as a “paradoxical journey of faith and futility”.

This paper builds on these challenges to the dominant evaluation motif by interrogating the standing claims of the effectiveness literature. These claims are that better knowledge (proof) of what reduces recidivism is increasingly available and that more research on this front is needed. In as much as there is an evolutionary component to these claims, it is worthwhile to consider them not just through statistical analysis, but through the longer lens of history. Studies and findings on the effectiveness of adult penal interventions in the U.S. are therefore situated within a historical framework, stretching from the early 19th Century to the present. Importantly, this paper is not seeking to distinguish what types of programs or strategies are most effective (i.e., those based in rehabilitation, deterrence, or incapacitation), nor is it providing a history of a particular penal philosophy (i.e., rehabilitation, deterrence, or incapacitation). Rather, the paper is offering a historical glimpse into the repeated application of a question that has been of predominant interest to American policymakers and a large segment of the community of criminological scholars. Ultimately, the functionality of this question and its recidivism answer is considered, along with alternative penal narratives that might prove more ‘useful’ going forward.

2. Materials and Methods

The current paper provides a sequential account of the questions and findings of 19th, 20th, and 21st Century seekers of what is effective in the adult penal system in America. The account reviews the numerous interventions to which this question has been applied and the methodologies, recidivism findings, and general conclusions associated with the asking of this question. These historical data are derived from a vast repository of published sources that show the actors, milestones, and indeed the controversies and uncertainties that have characterized this pursuit from its beginning. These sources include government reports, professional association meeting minutes, and legislative records, as well as scholarly reports, studies, and research literature reviews.

Under the general outlines of history, this question and its recidivism answers are analyzed using a narrative and statistical format. This combined approach enables assessment of the dual claims of the effectiveness literature indicated above. Given the voluminous literature that has been amassed over time, the analysis records in broad chronological strokes a cross-section of authoritative mainstream and signature findings. Therefore, the findings presented are neither exhaustive nor perfectly linear. However, they are couched within a deep understanding of the life of this pursuit.

3. Results

3.1. Effectiveness and Recidivism: Early 19th to Early 20th Century

The question of punishment being effective was put to official and systematic use at the beginning of the 19th Century. At this time, the loss of liberty through extended periods of confinement in prison had just replaced the physical and public punishments of Colonial America. The newly elevated question of efficacy acquired even greater importance when the first [makeshift] prisons became the ‘well-ordered’ institutions known as penitentiaries.

Throughout and beyond this institutional make-over, reformers and officials intended the penitentiary to be a humanitarian advancement over pre-modern forms of sanctioning

and the first prison facilities. However, their interest in whether a more humane and morally healthy system of punishment had evolved in the U.S. was not a purely benevolent one. It was one equally motivated by the utilitarian goal of a more effective system of punishment.

The humane treatment of inmates and conditions of confinement in the penitentiary were important factors under the theory that the “interior” of the inmate’s “new home” bore some relationship to recidivism [13]. This theorized relationship meant that the features of the institutional environment and the physical and moral health of the inmate were routinely documented by state government agencies. Official state reports documented the mental and physical status of inmates, their food type and allotment, plumbing and sanitation conditions, inmate work productivity, and inmate mortality rates and causes [14]. The moral health of the inmate was monitored through the proper provision of religious and intellectual instruction. For example, in an 1846 legislative report of the Joint Committee on the State Prison of New Jersey, inspectors remarked that “The first step to reform in every penitentiary must be the formation of industrious habits;” this reform of the inmate would be cultivated by the “imparting of wholesome advice and instruction” [15] (p. 384).

Methods of disciplining inmates for infractions were also judged with an eye toward their effect on post-prison behavior. In a Massachusetts legislative debate, an elected official posed the following question to correctional authorities: “. . . . How far does your discipline serve to make rogues in prison better, and rogues out of prison fewer?” [16] (p. 93). This question was prompted by data showing a 37 percent recidivism rate among released offenders. It was reported that “of the twenty-seven convicts released by pardon or remission of sentence, ten are known to have led a vicious life afterwards” [16] (p. 93).

In the attempt to ascertain effectiveness, the state of New York also compiled data on prison recommitments and the commission of new offenses by previously incarcerated offenders. In one analysis performed by state officials, recommitment rates for prisoners at Sing Sing were additionally examined by punishment severity (i.e., length of incarceration/time served), thus providing a crude empirical test of deterrence theory; this theory encompasses the idea that sanction severity is inversely related to crime. Based on figures for the years 1817 to 1842, the recommitment ratio, regardless of time served (1 year to Life), was variable, ranging from a high of 1 in 3 (33 percent) to a low of 1 in 17 (6 percent). The median rate of return to prison during this twenty-five-year period was 1 in 6 or 17 percent [17].

New Jersey studied effectiveness through these same means, but also factored in the role of institutional and offender characteristics in recidivism. The state’s governor observed that some inmates had two to four prior convictions, a fact prompting his request for information on the type of penitentiary to which they had been previously exposed. To determine the effectiveness of the two competing penitentiary models, the governor sought to compare the recommitment rates of those that had been previously confined in the “separate system” (continuous solitude) with those that had been confined in the “congregate system” (solitude except during work assignments) [15]. All in all, the state’s analysis of recidivism and recommitment rates considered both the prior records of the current inmates and their prior incarceration experience.

The commitment to the study of punishment’s efficacy proceeded in earnest in the decades that followed. This commitment was captured in a publication entitled “Prison Discipline as a Science.” The following excerpt addresses methods of prison discipline and how best to adjudicate what is effective in dealing with inmate behavior:

“Prison discipline must be studied as a science, with a sagacious, intelligent, and dispassionate examination of the facts and laws upon which it is based. It does not belong altogether within the realm of sentiment. It is rather within the domain of calm inquiry and the most careful experiment” [18] (p. 72).

The workhouse, an earlier custodial setting for the able-bodied poor, was also evaluated for effectiveness. One assessment of the workhouse recognized the relationship between the duration of exposure to this setting and effectiveness. This factor is referred to

in the current evaluation literature as treatment dosage. In a statewide Pennsylvania report, for example, it was observed that the average stay of less than 30 days in a workhouse was insufficient to affect behavioral change. It was suggested that the effectiveness of the workhouse should not be judged prematurely as “you can’t diagnose mentality and morality and administer treatment in thirty days” [19] (p. 37).

Elected officials and philanthropic reformers were not alone in their pursuit of an effective strategy. The new growing class of criminal justice professionals/professional penologists and scholars were equal participants in this endeavor. For example, in his presidential address at the annual meeting of the American Prison Association, Joseph Byers presented an informative set of statistics. He noted that in 1914, 23,303 individuals were under sentence in Massachusetts and that 57.6 percent were recidivists with an average of 7.5 prior convictions [20]. He further reported that of the 110,816 total recorded sentences, recidivists were responsible for 100,950 of them. Byers subsequently concluded that 57 percent of convicted criminals were responsible for 91 percent of the crime.

Employing the same logic that justified selective incapacitation in the 1980s, namely the use of lengthier prison sentences for those identified as high-risk and/or repetitive offenders, Byers proposed that these “feeble-minded” recidivists be targeted for special intervention. He reasoned that “science is telling us how to detect and experience is teaching us how to care for the feeble-minded” [20]. Yet, Byers also remarked that the research was too unsettled in its results and advised caution in policy until evidence was more conclusive.

At the turn of the 20th Century, the national social movement of progressivism inspired substantial reforms of the punishment system. State and federal reforms now operated under the philosophy of rehabilitation, which assumes that crime’s causes can be found in one’s (biological, psychological, or sociological) life history and ultimately cured through careful scientific study and tailored interventions. Two signature reforms to emerge from this era were the community-based sanctions of probation (alternative to prison) and parole (post-prison supervision).

With the addition of these penal options, the study of punishment’s effectiveness naturally widened in scope. For example, in Philadelphia, a 1915 report on probation showed the anticipated outcome of low recidivism rates. The study indicated that of the 1520 adult males placed on probation over three years, 6 percent qualified as failures [21]. A 1907 Massachusetts report calculated probation’s efficacy by highlighting cost savings, an outcome that would clearly be undermined by probationer recidivism. The report stated:

“[There has been a decrease of 1478 in the average prison population between 1897, when it reached its highest point, and 1906]”. A proportionate increase in the prison population would have added 1238 to the average, bringing the prison population up to 8978, and the decrease would have been 2716. The cost of food and clothing is somewhat in excess of a dollar a week. An estimate far below the actual cost of food and clothing shows a saving of at least \$150,000 a year. This reduction of the prison population, and consequent saving in cost of supporting prisoners, has been due almost wholly to the growth of the probation system.” [22] (p. 746).

Parole’s value as a penal strategy was also assessed through the outcome of recidivism. The Eastern State Penitentiary in Pennsylvania reported that of the 295 inmates paroled between September 1910 and December 1912, 63 cases (or 21 percent) were apparent failures [23]. The Western State Penitentiary reported that of the 276 paroled since the inception of parole, 223 (or 85.8) percent were successfully reporting at the time of the study [22]. Reports indicating low failure/recidivism rates were also compiled in Delaware, Illinois, Massachusetts, Washington, and New York [23].

The study of parole recidivism rates expanded throughout the 1930s with the formal adoption of parole in more states and the establishment of the Federal Bureau of Prisons in 1930. Recidivism rates for the federal system indicated that “not more than ten percent of discharged prisoners violate regulations during the period of parole” [24] (p. 70). Studies of the Uniform Crime Reports also showed that, generally, less than one percent of all men

being arrested were on parole at the time of arrest [24]. In the state of Washington, for example, of the 90,504 arrested in the first three months of 1935, 509 were on parole [24]. Recidivism data were also examined by offender/offense type. In the state of New York, for instance, recidivism rates for sex offenders were of particular interest. Data showed that of the 925 sex offenders released on parole over a period of seven years, eight (less than one percent) had been convicted of new offenses and twenty-five were returned by authorities on the likelihood that a future offense might be committed [24].

From the inception of parole, data collection also made the important distinction between failures/violations due to new criminal behavior and failures due to non-criminal behavior or rather technical violations (e.g., unemployment or consorting with bad influences). Statistics from the Eastern State Penitentiary in Pennsylvania reported that, in 1910, roughly 28 percent of paroled prisoners were returned for non-criminal violations/non-reporting as opposed to criminal activity [25]. A 1914 report of the Acting Committee of Pennsylvania noted that generally 80 out of 100 parolees succeeded, but of those who did not, few failures could be attributed to a new criminal act. It was generally understood that it was violations of other conditions of parole that brought them “back to the prison house” [26] (p.11). These and other reports, combined with offender testimony, led officials to conclude that many offenders returned to prison voluntarily because conditions outside were worse and insurmountable [13].

3.2. Effectiveness and Recidivism: Mid-20th to Early 21st Century

The question of effectiveness relative to crime reduction and recidivism accelerated and grew more sophisticated with the solidification of sociology and criminology as academic fields. From roughly the mid-20th Century on, numerous and larger studies were commissioned by state and federal governments, with countless more being conducted through the independent investigations of scholars. For decades, research centered on the effectiveness of rehabilitation, or at least the myriad strategies of a system ostensibly governed by this objective.

In 1966, for example, Walter C. Bailey conducted a content analysis of 100 reports of empirical evaluations of correctional treatment. The reports were derived from studies published primarily between 1940 and 1960. In his review of these reports, Bailey concluded that “evidence supporting the efficacy of correctional treatment is slight, inconsistent, and of questionable reliability” [27] (p. 157). He coupled this conclusion with the observation that his findings on efficacy, or rather the lack thereof, were unremarkable. That is, they comported with other research findings of the day, including Dalton’s [28] the ineffectiveness of probation officer counseling and Kirby’s [29] on the ineffectual treatment of criminals and delinquents.

In 1971, Robison and Smith published results on the effectiveness of correctional treatment in California. Their research, which was prepared for the California State Legislature, reviewed recidivism findings from ten studies of correctional programming. Overall, five intervention categories were considered, including imprisonment versus probation, treatment program in prison, intensity of parole/probation supervision, outright release to parole or society, and length of stay in prison. The authors ultimately concluded that no evidence existed that supported claims of the superior efficacy of one strategy over another [30].

That same year produced another report on the overall state of policy evaluation and recidivism. Levin’s report [31] summarized several prior studies looking at various factors that could potentially affect recidivism [32–36]. Collectively, these studies involved data gathered between 1958 and 1966, representing nearly the whole of California, the federal prison and parole system, parts of Pennsylvania, and Florida. From this collection of research findings, Levin concluded that offenders who had received probation generally had significantly lower rates of recidivism than similar offenders who had been incarcerated, and of those who had been incarcerated, offenders with shorter terms had lower recidivism rates than those with longer terms. In the various studies and literature reviews

he examined, recidivism/failure rates for probationers ranged between 10 and 40 percent, depending on the length of the study's follow-up period. For example, in a study based on all California Department of Corrections data, the average recidivism rate for post-prison releasees tracked for three years was 33 percent. Additional research from California, Washington, and Pennsylvania showed average recidivism rates of 72, 51, and 48 percent with tracking periods of thirty-six months, six to thirty months, and twenty-eight months, respectively.

In 1975, a study commissioned by the state of New York resulted in what remains the most renowned, polemical, and perhaps consequential report produced in criminology. Lipton, Martinson, and Wilks reviewed the results of 231 studies that were published in English between 1945 and 1967. The types of programs, practices, and variables examined in this review were wide-ranging, including cosmetic surgery, in-prison treatment, individual versus group counseling, type of institutional environment, medical/hormonal treatment, length of prison sentence, intensity of probation supervision, parole, and more. Martinson summarized the study's findings in the conservative leaning publication *The Public Interest*. His article opened with the following statement:

"With few and isolated exceptions, the rehabilitative efforts that have been reported so far have had no appreciable effect on recidivism. Studies that have been done since our survey was completed do not present any major grounds for altering that original conclusion." [37] (p. 25).

He further concluded that while success or partial successes had been found in these studies, no discernable pattern existed to affirm the "efficacy of any particular method of treatment" [37] (p. 49). Martinson equally argued that "it is just possible that some of our treatment programs are working to some extent, but that our research is so bad that it is incapable of telling" [37] (p. 49). This concession on the potential of certain strategies to be effective was nevertheless succeeded by the statement that their rigorous review gave "very little reason to hope that we have in fact found a sure way of reducing recidivism through rehabilitation" [37] (p. 49).

The political embrace of a sweeping "nothing works" interpretation of the Lipton et al., conclusions was instrumental in the death of rehabilitation as the dominant penal policy in the U.S. In the wake of this fatal pronouncement, rehabilitation adherents and scholars sought to revive the expired penal philosophy. Substantial reanalysis of the Lipton et al. study was conducted alongside new studies on the effectiveness of various correctional modalities. This effort was spearheaded primarily by a cohort of Canadian psychologist/criminologists. Some re-analyses found that some rehabilitative interventions did work to reduce recidivism [21–38]. New studies and literature reviews also found that certain interventions did work better than others [39]. However, a full explication of the observed variations in lower recidivism rates across interventions was lacking. A subsequent systematic review of rehabilitative programming confirmed that within the body of positive findings, significant and unexplored variations in recidivism rates did still exist [40].

The get-tough era that prevailed between 1980 and the early 2000s also witnessed an increase in research on deterrence, namely the study of the impact of the certainty and swiftness of sanctions on crime. Here too, definitive results on efficacy were elusive [41–48]. Harkening back to the 1817 report of Sing Sing's recommitment rates by time served and studies reviewed by Levin [31] and many others [30], Orsagh and Chen [49] reexamined the hypothesis that time served affects post-prison recidivism. They found that time served does affect recidivism, although, the direction of those effects varied by the social class of the inmate. Leading scholar Daniel Nagin [46] later concluded that while evidence in favor of deterrence was firmer than it had been for two decades, large and important gaps in knowledge persisted.

The confounding relationship between deterrence-based strategies and recidivism was reaffirmed in a more recent synopsis of what is known about the effects of the certainty and severity of sanctions (i.e., deterrence theory). In this synopsis, Bushway and Paternoster [50]

(p. 144) concluded that it was “extremely difficult to determine whether instrumental goals are being met, and if met through what mechanism.” The recent study of a community supervision program known as Project HOPE appears to bear further witness to this equivocal conclusion. The initial study of HOPE [51] used a randomized design and found that swift and certain, but not severe, responses to parole violations led to fewer future violations, arrests, and reimprisonment. This celebrated finding inspired 160 program replications across the United States [52]. Two ensuing randomized studies of HOPE [53,54] involving a total of five sites found no evidence of effectiveness. These findings were succeeded by a fourth quasi-experimental study that found evidence of positive effects [55].

The get-tough on crime era also generated substantial research on the impact of incapacitation on crime, or rather the intuitive theory that crime will decrease if more offenders are simply “locked up” for longer [56–62]. This research examined the fiscal and crime reduction benefits of selective (targeting recidivists through habitual offender and three-strikes laws) and general (mass) incapacitation. Some research, including that of economists, calculated that the social and criminal justice costs associated with criminal activity in fact exceeded the costs of incarceration. This finding justified the increased use of prison and length of stay as a preemptive strike against future crime [63–65]. Other scholars also demonstrated that increases in prison populations and/or sentence length could avert or prevent the incidence of overall crime and/or certain types of property crime [60,62,66].

Additional research countered claims of a crime reduction benefit from incarceration. Greenberg [67] argued that Zedlewski’s [65] findings on incapacitation effects were grossly overstated. MacKenzie [68] found that most studies testing for incapacitation effects had shown a small crime reduction benefit, but not necessarily at less financial expense. Criticism of the incapacitation thesis by Auerhahn [69] also claimed that certain incapacitation models performed poorly upon reanalysis. Meanwhile, a study of all prisoners released in the U.S. in 1994 highlighted another real drawback to incapacitation policies, namely the eventuality of a 67 percent recidivism rate within thirty-six months of release [70].

The incapacitation thesis has since been revisited in research on criminal careers. This body of work has found indirect [71] and direct support [72,73] for the incapacitation thesis. Yet, the question of incapacitation’s effect on crime is still regarded as unresolved [74]. The open status of this question is said to hinge upon better estimates of the onset, desistance, and offending frequency patterns of criminal careers.

Near the close of the 20th Century, another landmark comprehensive study of effectiveness was conducted. In 1996, Congress stipulated that an independent study employing “rigorous and scientifically recognized standards and methodologies” for determining effectiveness be performed [75] (p. 3). That study, entitled *Preventing Crime: What Works, What Doesn’t, What’s Promising* provided a systematic review of over 500 scientific evaluations of highly diverse federally funded state and local crime prevention programs for adults and juveniles.

In a summation of the findings, it was stated that “very few operational crime prevention programs have been evaluated using scientifically recognized standards and methodologies, including repeated tests under similar and different social settings” [75] (p. 1). However, the authors did conclude that there is now “minimally adequate evidence to establish a provisional list of what works, what doesn’t, and what’s promising” [75] (p. 1).

The research on adult penal interventions in the state-of-the-art study referenced above was reproduced and expanded in the 2006 book *What Works in Corrections* by Doris MacKenzie. This book has shaped much of the prevailing and enduring wisdom on what is effective in adult punishment in the U.S. The book consolidates in one place the results of twenty-five years of this pursuit through 21 systematic reviews (SRs) involving 241 studies of adult correctional interventions. This single-source compilation of research best facilitates assessment of the claims of the effectiveness literature, which are that better

knowledge (proof) of what is effective is increasingly available and that more research is still needed.

The publication dates of the studies contained this book fall between 1976 and 2001, with the bulk of the studies being published in the 1990s. The following analysis of recidivism rates is based on 12 of the original 21 SRs (SRs for domestic violence ($n = 3$) and sex offender interventions ($n = 4$) were excluded because they involve treatment that cannot be generalized to offenders as a group. The SR for Moral Reconciliation Therapy ($n = 1$) was excluded because 11 of the 14 studies were authored by the program's patent holder [76] (p. 116, Table 7.1). The SR for Prison-Based Drug Treatment ($n = 1$) was omitted because recidivism measures and subcategory information were so varied and numerous that distinctions between the two could not be made confidently [68] (p. 258, Table 12.2.) of the studies and findings reported in these 12 SRs, only high rigor studies scoring between 3 and 5 on the Maryland Scale of Scientific Methods (MSSM) were used; these scores indicate either random assignment of intervention participants or other assurances of experimental-control group comparability. Also, only one recidivism/failure measure per reviewed study was used here, though the original SRs reported subsample findings. Thus, if recidivism rates from a reviewed study were based on a 'new conviction,' subsample information on the percentage of new convictions resulting in prison readmission were omitted; this latter percentage reflects a judicial sentencing decision rather than intervention effectiveness (i.e., what works). Similarly, if recidivism rates from a reviewed study were based on 'all arrests,' recidivism rates for subsamples based on arrests by offense type were excluded from this analysis.

Last, recidivism measures based on technical violations, revocations for non-criminal infractions, or returns to prison for non-criminal conduct were excluded from the current analysis. If the reason for the revocation/violation was not stated in the reviewed study or the outcome measure merged violations for criminal conduct with technical violations, then the reviewed study's recidivism findings were retained for this analysis. Last, studies that measured recidivism as the mean number of offenses or survival days/months were omitted because of their incomparability with percentage-based recidivism outcomes; only four studies were affected by this exclusion (1 Life Skills program study; 3 Drug Court studies).

The analysis of the 12 SRs in Table 1 consists of the following elements. (Table 1 and the corresponding analysis are modified versions of a larger analysis by this author in *Rethinking Punishment*.) In the main findings column, the lowest and highest recidivism rate recorded in the SR for that intervention type is indicated for experimental and control group samples, irrespective of findings of statistical significance in the study of origin. In this same column, an [unweighted] average recidivism rate (Avg R) is provided that is based on all the studies in that SR category. (The Avg R statistic is not weighted by the sample size of the studies. The Avg R figure should only be interpreted within the context of the range of recidivism possibilities for any given program within an intervention category.) The number of studies used in the calculation of the recidivism range and Avg R are also reported. The last column of the table reports MacKenzie's conclusions on the effectiveness of each intervention type, which were supplemented by meta-analysis findings.

Table 1. Effectiveness and Recidivism: Analysis of 12 Systematic Reviews.

Systematic Review Intervention Category	Experimental Groups Recidivism Range X = Avg R N = # Study Samples	Control Groups Recidivism Range X = Avg R N = # Study Samples	MacKenzie's Conclusions on Intervention Effectiveness
Reasoning & Rehabilitation	13–63% X = 27% N = 11	0–70% X = 36% N = 14	Yes
Cognitive Restructuring	6–50% X = 32% N = 12	24–71% X = 43% N = 12	Yes
Vocational Education	7–59% X = 31% N = 7	10–66% X = 37% N = 8	Yes
Correctional Industry	7, 29% X = 18% N = 2	10, 34% X = 22% N = 2	No
Work/multi-component	22–71% 51% N = 8	19–71% X = 51% N = 9	No
Academic Education	4–62% X = 35% N = 16	12–75% X = 44% N = 18	Yes
Life Skills	8–48% X = 28% N = 3	17–70% X = 33% N = 4	No
Drug Court	5–71% X = 32% N = 34	10–81% X = 39% N = 38	Yes
Community drug treatment	10–58% X = 27% N = 6	15–56% X = 25% N = 6	Yes
Electronic monitoring	0.01–32% X = 20% N = 8	0.08–47% X = 28% N = 11	No
Intensive supervision probation	12–58% X = 35% N = 37	4–60% X = 35% N = 41	No
Boot Camp	4–63% X = 31% N = 30	3–77% X = 33% N = 30	No

3.2.1. Recidivism Pattern #1

A dominant pattern revealed in these data is the wide recidivism range that exists for every intervention category. Of the 12 intervention categories listed, 8 have a recidivism range of roughly 50 percentage points. This range holds for experimental and control groups, regardless of program orientation (rehabilitative or punitive) or (in)effectiveness. Importantly, this wide variation is not necessarily attributable to notable differences in the offender populations studied. For example, in the SR for reasoning and rehabilitation (R & R) programs, one study found a recidivism rate (r) of 13 percent, while another, using the same measure of recidivism, found an r of 63 percent. The study yielding a 13 percent r appears to have been based on comparisons between former prisoners, whereas the study yielding a 63 percent r was based on probationers/subjects who were compared to other

prisoners/controls. Wide ranging r values (0–70%) are also reported for the various R & R program control groups.

The SR for cognitive restructuring interventions exhibits a similar pattern. Recidivism rates for the studies within this intervention category ranged between 6 and 50 percent for subjects and 24 and 71 percent for controls. In the first instance, the wide range can likely be explained by the fact that the 6 percent r was based on a population of learning-disabled offenders supervised in the community, while the 50 percent r , which was statistically significant, was based on inmates. Nevertheless, a large variation in recidivism rates for this intervention's participants remains; another study in this SR based on inmates yielded a 24 percent r , which, incidentally, was not statistically significant. Thus, a 50 percent r constitutes a finding of effectiveness for cognitive restructuring, whereas a 24 percent r for subjects and controls counts as ineffective.

The SR for vocational-educational interventions yielded a recidivism range of 7 to 59 percent for subjects and 10 to 66 percent for controls. The 7 percent r was based on a study of former federal inmates, whereas the higher r of 59 percent was based on young adult males with prior records. Though these offender groups are objectively different, each group is high risk in its own right because of their prior records and/or prior incarceration history. The remaining studies in the vocational-educational review were based on inmate populations only. Based on these comparable groups, the recidivism range remained wide at 7 to 46 percent.

The intervention category of drug courts exhibited the widest recidivism range. Recidivism findings from drug court studies were as low as 5 percent and as high as 71 percent. The two studies yielding these highly disparate rates used the same measure of recidivism (re-arrest) and involved very similar program participants. For drug court control groups, that range was equally wide at 10 to 81 percent. The SR for community-drug treatment also showed wide ranging rates of 10 to 58 percent; they were based on offender groups that were the same in terms of important factors like being substance abusers and sufficiently low-risk and/or non-violent to remain in the community. The pattern of wide-ranging recidivism rates across studies also existed for interventions deemed to be more punitive in orientation; all were deemed ineffective by the original analysis. The recidivism range for electronic monitoring was 0.01 to 32 percent for subjects and 0.08 to 47 percent for controls. The pair of studies yielding this 32-point span for subjects was each based on rearrests and had no striking characteristic differences in their study populations.

The recidivism rates produced from studies on intensive supervision probation ranged from 12 to 58 percent for subjects and 4 to 60 percent for controls. These variations occurred within the same multi-site study but were based on two different sites (Atlanta and Milwaukee). The same recidivism measure and the same type of offender populations were used in both sites. Recidivism rates in the boot camp SR exhibited the same gulf. One study of boot camps yielded an r of 4 percent, while another yielded an r of 63 percent. The control group findings displayed an equally wide recidivism range of 3 to 77 percent. The foregoing analysis of the recidivism rates contained within the 12 SR's illustrates certain curious features of "what works" in the punishment of adults. Specifically, it shows some of the practical difficulties of relying on recidivism rates in the sphere of policy decision-making. For example, if the overarching goal of a program is a sizable reduction in recidivism, then the problem exposed by this analysis is one of unpredictability in what to expect. Should a jurisdiction expect a drug court that produces a 5 percent recidivism rate or a 71 percent recidivism rate that is effective from the standpoint of statistical significance? Similarly, should a jurisdiction anticipate a vocational-education program that yields a 7 percent or a 46 percent recidivism rate?

3.2.2. Recidivism Pattern #2

The second recidivism pattern comes into focus when recidivism rates for studies within an intervention type are averaged. This calculation shows that for each intervention type deemed effective, the percentage point gap between the averaged R for subjects and

controls is in single digits in all but one case. For example, based on the SR for drug courts, an intervention deemed effective, the averaged R (based on 32 studies) was 32 percent for subjects and 39 percent for controls. Similarly, the Avg R yielded by the SR of vocational-educational programs, another intervention deemed effective, was 31 percent for subjects and 37 percent for controls (based on 19 studies). For the intervention of community drug treatment (deemed effective), the difference in the Avg R was 2 percent but in favor of the controls. The percentage point difference in the Avg R was somewhat higher for R&R, cognitive restructuring, and academic education, nevertheless, the following fact remains. For every intervention category, the recidivism range (or percentage point difference in recidivism rates) for subjects was far greater than any difference in the Avg R between subjects and controls.

When the Avg R was examined for all intervention types in the community (punitive and rehabilitative), the Avg R for ineffective programs was sometimes lower or no different than the Avg R for effective programs. For example, based on the SR of boot camps (ineffective), the Avg R for subjects was 31 percent. For EM (ineffective), the Avg R for subjects was 20 percent. Yet, the Avg R for drug court studies (effective) and community drug treatment (effective) studies were 31 and 27 percent, respectively. Although different categories of offenders populated boot camps and drug courts, one could reason that because higher risk offenders participated in boot camps (their comparison groups were prisoners) and mostly-first time non-violent offenders populate drug courts, the Avg R for boot camp studies should be higher than the Avg R for drug courts. Moreover, descriptions of program participants provided by MacKenzie indicate that offender populations involved in EM studies were not discernably different from offender populations in community drug treatment studies.

3.2.3. Recidivism Pattern #3

The final recidivism pattern underscores the recurring difficulty of distinguishing the efficacy of any method of treatment or intervention over another. An example of this can be found in the percentage difference or lack thereof between the Avg R for control groups of ineffective interventions and the Avg R for subjects of effective interventions. For example, the Avg R for controls in correctional industry (ineffective) studies was 22 percent, which is lower than all the Avg R's for subjects of every intervention type deemed effective, institution and community-based. This seems to imply that, on average, an offender coming out of prison who has received no programming at all will do better than someone coming out of prison who has received an effective intervention or who is less risky or dangerous and received an effective intervention in the community. Looking at community-based programs only, the Avg R for controls in EM (ineffective) studies was 28 percent, which is lower than the Avg R of 32 percent for subjects in drug court (effective) studies and nearly identical to the Avg R of 27 percent for subjects in community drug treatment (effective) studies. The Avg R for boot camp controls (33%) also paralleled the Avg R (32%) for subjects in drug court studies.

4. Conclusions

This paper has followed the progression of the punishment question "What is effective?" and the recidivism rates it has generated over several centuries in the U.S. This question has been placed under the historical microscope in view of the foundational and categorical, yet still tentative, claims of the effectiveness literature. To reiterate, they are that better knowledge (proof) of what reduces recidivism is available and that more research is needed.

Several themes that distinguish the history of this perennial question serve as benchmarks for assessing these claims. A first and obvious theme is that the asking of what is effective in adult punishment has not only been extensive in duration, but unrelenting. The systematic pairing of offender recidivism with the question of what is effective has survived without interruption since the first decades of the 19th Century.

A second theme is that the pursuit of what is effective has been exhaustive in scope. The current analysis represents a small but illustrative picture of the tens of thousands of inquiries conducted over the past two centuries in the U.S alone. These inquiries have stretched across a vast penal landscape inhabited by many species of intervention. The interventions covered just in the studies cited here include the penitentiary, workhouse, probation, parole, diversion, work-release, cosmetic surgery, pharmaceutical medication, individual and group counseling, vocational and educational programming, community and institution-based substance abuse treatment, drug courts, home confinement, electronic monitoring, and boot camps.

A third theme is that the pursuit of what is effective has been sharpened by the analysis of several contingencies. These contingencies encompass the effects of differing research methodologies, treatment modalities, punishment and treatment settings, sentence lengths, type of service provider, duration of treatment, community supervision intensities, and probation and parole caseload sizes, just to name a few. For example, where the prison is concerned, three centuries of seekers have examined the relationship between recidivism and various conditions of confinement, methods of institutional discipline, type and length of confinement, prior record of the inmate, and a host of offense and offender characteristics.

Seekers of what is effective in the realm of non-custodial sanctions have been equally discriminating in their analyses. In the past and present, outcome indicators for probation and parole have included cost-effectiveness, rates of recidivism based on the commission of new offenses, and rates of absconding, non-reporting, and technical non-criminal violations. More specifically, analyses of probation and parole have been aligned with cost-savings and the offsetting of system expenditure through the financial remuneration of community-supervised offenders (e.g., restitution, fines, or fees) [19,77,78]. Throughout probation and parole's history, studies have also highlighted the difference between violations due to non-criminal and new criminal activity, with the former consistently proving to be more frequent in occurrence than the latter. These data have further shown that, for some offenders, the difficulty with reintegration has made reincarceration preferable to freedom. The list of contingent variables has only expanded with time, now including such factors as the impact of the probation officer's supervision style on recidivism [79].

A fourth and key theme in this history is the constancy of the variability of recidivism outcomes. This variability has persisted over time, within and across intervention types and studies, as well as research methodologies. For example, recidivism rates indicated by 19th Century analyses of prison's effectiveness ranged between 6 and 57 percent, which included rates of 37, 17, and 33 percent. This variation in recidivism ranges was also exhibited in 20th Century research on recidivism rates of post-prison releasees. Studies have shown recidivism rates of 33, 72, 48, and 67 percent across follow-up periods that differ by only 8 months. Where probation is concerned, government reports in the early years of its use indicated recidivism rates of roughly 6 percent, with some research in the mid-20th Century indicating recidivism rates of 10 percent. As reported here as well, 20th Century research on probation failures indicated ranges between 10 and 40 percent, depending on the follow-up period. In the early years of parole, government reports indicated recidivism rates between 1 and 21 percent for parolees. This theme of wide-ranging and unexplained variability is especially crystallized in the analysis of recidivism rates in Table 1.

The primary significance of this thematic history is an enlarged perspective on the claims of the effectiveness literature. Is better knowledge (proof) of what reduces recidivism available and is more research on this question of effectiveness needed? This delicate question calls to mind a similar question posed by the then late Reverend Louis Dwight in an 1856 article aptly titled "What Has Been Done and What Is to Do?" [80]. Although the question was posed rhetorically and with hope still in hand, Dwight openly contemplated whether we had "come to a resting place in our efforts to prevent crime . . . ?" [80] (p. 177).

In 2021, Dwight's question might be reframed as what has been gained and/or lost in this long pursuit. As for what has been gained, it must be acknowledged that the pursuit

of what is effective has yielded meaningful information for each generation and policy. It is also true that the research generating this knowledge has matured methodologically over many decades, culminating in the designation of certain practices as effective or “evidence-based” [5,76]. This certification assures policymakers that these practices have a better than random chance of producing the expected outcome of lower recidivism. For example, the research and policy institute Pew Trusts reports that programs that “have been proved to reduce recidivism through rigorous research” can be identified [81]. The Washington State Institute for Public Policy states their goal is to arm policymakers with knowledge of “programs and policies that have a proven ability to affect crime rates” [82] (p. 1). Similarly, the National Center for State Courts states that “unlike thirty years ago, we know—based on meticulous meta-analyses of rigorously conducted scientific research—that unlike incarceration, the right kinds of rehabilitation and treatment program . . . can reduce offender recidivism by conservative estimates of 10–20 percent” [83] (pp. 322–323).

With that being said, purported gains increasingly cast in the language of certitudes are problematic. Some leading voices in the field have argued that the evidence offered by criminological research is far short of what is needed to justify policy influence [7,84,85]. Others have similarly claimed that the rhetoric of proof that is permeating the “what works” discourse tends to overplay the current state of knowledge [8,12].

Ironically, the long journey to determine what is effective using recidivism has also produced a gain that is effectively a loss. It has been argued that the emphasis on effectiveness relative to recidivism has contributed to the ongoing expansion of the penal apparatus and growth of the penal population. Asking “what works” has cultivated the ongoing addition of new programs and interventions, each looking to achieve a more ideal yet unspecified threshold of recidivism reduction. This pattern of sustained system growth has been captured in the literature through many metaphors, the most legendary of which is net-widening [86–88]. Net-widening refers to the phenomenon of more and more of the general population being subject to some type of formal control that, in the absence of the penal reform, which is intended to reduce control and work better than the existing strategies, would have been subject to no control or less restrictive forms of control. What is basically a symbiotic relationship between the question of what is effective, recidivism, and the size of the punishment system has been eloquently depicted by Nils Christie [89], Stanley Cohen [88], and John Braithwaite [90]. The crux of their collective theoretical work is that the expansion of penal control is inexorable and without limit when the only or primary expressed end of punishment is effectiveness or utility.

It is recognized that, over the years, alternative narratives and questions have been advanced in adult punishment policy. Clearly, retribution and its questions of “what is fair” and “what is proportionate” have had a strong hearing throughout modern history. However, they are too often secondary to utilitarian concerns and have not been pursued with the same continuous empirical fervor. Restorative justice (RJ) and its question of “how do we set things right” for victims, offenders, and communities has gained little traction in the policy sphere, at least in the U.S. When RJ inspired programs (e.g., restitution or victim-offender mediation programs) have been evaluated, the question of effectiveness has still prevailed [91–94]. While it is beyond the scope of this paper to navigate a way forward, the words of Francis Allen still resonate and can be employed as a simple starting point. “No social institution as complex as those involved in the administration of criminal justice serve a single function or purpose” [94] (p. 99).

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Brown, M. *The Culture of Punishment: Prison, Society, and Spectacle*; New York University Press: New York, NY, USA, 2009.
2. Berman, D.A. Are Costs a Unique (and Uniquely Problematic) Kind of Sentencing Data? *Fed. Sentencing Rep.* **2012**, *24*, 159–160. [CrossRef]
3. Gottschalk, M. The past, present and future of mass incarceration in the United States. *Criminol. Public Policy* **2011**, *10*, 483. [CrossRef]
4. Blumstein, A. Linking evidence and criminal justice policy. *Criminol. Public Policy* **2013**, *12*, 721–730. [CrossRef]
5. Taxman, F.; Belenko, S. *Implementing Evidence-Based Practices in Community Corrections and Addiction Treatment*; Springer: New York, NY, USA, 2013.
6. Latessa, E.J. The challenge of change: Correctional programs and evidence-based practices. *Criminol. Public Policy* **2003**, *3*, 547. [CrossRef]
7. Cohen, S. The revenge of the null hypothesis: Evaluating crime control policies. *Crit. Criminol.* **1997**, *8*, 21–25.
8. Lee, L.C.; Stohr, M.K. A critique and qualified defense of “correctional quackery”. *J. Contemp. Crim. Justice* **2012**, *28*, 96–112. [CrossRef]
9. Heidemann, G.; Cederbaum, J.A.; Martinez, S. Beyond recidivism: How formerly incarcerated women define success. *Affilia* **2016**, *31*, 24–40. [CrossRef]
10. Gould, M.R. Rethinking our metrics: Research in the field of higher education in prison. *Prison J.* **2018**, *98*, 387–404. [CrossRef]
11. Andersen, T.S.; Scott DA, I.; Boehme, H.M.; King, S.; Mikell, T. What matters to formerly incarcerated men? Looking beyond recidivism as a measure of successful reintegration. *Prison J.* **2020**, *100*, 488–509. [CrossRef]
12. Klingele, C. The promises and perils of evidence-based corrections. *Notre Dame Law Rev.* **2016**, *91*, 537–584.
13. Lucken, K. *Rethinking Punishment: Challenging Conventions in Research and Policy*; Taylor & Francis: Abingdon, UK, 2017.
14. Pennsylvania Prison Society. Message from the governor of New Jersey to both houses of the seventieth legislature. *J. Prison Discip. Philanthr.* **1846**, *2*, 380–390.
15. The Philadelphia Society for Alleviating the Miseries of Public Prisons. Whippers and anti-whippers. *Pa. J. Prison Discip. Philanthr.* **1851**, *6*, 90–93.
16. Kuntz, W.F. *Criminal Sentencing in Three Nineteenth-Century Cities: Social History of Punishment in New York, Boston, and Philadelphia, 1830–1880*; Taylor & Francis: Abingdon, UK, 1988.
17. New Englander. Prison Discipline as a Science. *New Englander* **1873**, *122*, 70. Available online: <https://www.tandfonline.com/doi/full/10.1080/03071022.2019.1579977> (accessed on 5 March 2021).
18. Pennsylvania Prison Society. The county prisons of Pennsylvania. *J. Prison Discip. Philanthr.* **1920**, *8*, 1857–1920.
19. Byers, J.F. Prison reform: Address of the president of the American prison association. *J. Prison Discip. Philanthr.* **1916**, *55*, 48. [CrossRef]
20. V., A.H. Adult probation in Philadelphia. *J. Prison Discip. Philanthr.* **1857–1920** **1915**, *54*, 33.
21. Vanstone, M. The International Origins and Initial Development of Probation: An Early Example of Policy. *Transfer. Br. J. Criminol.* **2008**, *48*, 735–755. [CrossRef]
22. Pennsylvania Prison Society. Parole of prisoners in Pennsylvania. *J. Prison Discip. Philanthr.* **1857–1920** **1916**, *55*, 21.
23. Bates, S. How good is parole? A debate. *Forum Century C* **1938**, *2*, 69–70.
24. Parole work in Pennsylvania. *J. Prison Discip. Philanthr.* **1913**, *52*, 28.
25. Bailey, V. Penal legislation of 1915 in the commonwealth of Pennsylvania. *J. Prison Discip. Philanthr.* **1916**, *54*, 70.
26. Bailey, W.C. Correctional outcome: An evaluation of 100 reports. *J. Crim. Law Criminol.* **1966**, *57*, 153–160. [CrossRef]
27. Dalton, R.H. Value and use of counseling techniques in the work of probation officers. *Fed. Probation.* **1952**, *16*, 17.
28. Kirby, B.C. Measuring effects of treatment of criminals and delinquents. *Sociol. Soc. Res.* **1954**, *38*, 368–374.
29. Robison, J.; Smith, G. The Effectiveness of Correctional Programs. *Crime Delinq.* **1971**, *17*, 67–80. [CrossRef]
30. Levin, M.A. Policy evaluation and recidivism. *Law Soc. Rev.* **1971**, *6*, 17–46. [CrossRef]
31. Beattie, R.H.; Bridges, C.K. *Superior Court Probation and/or Jail Sample*; Sacramento, Cal.: Bureau of Criminal Statistics, California Department of Justice: Sacramento, CA, USA, 1970.
32. Davis, G.F. A study of adult probation violation rates by means of the cohort approach. *J. Crim. Law Criminol. Police Sci.* **1964**, *55*, 70–85. [CrossRef]
33. England, R.W., Jr. What is responsible for satisfactory probation and post-probation outcome. *J. Crim. Law Criminol. Police Sci.* **1957**, *47*, 667. [CrossRef]
34. Kolodney, S.E.; Patterson, P.L.; Daets, D.; Marx, R.L. *A Study of the Characteristics and Recidivism Experience of California Prisoners*; Public Systems Incorporated: San Jose, CA, USA, 1970.
35. Glaser, D. *The Effectiveness of a Prison and Parole System*; Bobbs-Merrill: Indianapolis, IN, USA, 1964.
36. Martinson, R. What works?—Questions and answers about prison reform. *Public Interest* **1974**, *35*, 22.
37. Gendreau, p. Treatment in corrections: Martinson was wrong. *Can. Psychol.* **1981**, *22*, 332–338. [CrossRef]
38. Cullen, F.T.; Gilbert, K.E. *Reaffirming Rehabilitation*; Anderson: Cincinnati, OH, USA, 1982.
39. Lipsey, M.W.; Cullen, F.T. The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annu. Rev. Law Soc. Sci.* **2007**, *3*, 297–320. [CrossRef]

40. Katz, L.; Levitt, S.D.; Shustorovich, E. Prison conditions, capital punishment, and deterrence. *Am. Law Econ. Rev.* **2003**, *5*, 318–343. [\[CrossRef\]](#)
41. Nagin, D.S. Criminal Deterrence Research at the Outset of the Twenty-First Century. *Crime Justice* **1998**, *23*, 1–42. [\[CrossRef\]](#)
42. Paternoster, R. The deterrent effect of the perceived certainty and severity of punishment: A review of the evidence and issues. *Justice Q.* **1987**, *4*, 173–217. [\[CrossRef\]](#)
43. Weisburd, D.; Waring, E.; Chayet, E. Specific deterrence in a sample of offenders convicted of white-collar crimes. *Criminology* **1995**, *33*, 587–607. [\[CrossRef\]](#)
44. Orsagh, T.; Chen, J.-R. The effect of time served on recidivism: An interdisciplinary theory. *J. Quant. Criminol.* **1988**, *4*, 155–171. [\[CrossRef\]](#)
45. Cullen, F.T.; Gendreau, P. From nothing works to what works: Changing professional ideology in the 21st century. *Prison J.* **2001**, *81*, 313–338. [\[CrossRef\]](#)
46. Palmer, T. Martinson Revisited. *J. Res. Crime Delinq.* **1975**, *12*, 133–152. [\[CrossRef\]](#)
47. Palmer, T. The effectiveness issue today: An overview. *Fed. Probat.* **1983**, *47*, 3–10.
48. Van Voorhis, P. Correctional effectiveness: The high cost of ignoring success. *Fed. Probat.* **1987**, *51*, 56–62.
49. Bushway, S.; Paternoster, R. The impact of prison on crime. In *Do Prisons Make Us Safer?* Raphael, S., Stoll, M.A., Eds.; Russell Sage Foundation: New York, NY, USA, 2009; pp. 119–150.
50. Hawken, A.; Kleiman, M. *Managing Drug Involved Probationers with Swift and Certain Sanctions: Evaluating Hawaii's HOPE: Executive Summary*; National Criminal Justice Reference Services: Washington, DC, USA, 2009.
51. Oleson, J.C. HOPE Springs Eternal. *Criminol. Public Policy* **2016**, *15*, 1163–1183. [\[CrossRef\]](#)
52. O'Connell, D.J.; Brent, J.J.; Visser, C.A. Decide Your Time. *Criminol. Public Policy* **2016**, *15*. [\[CrossRef\]](#)
53. Lattimore, P.K.; MacKenzie, D.L.; Zajac, G.; Dawes, D.; Arseneault, E.; Tueller, S. Outcome findings from the HOPE Demonstration Field Experiment: Is swift, certain, and fair an effective supervision strategy? *Criminol. Public Policy* **2016**, *15*, 1103–1141. [\[CrossRef\]](#)
54. Hamilton, Z.; Campbell, C.M.; van Wormer, J.; Kigerl, A.; Posey, B. Impact of swift and certain sanctions: Evaluation of Washington State's policy for offenders on community supervision. *Criminol. Public Policy* **2016**, *15*, 1009–1072. [\[CrossRef\]](#)
55. Blumstein, A.; Cohen, J. Estimation of individual crime rates from arrest records. *J. Crim. Law Criminol.* **1979**, *70*, 561. [\[CrossRef\]](#)
56. Dilulio, J.J.; Piehl, A.M. Does prison pay? The stormy national debate over the cost-effectiveness of imprisonment. *Brook. Rev.* **1991**, *9*, 28–35. [\[CrossRef\]](#)
57. Greenberg, D.F. The incapacitative effect of imprisonment: Some estimates. *Law Soc. Rev.* **1975**, *9*, 541–580. [\[CrossRef\]](#)
58. Greenwood, P.W.; Abrahamse, A.F. *Selective Incapacitation*; Rand: Santa Monica, CA, USA, 1982; pp. 11–26.
59. Marvell, T.B.; Moody, C.E. Prison population growth and crime reduction. *J. Quant. Criminol.* **1994**, *10*, 109–140. [\[CrossRef\]](#)
60. Spelman, W. What Recent Studies Do (and Don't) Tell Us about Imprisonment and Crime. *Crime Justice* **2000**, *27*, 419–494. [\[CrossRef\]](#)
61. Spelman, W. *Criminal Incapacitation*; Plenum Press: New York, NY, USA, 1994.
62. Cohen, M.A. Pain, suffering, and jury awards: A study of the cost of crime to victims. *Law Soc. Rev.* **1988**, *22*, 537. [\[CrossRef\]](#)
63. Levitt, S.D. The effect of prison population size on crime rates: Evidence from prison overcrowding litigation. *Q. J. Econ.* **1996**, *111*, 319–351. [\[CrossRef\]](#)
64. Zedlewski, E.W. *Making Confinement Decisions*; US Department of Justice, National Institute of Justice: Washington, DC, USA, 1987.
65. Visser, C.A. Incapacitation and crime control: Does a "Lock'em up" strategy reduce crime? *Justice Q.* **1987**, *4*, 513–543. [\[CrossRef\]](#)
66. Greenberg, D.F. The cost-benefit analysis of imprisonment. *Soc. Justice* **1990**, *17*, 49–75.
67. MacKenzie, D.L. *What Works in Corrections*; Cambridge University Press: New York, NY, USA, 2006.
68. Auerhahn, K. Selective incapacitation and the problem of prediction. *Criminology* **1999**, *37*, 703–734. [\[CrossRef\]](#)
69. Langan, P.A.; Levin, D.J. Recidivism of prisoners released in 1994. *Fed. Sentencing Report.* **2002**, *15*, 58–65. [\[CrossRef\]](#)
70. Eggleston, E.P.; Laub, J.H.; Sampson, R.J. Methodological sensitivities to latent class analysis of long-term criminal trajectories. *J. Quant. Criminol.* **2004**, *20*, 1–26. [\[CrossRef\]](#)
71. Sweeten, G.; Apel, R. Incapacitation: Revisiting an Old Question with a New Method and New Data. *J. Quant. Criminol.* **2007**, *23*, 303–326. [\[CrossRef\]](#)
72. Bhati, A.S. Estimating the number of crimes averted by incapacitation: An information theoretic approach. *J. Quant. Criminol.* **2007**, *23*, 355–375. [\[CrossRef\]](#)
73. Piquero, A.R.; Blumstein, A. Does Incapacitation Reduce Crime? *J. Quant. Criminol.* **2007**, *23*. [\[CrossRef\]](#)
74. Sherman, L.W.; Gottfredson, D.; MacKenzie, D.; Eck, J.; Reuter, P.; Bushway, S. *Preventing Crime: What Works, What Doesn't, What's Promising*; U.S. Department of Justice, National Institute of Justice: Washington, DC, USA, 1998.
75. Laster, R.E. Criminal restitution: A survey of its past history and an analysis of its present usefulness. *Univ. Richmond Law Rev.* **1970**, *5*, 71.
76. McSparron, J. Community correction and diversion: Costs and benefits, subsidy modes, and start-up recommendations. *Crime Delinq.* **1980**, *26*, 226–247. [\[CrossRef\]](#)
77. Seiter, R.P.; West, A.D. Supervision styles in probation and parole: An analysis of activities. *J. Offender Rehabil.* **2003**, *38*, 57–75. [\[CrossRef\]](#)

78. Dwight, L. What has been done and what is to do. *Pa. J. Prison Discip. Philanthr.* **1856**, *11*, 177.
79. Dube, S.; White, D. Iowa Uses Research and Data to Identify Corrections Programs That Work. 17 September 2018. Available online: <https://www.pewtrusts.org/en/research-and-analysis/articles/2018/09/17/iowa-uses-research-and-data-to-identify-corrections-programs-that-work> (accessed on 1 March 2021).
80. Washington State Institute for Public Policy. *Evidence-Based Adult Corrections Programs: What Works and What Does Not*; Institute for Public Policy: Olympia, WA, USA, 2006.
81. Warren, R.K. Evidence-Based Sentencing. *Fed. Sentencing Rep.* **2008**, *20*, 322–325. [[CrossRef](#)]
82. Tittle, C.R. The Arrogance of Public Sociology. *Soc. Forces* **2004**, *82*. [[CrossRef](#)]
83. Wellford, C.F. History and evolution of criminology. In *21st Century Criminology: A Reference Handbook*; Miller, J.M., Ed.; Sage: Thousand Oaks, CA, USA, 2009; pp. 10–17.
84. Blomberg, T. Diversion and accelerated social control. *J. Crim. Law Criminol.* **1977**, *68*, 274–282. [[CrossRef](#)]
85. Austin, J.; Krisberg, B. NCCD research review: Wider, stronger, and different nets: The dialectics of criminal justice reform. *J. Res. Crime Delinq.* **1981**, *18*, 165–196. [[CrossRef](#)]
86. Cohen, S. *Visions of Social Control*; Polity Press: Cambridge, UK, 1985.
87. Braithwaite, J. Beyond positivism: Learning from contextual integrated strategies. *J. Res. Crime Delinq.* **1993**, *30*, 383–399. [[CrossRef](#)]
88. Christie, N. *Crime Control as Industry*; Routledge: Abingdon, UK, 1993.
89. De Beus, K.; Rodriguez, N. Restorative justice practice: An examination of program completion and recidivism. *J. Crim. Justice* **2007**, *35*, 337–347. [[CrossRef](#)]
90. Bonta, J.; Jessemann, R.; Rugge, T.; Cormier, R. Restorative justice and recidivism: Promises made, promises kept. In *Handbook of Restorative Justice: A Global Perspective*; Routledge: Abingdon, UK, 2006; Volume 108, p. 120.
91. Latimer, J.; Dowden, C.; Muise, D. The effectiveness of restorative justice practices: A meta-analysis. *Prison J.* **2005**, *85*, 127–144. [[CrossRef](#)]
92. Strang, H.; Sherman, L.W.; Mayo-Wilson, E.; Woods, D.; Ariel, B. Restorative Justice Conferencing (RJC) Using Face-to-Face Meetings of Offenders and Victims: Effects on Offender Recidivism and Victim Satisfaction. A Systematic Review. *Campbell Syst. Rev.* **2013**, *9*, 1–59. [[CrossRef](#)]
93. Allen, F. Legal values and the rehabilitative ideal. In *Why Punish? How Much?* Tonry, M., Ed.; Oxford University Press: New York, NY, USA, 2010; pp. 97–105.
94. Lipton, D.S.; Martinson, R.; Wilks, J. *The Effectiveness of Correctional Treatment: A Survey of Treatment Evaluation Studies*; Praeger Publishers: Santa Barbara, CA, USA, 1975.