

1-15-2018

When Bad News Arrives: Project HOPE in a Post-Factual World

Francis T. Cullen

Travis C. Pratt

Jillian J. Turanovic

Leah C. Butler

Follow this and additional works at: <https://digitalcommons.unomaha.edu/criminaljusticefacpub>

 Part of the [Criminology Commons](#)



When Bad News Arrives: Project HOPE in a Post-Factual World

Francis T. Cullen¹, Travis C. Pratt¹, Jillian J. Turanovic², and Leah Butler¹

¹University of Cincinnati, OH, USA

²Florida State University, Tallahassee, FL, USA

Abstract

On the basis of limited empirical evidence, advocates of Project HOPE (Hawaii's Opportunity Probation with Enforcement) have succeeded in spreading the model to a reported 31 states and 160 locations. A recent randomized control experiment across four sites has revealed negative results: no overall effect on recidivism. In this context, we examine how prominent advocates of Project HOPE have coped with the arrival of this "bad news." Despite null findings from a "gold standard" evaluation study, advocates continue to express confidence in the HOPE model and to support its further implementation. The risk thus exists that Project HOPE is entering a post-factual world in which diminishing its appeal—let alone its falsification—is not possible. It is the collective responsibility of corrections researchers to warn policy makers that the HOPE model is not a proven intervention and may not be effective in many agencies. It is also our responsibility to create a science of community supervision that can establish more definitively best practices in this area.

Keywords

deterrence, offender supervision, parole, probation, Project HOPE

In 2004, First Circuit Court Judge Steve S. Alm envisioned, designed, and then led the implementation of Hawaii's Opportunity Probation with Enforcement program—now commonly known as Project HOPE. Judge Alm was troubled that so many offenders did not comply with their conditions of probation, failing drug tests or not showing up for such tests. After repeated violations with few consequences, the time would come when

an exasperated judge would “lower the boom” by revoking their probation and sending them to prison. Alm reasoned that a commonsensical solution would be to sanction each violation immediately—that is, to engage in swift and certain punishment. Another key insight was to resist the temptation to impose severe penalties. Instead, sanctions would be modest (a few days in jail) and graduated (increasing as violations increased). All offenders would be warned at the outset of their probation about the sanctioning system, and all would be punished the same—that is, fairly.

Project HOPE thus was erected on the principles of swift–certain–fair punishment—referred to by the acronym SCF probation. Criminologically, SCF probation (or parole) is based on Cesare Beccaria’s classical school, assuming that sanctions that were imposed swiftly, with certainty, and with fairness would deter misconduct and motivate compliance. By contrast, under the HOPE model, rehabilitation was de-emphasized, to be given only to those who asked for it or for those who failed multiple drug tests (Hawken, 2010b; Hawken & Kleiman, 2009).

Project HOPE might have remained a local innovation that did not migrate to the mainland, but two factors, among others (Duriez, Cullen, & Manchak, 2014), coalesced to transform the program into a national and potentially international movement. First, evaluation research indicated that Project HOPE worked to reduce failed drug tests and arrests (Hawken & Kleiman, 2009). The intervention could claim to be “evidenced based.” Second, three special people were prominently involved in the program’s assessment—Judge Alm, the inventor, and two evaluators, Angela Hawken and Mark A. R. Kleiman. Alm was important because he is charismatic and committed to spreading the good news about his Project HOPE. Hawken and Kleiman are not typical armchair criminologists, but rather policy analysts now both on the faculty at the Marron Institute of Urban Management at New York University. Their vitas indicate that Hawken was trained in economics at the University of the Witwatersrand in Johannesburg, South Africa, prior to receiving a PhD in policy analysis at the RAND Graduate School. Kleiman, a noted public intellectual, earned a doctorate in public policy from Harvard University’s John F. Kennedy School of Government.

Together and independently, this trinity of HOPE advocates was persuasive in trumpeting the intervention. Judge Alm could tell stories about offenders who attributed

their reform—indeed, saving their lives—to his program. Hawken and Kleiman had the scholarly expertise to show that the intervention was effective, based on indisputable deterrence principles, and a solution to broken probation and parole systems. They were masterful in explaining the value of the SCF program to policy makers and, especially in the case of Hawken, of supplying technical expertise to implement programs. In 2014, it was estimated that Project HOPE—with the “H” changed to “Honest” when it moved to the mainland—was in 40 jurisdictions (Pearsall, 2014). A more recent estimate was 160 jurisdictions, including 31 states (Alm, 2016; Bartels, 2016). International interest in HOPE is percolating as well (Bartels, 2016; Oleson, 2016).

Project HOPE advocates were so effective that few criminologists expressed reservations about the intervention (for exceptions, see Clear & Frost, 2014; Cullen, Manchak, & Duriez, 2014; Duriez et al., 2014; Klingele, 2015). The acceptance or silence about the program was remarkable in light of criminologists' traditional rejection of deterrence-oriented, punishment-based interventions. Still, the HOPE model offered a trade-off that seemed to attract scholars ordinarily suspicious of punitiveness: By punishing swiftly, certainly, and briefly (short jail stays), offenders would be kept out of prison—the one outcome most dreaded by progressive criminologists. HOPE thus offered compassionate punishment—applied, much as parents would do for their child, for the offender's own good.

What Project HOPE advocates and acceptors did not envision was that subsequent evaluations might produce less favorable results, even to the point of questioning the program's effectiveness. Alas, such bad news is precisely what has arrived (Cook, 2016; Cullen, Pratt, & Turanovic, 2016, 2017). As will be discussed, particularly troubling is a four-site, randomized controlled trial (RCT), “gold standard” evaluation showing that the HOPE model has virtually no effects on offender recidivism (Lattimore, Dawes, et al., 2016; Lattimore, MacKenzie, et al., 2016). From what we can discern, however, Alm, Hawken, and Kleiman seem unfazed by this bad news and appear to be proceeding vigorously in their evangelizing for HOPE.

Such confidence in the face of negative facts is perhaps understandable. As Kuhn (1970) taught us, anomalous findings can be dismissed when members of a dominant paradigm see their model as indisputable and see alternative realities as impossible.

Similarly, all of us are guilty of confirmation bias—of paying attention to good news and ignoring bad news—and likely more so when we have invested much in an idea (Kahneman, 2011). Still, the message of this essay is that inconvenient truths cannot just be shoved aside. It is our collective responsibility to ensure that Project HOPE is not allowed to flourish unchallenged in a post-factual world. Facts must matter (see Gilbert, 1997).

This article proceeds in four parts. The first section suggests that from the beginning of the HOPE movement, its advocates failed to consider the limits of initially positive evaluation evidence and substantial research evidence documenting the limits of deterrence-oriented correctional interventions. Second, the arrival of bad evaluation news and why it cannot be ignored are discussed. Third, the reaction to this bad news by Alm, Hawken, and Kleiman is reviewed, showing why it is now unlikely that they will stop proselytizing in favor of HOPE programs. Fourth, this article concludes by calling for a moratorium on any further implementation of Project HOPE interventions until a strong body of supportive evaluation evidence is forthcoming.

As a prelude to this analysis, we want to emphasize that Steven Alm is a remarkable jurist and that Angela Hawken and Mark Kleiman are first-rate scholars and policy analysts. All three advocates are passionate about HOPE and believe that it will improve the effectiveness of community supervision agencies and the lives of offenders. They are good people trying to do good things. Furthermore, from our writings, we are now seen as HOPE doubters, and thus it is a fair criticism that we might manifest our own brand of confirmation bias, seeing the HOPE glass as half empty rather than as half full (Cullen et al., 2014; Cullen et al., 2016, 2017; Duriez et al., 2014). Rest assured that Hawken and Kleiman could offer a rebuttal to this article that some might find persuasive. As with any scientific enterprise, our audience should practice organized skepticism and see which side of this debate is correct (Merton, 1973).

All this said, we are not mere naysayers but rather are genuinely concerned that the effectiveness of Project HOPE is being oversold, and its weak empirical base is misunderstood by consumers of the intervention. If nothing else, our goal is to encourage policy makers and practitioners to slow down and read all the evidence on HOPE so as to avoid impulsively jumping on a bandwagon that may eventually crash.

The Dangers of Advocacy

Reading their works and hearing them speak, it is clear that Alm, Hawken, and Kleiman exude confidence that Project HOPE is erected on a solid empirical and theoretical foundation. They are not injudicious and warn that implementation is a challenge that must be overcome in a nuanced way. But they manifest no hint of doubt about HOPE's effectiveness. At least as far as we can see, they have never asked: What if we are wrong? They should have. Even before bad news arrived, three considerations should have diminished their hubris to the point of humility.

First, even if the initial evaluation of Project HOPE conducted by Hawken and Kleiman (2009) was flawless and peer reviewed—neither of which is the case—it still would have been risky to disseminate the HOPE model for implementation elsewhere. The wisdom of exercising caution in part is tied to the fact that Steven Alm is a special jurist and that Hawaii has unique cultural characteristics (e.g., nearly two thirds of participants in the Hawken & Kleiman, 2009, RCT evaluation were Asian/Polynesian). More worrisome is the growing realization that many classic studies simply do not replicate—that initial results prove idiosyncratic and not representative of the true empirical reality (Ioannidis, 2005; Lehrer, 2010). The replication crisis is now well known across scientific fields and has occurred within research on criminal sanctions, including intensive supervision programs (Hyatt & Barnes, 2017; Petersilia & Turner, 1993) and domestic violence arrest experiments (see, e.g., Dunford, Huizinga, & Elliott, 1990).

Hawken and Kleiman's (2009) evaluation of Project HOPE is easily available online and can be read in detail by interested parties (see <https://www.ncjrs.gov/pdffiles1/nij/grants/229023.pdf>). It reports on two evaluations—one a quasi-experimental pilot study and the other an RCT—both of which achieved positive results. We do not question the findings. However, Hawken and Kleiman (2009) outlined the “study limitations” (pp. 48-50) of their pilot study, which apply as well to the RCT. Three are particularly important: (1) a 12-month follow-up period that precluded assessing whether treatment effects would persist after offenders completed their probation and were no longer under HOPE supervision, (2) training of Hawaiian probation officers in core correctional practices such as cognitive-behavioral therapy and motivational interviewing that may not occur in other jurisdictions, and (3) the inability to decompose

“essential elements” of the HOPE model to know what was producing the effects (e.g., drug testing, certain sanctions, or the interaction of the two). We can add one more: (4) the use of failed drug tests and missed appointments as triggers for sanctions, which may not be appropriate for nondrug-involved offenders who violate other supervision conditions (e.g., associating with known felons, curfew restrictions).

Taken together, these limitations defined Project HOPE as a promising—not a proven—program that in 2009 merited much more research to establish the generalizability of, and reasons for, its effects. No basis existed for what subsequently occurred: the widespread implementation of the model without systematic research. To our knowledge, Hawken, Kleiman, and Alm did not at any point call for a halt to this program dissemination on the grounds that the evidence was not yet sufficiently developed to justify it.

Second, perhaps thinking that Project HOPE’s unique features make it the exception to the rule, its advocates seem to be unconcerned by (or unaware of) the consistent finding of literature reviews, including meta-analyses, that deterrence-oriented correctional interventions have weak, null, or even iatrogenic effects on recidivism (Lipsey & Cullen, 2007; MacKenzie, 2006). By contrast, programs emphasizing human services to offenders are among the most effective (Bonta & Andrews, 2017; Lipsey, 2009). McGuire (2013), for example, has assessed 100 meta-analyses or systematic reviews on “what works to reduce re-offending” (p. 20). The data revealed that “the only recurrently negative mean effect sizes reported to date are those obtained from criminal sanctions or deterrence-based methods. Punitive sanctions repeatedly emerge as a failed strategy for altering offenders’ behaviour” (McGuire, 2013, p. 30). MacKenzie and Farrington (2015) were equally conclusive: “Interventions based on surveillance, control, deterrence, or discipline are ineffective” (p. 565). The weakness in such deterrence programs is that the intervention used does not target for change and is not responsive to the strongest risk factors that predict recidivism (Bonta & Andrews, 2017).

Third, if Project HOPE is the one deterrence program that is effective, this claim would rest on the fact that certain and swift sanctions induce offender compliance. Starting in their 2009 report, Hawken and Kleiman have implied that the existing

research substantiates this assertion. If our reading of the literature is accurate, this is not the case; the evidence is equivocal or scarce. Unfortunately, it appears that HOPE advocates are unaware of contrary research and, at least in some instances, misread or place excessive confidence in works that they cite. Two points merit consideration.

First, as Paternoster and Bachman (2013) noted, it is possible to “conclude with reasonable confidence” (p. 666) that perceived certainty of punishment is inversely related to offending. Still, criminal decision-making is complex and can be specified by characteristics of the individual (e.g., psychological traits) and of the sanction (e.g., the threat of informal sanctions is more effective than of formal sanctions). There is much that is unknown. “Any review of a body of work such as this is a humbling experience,” note Paternoster and Bachman (2013),

because it clearly shows that despite decades of research on deterrence, there is much we still do not know—other than that our theories and the research that they generate may still be primitive and in great need of modification and development. (p. 666)

Most important, research shows that the size of the perceived certainty effect becomes weak when other predictors of crime (e.g., antisocial values, social bonds) are introduced. In a meta-analysis, Pratt, Cullen, Blevins, Daigle, and Madensen (2006) showed that once variables from competing theories were controlled, the effect size for certainty (measured as a mean standardized β regression coefficient) decreased from $-.232$ to $-.072$.

The Pratt et al. (2006) study was available to Hawken and Kleiman (2009) when they undertook their literature review and claimed that the “HOPE program has a strong theoretical base” and that the core elements of the project “are well supported by prior research” (p. 9). In fact, by year’s end 2016, Google Scholar showed that the Pratt et al. meta-analysis, which would have challenged these conclusions, had been cited 264 times. Let us also consider two empirical studies that Hawken and Kleiman discovered and cited regarding the effectiveness of perceived certainty: Grasmick and Bryjak (1980) and Paternoster (1989). Thus, the Grasmick and Bryjak study is dated and misspecified because it contains no controls for other known crime predictors. Regardless, the findings it reports are contrary to the HOPE model: Severity deters and the effects of

certainty are weakest at lower levels of severity. In the Paternoster (1989) study, the abstract notes that the “observed effects” for perceived certainty “were marginal” and significant for only “some offenses” (p. 7). Paternoster examined the effects of certainty (perceived likelihood of being caught by the police) when committing four offenses (marijuana use, drinking liquor, petty theft, and vandalism) across models reported in six tables. Notably, of the 24 possible effects, perceived certainty was significant in only six, or 25%, of the analyses.

Second, Hawken and Kleiman (2009) asserted that the “immediacy, or celerity, of a sanction is also vital for shaping behavior” (p. 9). It is puzzling, however, that the only work cited in support of this key program component is Farabee’s (2005) monograph published by the American Enterprise Institute, a conservative think-tank. In this document, Farabee mainly advocates for punishment over rehabilitation but supplies only a partial review of the deterrence literature (see Cullen, Smith, Lowenkamp, & Latessa, 2009). As far as we can tell, the main evidence he cites in favor of celerity is an off-hand quote from James Q. Wilson (1997): “When raising children, we do not say: ‘Because [you’ve misbehaved], you have a 50-50 chance nine months from now of being grounded’” (Farabee, 2005, p. 60). With slight variation in wording, this quote is repeated by Hawken and Kleiman (2009)—again to show why swiftness matters. Alas, criminological argument through contrived anecdotes does not take the place of solid evidence. We could just as easily retort that no parents would say, “Because you’ve misbehaved, we are confining you to your room immediately for two days”—as opposed to disciplining their child verbally within a context of love and showing alternative courses of prosocial action (see Wright & Cullen, 2001). In both cases, the statements are silly.

More tellingly, in their review, Paternoster and Bachman (2013) concluded that “most research has shown that the celerity of punishment is not strongly or not at all related to offending decisions” and that “we know very little conceptually or empirically about celerity” (p. 666; see also Pratt & Turanovic, in press). There is a body of evaluation studies in experimental psychology showing that celerity of punishment can shape behavior but that its effects begin to diminish if the sanction is not applied almost immediately (seconds or minutes) after an act has occurred (Abramowitz & O’Leary,

1990; Lerman & Vorndran, 2002). That the effects of punishment can decay quickly led Pratt and Turanovic (in press) to conclude that achieving this level of immediacy in the criminal justice system—where the “perfect tyranny” of the laboratory is traded for the reality of correctional supervision—is “a practical impossibility.” At the very least, no research exists that has established how swiftly a sanction must be applied to increase the deterrence of offenders under community supervision. This remains an unknown fact, a troubling reality for the theory underlying Project HOPE.

In sum, the purpose of this section is to suggest that Project HOPE has never enjoyed strong empirical confirmation. It was applied effectively in one location by its inventor, but the criminological evidence on deterrence and its core elements—certainty and swiftness of punishment—have, from its inception, remained weak if not unsupported. As will be discussed later, Hawken and Kleiman counter that evaluations of other, non-HOPE programs using SCF principles have produced positive results (Hawken et al., 2016; Kleiman, Kilmer, & Fisher, 2014). Our position, however, is that the existing research simply did not justify the spread of Project HOPE to 40 jurisdictions, let alone 160. Fortunately, a gold standard, multisite RCT was conducted and just recently published that should go a long way in deciding whether continued advocacy of, or a swift call to halt the program’s dissemination, is in order (Lattimore, Dawes, et al., 2016; Lattimore, MacKenzie, et al., 2016). Indeed, as Hawken et al. (2016) noted in anticipation in May of 2016,

HOPE-style supervision is relatively new and only a handful of studies (of varying quality) assess the effectiveness of this approach. This underscores the importance of the HOPE DFE (in four states) that is supported by DOJ (the evaluation is led by RTI and Penn State). The experiment entails a 1600-subject randomized controlled trial, with findings expected in 2016. (p. 27)

Bad News

Near the end of 2016, *Criminology & Public Policy* (Volume 15, Number 4) published the “HOPE Collection”—three major evaluations of the HOPE model and then seven reaction articles that included, among others, commentaries by Alm, Hawken, Kleiman, and us (Cullen et al., 2016). For all relevant details, the volume should be

consulted. But the punch line was clear: The news for HOPE was not good (Cook, 2016). A large statewide quasi-evaluation in Washington produced small but significant results (Hamilton, Campbell, van Wormer, Kigerl, & Posey, 2016). However, across five sites, one in Delaware (O'Connell, Brent, & Visher, 2016) and four in the evaluation conducted by Lattimore, MacKenzie, et al. (2016), the effects on recidivism were virtually null. Due to its quality and concerted effort to replicate the HOPE model, Lattimore and colleagues' experimental study leveled a particularly hard blow. More bad news then arrived from our meta-analysis of HOPE evaluations (Cullen et al., 2016). Each of these unhappy messages—from our team and from Lattimore's team—is considered below.

Meta-Analysis of HOPE

As a means of summarizing the HOPE evaluation research, we reported in our *Criminology & Public Policy* commentary a meta-analysis of the evaluations across the six sites reported in the Hamilton et al.; O'Connell et al.; and Lattimore, MacKenzie et al. articles. To be fair, we also included in the analysis two evaluations from Hawaii—the original RCT (Hawken & Kleiman, 2009) and a recent replication (Hawken et al., 2016). The results reveal weak to null effects, with all confidence intervals containing zero (Cullen et al., 2016, p. 5, Table 1). Specifically, the overall effect size was .070 (using the absolute risk reduction—ARR—effect size). The results from Hawaii (.195) and for drug offenses (.156) were stronger. But the mean effects for other offenses (.038) and for programs on the mainland (.033) were tiny. Our message was that no program with this empirical foundation warrants implementation.

At the 2016 American Society of Criminology (ASC) conference, a session was organized at which the findings of three evaluation studies contained in *Criminology & Public Policy* were presented, as well as commentaries by Cullen, Hawken, and Kleiman (ASC, 2016). In response, Hawken and Kleiman each complained that the meta-analysis was flawed for not including a number of evaluation studies showing that SCF programs were effective. Our purpose, however, was to quantitatively synthesize the evaluation findings of rigorous studies of the HOPE model—not generic SCF programs that, as noted above, Hawken et al. (2016) had characterized as small in

number and varying in quality.

Table 1. Summary of Studies Specified as Supporting the HOPE Model in the Face of Negative Evidence.

Study	Setting and intervention	Research design	Key findings	Credible challenge
Carns and Martin (2011)	Alaska: PACE, based explicitly on HOPE model, staff trained in Hawaii	Random assignment to PACE (n = 63) or control group (unspecified sample size)	Data on outcomes could not be collected on control group subjects; thus, no comparison was possible	No
Caulkins and Dupont (2010)	South Dakota: 24/7 Sobriety Project	Narrative review of the outcome evidence regarding the effectiveness of 24/7	No evaluation was conducted	No
Cavanaugh, Godfrey, and Harney (1994)	Oregon: Multnomah County Structured Sanctions Process	Process evaluation of the program's implementation	No outcome evaluation was conducted with respect to offender behavior	No
Cavanaugh and Harrell (1995)	Oregon: Multnomah County Drug Testing Program	Randomly assigned drug offenders to treatment (n = 168) and control (n = 83) conditions	Universally null treatment effects on recidivism	No
Devall, Lanier, and Hartmann (2015)	Michigan: Swift and Sure Sanctions Probation Program, based explicitly on HOPE model	Propensity score matching (n = 379 treatment and 379 controls)	Treatment effect of 9%; unspecified magnitude of both positive and negative effects by offense category	No
Dupont and Wish (1992)	National Effort: Operation Tripwire for drug abusing (primarily heroin) offenders	Narrative review of selected drug use and treatment literature	No evaluation was conducted	No
Grommon, Cox, Davidson, and Bynum (2013)	Midwest Urban County: Drug testing with immediate results and sanctions	Random assignment to treatment and control conditions (n = 529 total offenders)	Mixed with some statistically significant treatment effects of between 5% and 8% on recidivism	No
Harrell and Roman (2001)	Washington, DC: Graduated Sanctions Program for felony drug offenders	Random assignment to treatment and control conditions (n = up to 916)	Randomization failed (treatment group was disproportionately female and older); mixed null and some treatment effects of up to 11%	No
Harrell, Roman, Bhati, and Parhasarathy (2003)	Maryland: Break the Cycle Initiative for drug offenders	Retrospective quasi-experimental design of offenders (n = 5,600) drawn pre- and postintervention	Null overall effects mixed with significant (but small) treatment effects	No
Hawken and Kleiman (2011)	Washington: Intensive Supervision Program	Randomly assigned offenders to treatment (n = 40) or control (n = 40) conditions	Null effects mixed with unclear treatment effects (limited information reported) and insufficient statistical power	No
Kilmer, Nicosia, Heaton, and Midgette (2013)	South Dakota: 24/7 Sobriety Project	County-level time series design comparing treatment and control counties	Null effects mixed with small treatment effect of up to a maximum 12% on others	No
Kubas, Kayabas, and Vachal (2015)	North Dakota: 24/7 Sobriety Project	Single-group design tracking outcomes of 1,893 DUI offenders	No treatment effects estimated because no comparison group was assessed	No
Kunkel and White (2013)	Arkansas: SWIFT courts, based explicitly on HOPE model	Propensity score matching on 6 factors (n = 54 treatment and 54 controls; nearly all substance abusers)	Treatment effects between 4% and 5% on recidivism; insufficient statistical power to detect significance	No
Loudenburg, Drube, and Leonardson (2013)	South Dakota: 24/7 Sobriety Program	Quasi-experimental with nonparticipating probationers matched on DUI level (n = 4,009 total offenders)	Mixed null effects, with largest treatment effects of up to 6% on DUI reoffending	No
Shannon (2013); Shannon, Hulbig, Birdwhistell, Newell, and Neal (2015)	Kentucky: SMART Probation program, based explicitly on HOPE model	Quasi-experimental design with 309 treatment and 300 lower risk control subjects	Mix of null and small treatment effects; insufficient data to make comparisons on revocations	No
Snell (2007)	Texas: Sanctions Court for felony probationers	Quasi-experimental design with an unspecified sample size with control group drawn from time period prior to intervention	Null effects mixed with a significant treatment effect on recidivism of 4%	No
Steiner, Makarios, Travis, and Meade (2012)	Ohio: Community-based sanctions program for parolees	Analyses of 995 parole violators' behavior over 1 year	Consistent deterrent effects of various magnitude of sanctions on rearrest	Yes
Stevens-Martin (2014)	Texas: SWIFT courts, based explicitly on HOPE model	Single-group posttest on 59 felony offenders	No treatment effects estimated because no pretest was given and no comparison group was assessed	No
Virginia Criminal Sentencing Commission (2014)	Immediate Sanction Probation Pilot Project	Pilot project with 64 offenders	No treatment effects estimated because no pretest was given and no comparison group was assessed	No
Weinrath, Doerksen, and Watts (2015)	Manitoba: COHROU Program	Single-group pretest-posttest design (n = 409)	No treatment effects could be estimated because no comparison group was assessed	No

Note. HOPE = Hawaii's Opportunity Probation with Enforcement; PACE = Probation Accountability with Certain Enforcement; DUI = driving under the influence; SMART = Supervision Motivation Accountability Responsibility and Treatment; COHROU = Criminal Organization High Risk Offender Unit.

Nonetheless, we have since consulted various writings by Hawken and Kleiman to track down SCF studies they have cited as lending support to the HOPE model. We

may not have found all studies, but we are able to report 20 in Table 1. Recall that our meta-analysis revealed a mean overall effect size of .070, with results other than Hawaii at .033. We wanted to see whether these SCF studies presented a credible challenge to these findings. To do so, a study would have to answer these three questions positively: (1) Was there a consistent “treatment effect” that was appreciably different from the .030 to .070 range that we identified in our study? (2) If so, could that treatment effect be designated as “statistically significant”? This speaks to the problem of low statistical power in certain studies where small samples are more likely to lead to inferential errors and false positives. (3) If the treatment effect was statistically significant, was it produced by a study with a research design that had an appropriate comparison group? In the least, it would have to use a quasi-experimental design in which the groups were roughly equivalent or in which the authors controlled statistically the key differences between the groups. All three of these conditions would need to be met—and would need to be met for a rather large group of studies—before any substantively meaningful challenge could be made to our core findings or conclusions.

Based on these three conditions, the far right column in Table 1 reports our assessment of the SCF studies. As can be seen, taken together, these evaluations do not offer a credible challenge to the central conclusion of our meta-analysis: HOPE-style supervision is unlikely to achieve meaningful change in offenders’ behavior. To be fair, the study by Steiner, Makarios, Travis, and Meade (2012) demonstrated rather consistent deterrent effects of sanctions in their analysis of Ohio parolees. It is also important to note that this study was strong methodologically. Few of the others were. These studies are riddled with such a wide array of methodological problems that a careful read of this work should inspire skepticism—not confidence—in their collective results. In looking closely, we find that there are several studies that failed to include a comparison group (so “treatment effects” could not even be estimated), as well as those plagued with basic research design problems such as randomization failures and insufficient statistical power. There is also the widespread presence of null results across these studies—null results that tend to be conveniently ignored by HOPE advocates as if they do not exist (Hawken, 2016). But they do. And even when statistically significant, the magnitude of these effects is at times misinterpreted— both

by the scholars who produced them and by the HOPE advocates who repeat them—and is typically small (usually with an ARR effect size below .10) and roughly in the neighborhood of what were found in the meta-analysis of HOPE studies (Cullen et al., 2016). Given our read of this literature, it therefore strikes us as more than a bit of a stretch to say only that “there are several studies demonstrating positive results” (Hawken, 2016, p. 1231).

Lattimore and Colleagues’ Demonstration Field Experiment (DFE)

RCT studies across multiple sites are rare in corrections (see, e.g., Petersilia & Turner, 1993). When such “gold standard” evaluations occur and especially if they produce consistent results, they yield the best evidence available on a program’s effectiveness. Indeed, the initial favorable Hawaiian results for Project HOPE earned the HOPE model a multisite RCT replication called a “demonstration field experiment”—or “DFE”—funded by the National Institute of Justice and the Bureau of Justice Assistance. The results have been reported in the *Criminology & Public Policy* article (Lattimore, MacKenzie, et al., 2016) and in a summary of findings (Lattimore, Dawes, et al., 2016).

Initiated in August 2012 and focusing on high-risk and medium-risk probationers, Project HOPE was implemented in four disparate locations: Saline County (Benton), Arkansas; Essex County (Salem), Massachusetts; Clackamas County (Oregon City), Oregon; and Tarrant County (Ft. Worth), Texas. Participants had to meet the eligibility criterion of having “at least one year of probation remaining,” with HOPE offenders compared with offenders receiving “probation as usual” (PAU) (Lattimore, Dawes, et al., 2016, p. 2). Considerable effort was devoted to ensuring fidelity to the HOPE model, with each site receiving “technical assistance and the presence of full-time, on-site HOPE program coordinators” (p. 8). Most important, it appears that Angela Hawken and Judge Alm were involved in helping to set up the HOPE programs and in ensuring program fidelity. When we (Cullen) contacted Pamela Lattimore for specific details, she was kind enough to provide the following summary (personal communication, November 21, 2016; see also Hawken, 2016, p. 1232; Lattimore, MacKenzie, et al., 2016, pp. 1130-1131):

Angela [Hawken] and her team were deeply involved in setting up the four DFE

sites. They and Judge Alm made repeated visits to the sites. The fidelity measures that I mentioned in my [ASC] presentation and that we report in the article were established in conjunction with Angela and her team and the data were actually collected by the site Program Coordinators who were hired by the sites but were “supervised” by Angela and her team. We had monthly calls with Angela and her team, along with NIJ and BJA, during the first 4+ years of the DFE so we were well aware of their contacts with the teams. We also made multiple site visits (Gary [Zajac] led those) and had contact (daily) with the sites through our on-site research coordinators. So, I am confident that we had a good view of what was going on with the sites.

Let us be clear what was riding on these results. The studies were implemented with a high degree of expertise and fidelity to the HOPE model—with direct assistance from Hawken and Alm. Researcher-led demonstration projects also tend to produce larger positive effects than routine practice programs (Lipsey, 2003). In a sense, the decks were stacked to give HOPE every chance to succeed. If it did not, then confidence in the model as an intervention that could be rolled out to tens of agencies would be undermined. In all likelihood, this is why, as noted, Hawken et al. (2016) recognized the “importance of the HOPE DFE (in four states)” (p. 27).

Alas, bad news was delivered from Lattimore and colleagues’ DFE in the latter part of 2016 (for a summary of findings, see Lattimore, Dawes, et al., 2016). First, it was found that “recidivism outcomes were similar for those in HOPE programs and probation as usual (PAU)” (Lattimore, Dawes, et al., 2016, p. 1). Second, the results showed that “overall, there were *no significant differences* ($p < .05$) *between the HOPE and PAU groups in the likelihood of primary outcomes including arrest, probation revocation, or new conviction*” (Lattimore, Dawes, et al., 2016, p. 2, emphasis in original). HOPE probationers experienced lower recidivism arrests, but the difference was small—“0.8 compared with 0.7 ($p = .06$)” (Lattimore, Dawes, et al., 2016, p. 3). Thus, the stark conclusion of the DFE was as follows: “*it is unlikely that HOPE/SCF can produce lower recidivism or lower costs than PAU*” (Lattimore, Dawes, et al., 2016, p. 1, emphasis in original).

Explaining Failure: To Be or Not To Be HOPE

The Lattimore et al. study was a critical litmus test for Project HOPE. If the results had been positive, they would have been touted by advocates as proving the viability of the model. The consistent negative findings, however, cast substantial doubt that the HOPE approach to community supervision is likely to be effective across diverse agency settings. Heterogeneous results are common in evaluation research, so it would be surprising if the model did not work in some locations—especially if the comparison PAU in question was ineffective if not iatrogenic. But now the best available evidence suggests that HOPE is an intervention that, even when diligently implemented, is likely to fail. This is consequential because Project HOPE does not require just adding a new “program” but a systemic renovation of probation or parole that involves drug tests, the immediate imposition of graduate sanctions, the allocation of court and jail resources, and the replacement of a rehabilitative philosophy in favor of a deterrence-oriented (SCF) philosophy.

As the project’s main advocates, Alm, Hawken, and Kleiman could have viewed the Lattimore and colleagues’ RCT and called for a moratorium on HOPE’s further implementation, at least until more research could be done to figure out how to make it work. They did not. Instead, although divergent explanations were given for the DFE results, they ultimately concurred in downplaying the evaluation evidence and justified the continued spread of HOPE. All three wrote fascinating essays for the “HOPE Collection.” We present the central message of their rebuttals; readers should consult the essays for all details. In addition, we draw on Hawken and Kleiman’s presentations at the ASC meeting and a subsequent web posting by Kleiman.

Alm: It Did Not Work Because It Was Not HOPE

According to Judge Alm, the interventions in the four DFE (and Delaware RCT) sites failed because they did not include special elements of the HOPE model. Alm (2016) contended that HOPE was never a pure deterrence approach; in fact, he pointed out that initially he “had never heard of Mark Kleiman, Angela Hawken, or Daniel Nagin” (p. 1197). Instead, his ideas about probation supervision were based on his common-sense ideas about parenting his son. The first principle of HOPE was, in a sense, the

first principle of parenting: Provide a caring and supportive environment within which discipline is to be applied. In probation, the judge and officers thus must have a “therapeutic alliance” with offenders so as to develop quality relationships. Supervising officers also must be trained in good parenting, so to speak, by being taught evidence-based practices. Sanctions are the final, missing element to the HOPE recipe: High-risk probationers, the target group for the program, need these certain slaps on the wrist to be reminded that bad behavior has consequences and that they must develop an internal sense of accountability. “The HOPE strategy has always been much more than [SCF sanctions],” notes Alm. “HOPE is caring and therapeutic. HOPE is procedural justice in action with rehabilitation of the probationer being the goal” (Alm, 2016, p. 1202).

Such “therapeutic jurisprudence” (Bartels, 2016) is not that far apart from the Risk– Need–Responsivity (RNR) model proposed by Andrews, Bonta, Gendreau, and others (Bonta & Andrews, 2017). Alm notes that it was the combination of these elements that produced success in Hawaii and failure in the DFE sites. These programs did not succeed, Alm “came to the realization,” because “the evaluators . . . unfortunately have a basic misunderstanding of the HOPE strategy” (Alm, 2016, p. 1195). The problem? “They seemingly believe that HOPE is simply a sanctions-only program that is solely concerned with imposing jail sanctions on probationers with no interest in rehabilitation” (Alm, 2016, p. 1195). The difficulty, of course, is that Judge Alm cannot control the implementation of HOPE in jurisdictions where judicial empathy is lacking, staff training in treatment skills is absent, and the program is sold as an SCF deterrence intervention. If “misunderstanding” occurs among talented evaluators, which include noted criminologists, what will happen elsewhere? Is this not reason to be concerned?

Hawken and Kleiman: It Did Not Work Because It Was HOPE

Angela Hawken (2016) has offered a judicious response, noting that Lattimore and colleagues’ “DFE . . . is a high-quality study that will carry substantial weight” (p. 1231). For her, this study needs to be placed in the “broader growing literature concerning the HOPE strategy,” some of which she briefly summarizes and reports to have positive results (Hawken, 2016, p. 1231). As noted above, we differ in how we

would characterize this research. But Hawken (2016) certainly has enough grounds to argue, as she does, that “we should be open to testing new approaches and to identifying what works and under what conditions” (p. 1231). We would point out, however, the difference between conducting careful tests and the risk that Project HOPE will be taken to scale prematurely before all the evidence has arrived—something that is now in progress.

But why did the interventions fail in the four RCT sites? According to Hawken (2016), one possibility is that the agencies who self-selected into the DFE were “unusually well-functioning probation departments,” which made “it harder to demonstrate an effect” (p. 1232). Of course, the implication is that good agencies do not need HOPE and should continue their current supervision practices.

Hawken’s more stunning response, however, is that the four sites in the study were strictly mandated to use the HOPE model, which was then implemented with fidelity. Here, she parts company from Judge Alm who, as noted, critiqued the sites for not implementing the therapeutic side of the HOPE model. For Hawken, the generalizability of the HOPE model from Hawaii to the mainland was always unlikely. She believes that “all implementation is local” so that the unique features of each agency and community (e.g., system and resource differences, stakeholder preferences, and offender perceptions) must be considered when designing a program that works for them. HOPE was effective in Hawaii because Hawaiians had started with a “conversation” about how to implement SCF within their probation system (Hawken, 2016, p. 1238). In essence, Hawaiian HOPE may not succeed in agencies in Arkansas, Massachusetts, Oregon, and Texas unless the local circumstances in these places are taken into account.

Alas, individualizing interventions in each locale to improve their efficacy is a daunting challenge with uncertain empirical results. That aside, a key point should be recognized: Advocates are now giving us incompatible advice regarding how best to use Project HOPE in the future. For Alm, HOPE will fail unless implemented as designed, including a strong therapeutic component. For Hawken, HOPE will fail unless it abandons the rigid strictures of the Hawaiian HOPE model and is modified to build on the strengths and meet the needs of each agency. The lack of commitment to rehabilitation

in the DFE four sites did not seem to concern her. In the past, she has talked about “economizing on treatment” (Hawken & Kleiman, 2009, p. 4) and about using “behavioral triage”—that is, the allocation of treatment based not on risk level but on the failure to comply with supervision conditions (Hawken, 2010a).

Mark Kleiman sides with Hawken, echoing her points. Similar to Hawken, rehabilitation seems of secondary importance—not part of HOPE’s DNA as Alm would have it, but something to be given to those who ask for it or who fail a bunch of drug tests. There is no discussion of the crucial need for therapeutic alliances with offenders or training all officers in core correctional practices. Again, HOPE is not seen as a proven model of universal effectiveness. Rather, Kleiman (2016a) asserts that “SCF is an operating concept for corrections, not a specific program,” that must be “fit . . . to local conditions.” He paints a rosy picture of existing SCF evaluations. He sees studies favorable to SCF as awesome and those against it as sad. Thus, the evaluation of the Washington State program is referred to as “glowing” and the Hawaii HOPE program is called an “impressive success,” whereas the Lattimore et al. study is picked apart and said to go “far beyond the data” (Kleiman, 2016a; see Gottfredson, 1979). We have made it clear what we think of the relative merits of these two sets of findings; readers can decide on their own who is correct (see also Lattimore, 2017). Finally, Kleiman repeats Hawken’s view that HOPE’s effects may have been limited because the four comparison agencies in Lattimore and colleagues’ DFE were well run. Because “supervision as usual” (SAU) in the run-of-the-mill agency is inadequate, HOPE supposedly would fare much better if compared with them. Of course, there is another way of looking at this claim. Why bother to implement the HOPE/SCF model at all? Why not just look at the high-performance SAU agencies and model them?

For Kleiman, SCF is a proven principle to be used in community and prison settings to modify offender behavior. Still, he hedges his bet considerably, almost to the point of making SCF nonfalsifiable:

Even if, on theoretical grounds, it seems likely that some version of SCF could be designed that would outperform SAU under most circumstances, it does not follow that any specific SCF program will outperform the SAU actually in place in any specified jurisdiction for any specified class of offenders. Conversely, all that

can be learned from an SCF program that fails to improve on previous practice is that the program in question, applied as and where it was, did not succeed.

(Kleiman, 2016b, p. 1189)

His solution is to pursue the “vital research task” of fine-tuning this approach so as to allow it to work more and more effectively (Kleiman, 2016b, p. 1191). In this context, he suggests that it makes little sense to try “to answer the unanswerable and impractical question, ‘Does it work?’” (Kleiman, 2016b, p. 1191).

By contrast, the “what works” question is precisely what we believe must be addressed *before telling jurisdictions to establish SCF/HOPE programs*. We are not inalterably opposed to further testing, and Kleiman (2016b), in particular, has raised important factors to consider (e.g., limiting supervision conditions, use of positive incentives and nonjail sanctions, and focusing only on offenders who have a record of persistent supervision failure). We are troubled, however, by any evangelizing efforts to sell the HOPE-like model as a *new-and-improved, empirically proven supervision paradigm that every agency should adopt*.

Right now—especially in light of the DFE HOPE trials which report the best data available—all that can be said is that the model might or might not reduce recidivism, and, if it does, its effects could be large but probably will be small and may decay over time. Its advocates cannot advise agencies considering this intervention whether to comply with HOPE principles, with its strong therapeutic commitment, or to implement individualized versions of SCF principles. Indeed, they cannot even tell agencies that are fairly well run whether SCF/HOPE will be more effective than their current SAU. In this sea of uncertainty in which key empirical facts remain unknown, we believe that it is unduly risky to disseminate this model widely, especially if it is not accompanied by a bright label warning that it may not work (Cullen et al., 2014).

Conclusion: Project HOPE in a Post-Factual World

The advocates of the project were cruising down Highway HOPE when the bad news from the Lattimore and colleagues’ study arrived. They could have looked at these findings, hit the brakes, and pulled over to the side of the road to contemplate where they should next head. Instead, they put their foot on the gas pedal, turned up the stereo

to drown out the noise, and accelerated down the road at a high speed. Undaunted by empirical reality, they did not consider that a cliff might lie ahead and that HOPE's correctional future might end in a crash reminiscent of *Thelma and Louise*.

Indeed, Judge Alm has retired so as to "take on a new role, based in Washington, DC, as a legal consultant to the states, U.S. Department of Justice and Congress on the implementation of HOPE" (Bartels, 2016, p. 30). Alm chose DC because, as he states, "it'll be easier to promote the HOPE project from there, and there may be a HOPE institute that will get funded" (Yamane, 2016). Angela Hawken's biographical description on the website of New York University's Marron Institute notes her role in testing SCF strategies and reports that she "works with agencies in 14 states that are implementing and evaluating SCF in community corrections and in custody" (see <http://marroninstitute.nyu.edu/people/angela-hawken1>). And nothing that Mark Kleiman has said so far hints that his confidence in SCF has been diminished in any way. Perhaps speaking for the group, Judge Alm (2016) has advised those considering HOPE as follows: "Don't be discouraged by nay-sayers" (p. 1212).

We disagree, especially because a lot is at stake. The field needs those who invent programs and then advocate for their implementation (Sherman, 2011). But the field also needs people like us—informed naysayers who want to see a lot of persuasive evidence before offenders are subjected to yet another bad intervention that is intuitively appealing yet empirically vacuous. Remember, both prisons and boot camps were invented with the best of intentions (Cullen, Blevins, Trager, & Gendreau, 2005; Rothman, 1971). The continued implementation of Project HOPE means that this program now exists in a post-factual world. We do not believe that there has ever been sufficient positive evidence to justify the widespread dissemination of the program. More disquieting, after the publication of Lattimore and colleagues' DFE, there is now strong negative evidence that, even when implemented with fidelity, the HOPE model may not be effective.

It may be too much to expect HOPE advocates to admit the obvious—that HOPE cannot be said to work. Importantly, they are not only advocates but also HOPE experts capable of advising others on the nature and implementation of the program. Policy

makers and agencies look to them for answers on how best to reform community supervision. And as Kahneman (2011) observed, “an expert worthy of the name is expected to display high confidence” (p. 263). For the rest of us, however, it is our collective obligation to be criminological curmudgeons—to caution that Project HOPE lacks consistent empirical support and thus is likely to be yet another in a long line of deterrence-oriented interventions that, in the end, offers false hope (Duriez et al., 2014). The time may have come to consider a HOPE moratorium.

Although calling for more research risks triteness, the fact remains that this is precisely what is needed: a science of community supervision. We believe that the existing evidence regarding effective supervision favors a human services approach, officers’ building quality relationships with offenders, and the use of cognitive-behavioral techniques and core correctional practices in meetings with supervisees (Bonta & Andrews, 2017; Cullen, Jonson, & Mears, 2017; Paparozzi & Gendreau, 2005). We also are persuaded that probation and parole should move away from the zero-tolerance enforcement of supervision conditions. The alternative option—“zero-based condition setting”—would eliminate boilerplate lists of conditions and instead impose only the few conditions that relate to an offender’s risk for recidivating (Corbett, 2015, p. 1729; see also Doherty, 2016; Klingele, 2013). Still, we must admit that even with more than 4.6 offenders on probation and parole (Kaeble & Bonczar, 2016), it is inexplicable that the existing knowledge on supervision effectiveness is so limited. Even if Project HOPE is likely not the best option, much more research is needed to find out what is.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

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

References

- Abramowitz, A. J., & O'Leary, S. G. (1990). Effectiveness of delayed punishment in an applied setting. *Behavior Therapy*, *12*, 231-239.
- Alm, S. S. (2016). HOPE probation: Fair sanctions, evidence-based principles, and therapeutic alliances. *Criminology & Public Policy*, *15*, 1195-1214.
- American Society of Criminology. (2016, November 17). Session 539—Does swift, certain, and fair work? Findings from two RCT's. In *Annual meeting program* (pp. 181-182). Columbus, OH: Author.
- Bartels, L. (2016). Looking at Hawaii's Opportunity with Probation Enforcement (HOPE) pro-gram through a therapeutic jurisprudence lens. *QUT Law Review*, *16*(3), 30-49.
- Bonta, J., & Andrews, D. A. (2017). *The psychology of criminal conduct* (6th ed.). New York, NY: Routledge.
- Carns, T. W., & Martin, S. (2011). *Anchorage PACE: Probation Accountability with Certain Enforcement—A preliminary evaluation of the Anchorage Pilot PACE Project*. Anchorage: Alaska Judicial Council. Retrieved from <http://www.ajc.state.ak.us/reports/pace2011.pdf>
- Caulkins, J., & Dupont, R. (2010). Is 24/7 sobriety a good goal for repeat driving under the influence (DUI) offenders? *Addiction*, *105*, 575-577.
- Cavanaugh, D. P., Godfrey, K., & Harney, S. (1994). *A formative and process evaluation of the Multnomah County Department of Community Corrections' structured sanctioning process*. Los Angeles, CA: BOTEK Analysis Corporation.
- Cavanaugh, D. P., & Harrell, A. V. (1995). *Evaluation of the Multnomah County drug testing and evaluation program*. Washington, DC: National Institute of Justice, U.S. Department of Justice.
- Clear, T. R., & Frost, N. A. (2014). *The punishment imperative: The rise and failure of mass incarceration in America*. New York: New York University Press.
- Cook, P. J. (2016). Behavioral science critique of HOPE. *Criminology & Public Policy*, *15*, 1155-1161.
- Corbett, R. P., Jr. (2015). The burdens of leniency: The changing face of probation. *Minnesota Law Review*, *99*, 1697-1733.
- Cullen, F. T., Blevins, K. R., Trager, J. S., & Gendreau, P. (2005). The rise and fall of

- boot camps: A case study in common-sense corrections. *Journal of Offender Rehabilitation*, 40(3-4), 53-70.
- Cullen, F. T., Jonson, C. L., & Mears, D. P. (2017). Reinventing community corrections. In M. Tonry & D. S. Nagin (Eds.), *Crime and justice: A review of research, Vol. 46—Reinventing American criminal justice* (pp. 27-93). Chicago, IL: University of Chicago Press.
- Cullen, F. T., Manchak, S. M., & Duriez, S. A. (2014). Before adopting Project HOPE, read the warning label: A rejoinder to Kleiman, Kilmer, and Fisher's comment. *Federal Probation*, 78(2), 75-77.
- Cullen, F. T., Pratt, T. C., & Turanovic, J. J. (2016). It's hopeless: Beyond zero-tolerance supervision. *Criminology & Public Policy*, 15, 1215-1227.
- Cullen, F. T., Pratt, T. C., & Turanovic, J. J. (2017). The failure of swift, certain, and fair supervision: Choosing a more hopeful future. *Perspectives*, 41(3), 66-78.
- Cullen, F. T., Smith, P., Lowenkamp, C. T., & Latessa, E. J. (2009). Nothing works revisited: Deconstructing Farabee's "Rethinking Rehabilitation." *Victims & Offenders*, 4, 101-123.
- DeVall, K., Lanier, C., & Hartmann, D. (2015). *Evaluation of Michigan's Swift & Sure Sanctions Probation Program*. Retrieved from <http://courts.mi.gov/Administration/admin/op/problem-solving-courts/Documents/SS-Eval.pdf>
- Doherty, F. (2016). Obey all laws and be good: Probation and the meaning of recidivism. *Georgetown Law Journal*, 104, 291-354.
- Dunford, F. W., Huizinga, D., & Elliott, D. S. (1990). The role of arrest in domestic assault: The Omaha police experiment. *Criminology*, 28, 183-206.
- Dupont, R. L., & Wish, E. D. (1992). Operation tripwire revisited. *Annals of the American Academy of Political and Social Science*, 521, 91-111.
- Duriez, S. A., Cullen, F. T., & Manchak, S. J. (2014). Is Project HOPE creating a false sense of hope? A case study in correctional popularity. *Federal Probation*, 78(2), 57-70.
- Farabee, D. (2005). *Rethinking rehabilitation: Why can't we reform our criminals?* Washington, DC: American Enterprise Institute Press.

- Gilbert, N. (1997). Advocacy research and social policy. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 22, pp. 101-148). Chicago, IL: University of Chicago Press.
- Gottfredson, M. R. (1979). Treatment destruction techniques. *Journal of Research in Crime and Delinquency*, 16, 39-54.
- Grasmick, H. G., & Bryjak, G. J. (1980). The deterrent effect of perceived severity of punishment. *Social Forces*, 59, 471-491.
- Grommon, E., Cox, S. M., Davidson, W. S., II, & Bynum, T. S. (2013). Alternative models of instant drug testing: Evidence from an experimental trial. *Journal of Experimental Criminology*, 9, 145-168.
- Hamilton, Z., Campbell, C. M., van Wormer, J., Kigerl, A., & Posey, B. (2016). The impact of swift and certain sanctions: An evaluation of Washington State's policy for offenders on community supervision. *Criminology & Public Policy*, 15, 1009-1072.
- Harrell, A., & Roman, J. (2001). Reducing drug use and crime among offenders: The impact of graduated sanctions. *Journal of Drug Issues*, 31, 207-231.
- Harrell, A., Roman, J., Bhati, A., & Parhasarathy, B. (2003). *The impact evaluation of the Maryland Break the Cycle Initiative*. Washington, DC: Urban Institute Justice Policy Center.
- Hawken, A. (2010a). HOPE for probation: How Hawaii improved behavior with high-probability, low-severity sanctions. *Journal of Global Drug Policy and Practice*, 4(3). Retrieved from <http://hopehawaii.net/assets/the-journal-of-global-drug-policy—practice.pdf>
- Hawken, A. (2010b). The message from Hawaii: HOPE for probation. *Perspectives*, 34(3), 36-49.
- Hawken, A. (2016). All implementation is local. *Criminology & Public Policy*, 15, 1229-1239. Hawken, A., & Kleiman, M. (2009). *Managing drug involved probationers with swift and certain sanctions: Evaluating Hawaii's HOPE*. Washington, DC: National Institute of Justice, U.S. Department of Justice.
- Hawken, A., & Kleiman, M. A. R. (2011). *Washington intensive supervision program evaluation report*. Retrieved from <http://www.scfcenter.org/resources/Research>

/201112%20 Washington%20Intensive%20Supervision%20Program%
20Evaluation%20Report.pdf

Hawken, A., Kulick, J., Smith, K., Mei, J., Zhang, Y., Jarman, S., & Vial, T. (2016).

HOPE II: A follow-up to Hawaii's HOPE evaluation. Washington, DC: National Institute of Justice, U.S. Department of Justice.

Hyatt, J. M., & Barnes, G. C. (2017). An experimental evaluation of the impact of intensive supervision on the recidivism of high-risk probationers. *Crime & Delinquency*, 63, 3-38.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2, 696-701.

Kaeble, D., & Bonczar, T. P. (2016). *Probation and parole in the United States, 2015*. Washington, DC: Bureau of Justice Statistics, U.S. Department of Justice.

Kahneman, D. (2011). *Thinking, fast and slow*. New York, NY: Farrar, Straus, and Giroux.

Kilmer, B., Nicosia, N., Heaton, P., & Midgette, G. (2013). Efficacy of frequent monitoring with swift, certain, and modest sanctions for violations: Insights from South Dakota's 24/7 Sobriety Project. *American Journal of Public Health*, 103(1), e37-e43.

Kleiman, M. A. R. (2016a). Experimental findings on HOPE probation. *The Reality-Based Community*. Retrieved from <http://www.samefacts.com/2016/11/everything-else/experim-ental-findings-on-hope-probation/>

Kleiman, M. A. R. (2016b). Swift-certain-fair: What do we know now, and what do we need to know? *Criminology & Public Policy*, 15, 1185-1193.

Kleiman, M. A. R., Kilmer, B., & Fisher, D. T. (2014). Response to Stephanie A. Duriez, Francis T. Cullen, and Sarah M. Manchak: Theory and evidence on the swift-certain-fair approach to enforcing conditions of community supervision. *Federal Probation*, 78(2), 71-73.

Klinge, C. (2013). Rethinking the use of community supervision. *Journal of Criminal Law and Criminology*, 103, 1015-1069.

Klinge, C. (2015). What are we hoping for? Defining purpose in deterrence-based correctional programs. *Minnesota Law Review*, 99, 1631-1662.

- Kubas, A., Kayabas, P., & Vachal, K. (2015). *Assessment of the 24/7 Sobriety Program in North Dakota: Participant behavior during enrollment*. Fargo: Upper Great Plains Transportation Institute, North Dakota State University.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd and enlarged ed.). Chicago, IL: University of Chicago Press.
- Kunkel, T. L., & White, M. T. (2013). *Arkansas SWIFT courts: Implementation assessment and long-term evaluation plan*. Williamsburg, VA: National Center for State Courts.
- Lattimore, P. K. (2017). Response to the comments of Mark Kleiman in regard to "Outcome findings from the HOPE demonstration field experiment" by Lattimore, MacKenzie, Zajac, et al. *Criminology & Public Policy*, Vol. 15(4), 1-39 (Unpublished manuscript). Research Triangle Park, NC: RTI International.
- Lattimore, P. K., Dawes, D., Tueller, S., MacKenzie, D. L., Zajac, G. X., & Arsenault, E. (2016). *Summary findings from the national evaluation of the Honest Opportunity Probation with Enforcement demonstration field experiment: The HOPE DFE evaluation*. Research Triangle Park, NC: RTI International and the Penn State Justice Center for Research.
- Lattimore, P. K., MacKenzie, D. L., Zajac, G. X., Dawes, D., Arsenault, E., & Tueller, S. (2016). Outcome findings from the HOPE demonstration field experiment: Is swift, certain, and fair an effective supervision strategy? *Criminology & Public Policy*, 15, 1103-1139.
- Lehrer, J. (2010, December 13). The truth wears off: Is there something wrong with the scientific method? *The New Yorker*, pp. 52-57.
- Lerman, D. C., & Vorndran, C. M. (2002). On the status of knowledge for using punishment: Implications for treating behavior disorders. *Journal of Applied Behavior Analysis*, 35, 431-464.
- Lipsey, M. W. (2003). The confounded moderators in meta-analysis: Good, bad, and ugly. *Annals of the American Academy of Political and Social Science*, 587, 69-81.
- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims & Offenders*, 4, 124-147.

- Lipsey, M. W., & Cullen, F. T. (2007). The effectiveness of correctional rehabilitation: A review of systematic reviews. *Annual Review of Law and Social Science*, 3, 297-320.
- Loudenburg, R., Drube, G., & Leonardson, G. (2013). *South Dakota 24/7 Sobriety Program evaluation findings report*. Salem, SD: Mountain Plains Evaluation LLC.
- MacKenzie, D. L. (2006). *What works in Corrections: Reducing the criminal activities of offenders and delinquents*. New York, NY: Cambridge University Press.
- MacKenzie, D. L., & Farrington, D. P. (2015). Preventing future offending of delinquents and offenders: What have we learned from experiments and meta-analyses? *Journal of Experimental Criminology*, 15, 565-595.
- McGuire, J. (2013). "What works" to reduce re-offending: 18 years on. In L. A. Craig, L. Dixon, & T. A. Gannon (Eds.), *What works in offender rehabilitation: An evidence-based approach to assessment and treatment* (pp. 20-49). Chichester, UK: Wiley-Blackwell.
- Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations* (Norman W. Storer, Ed.). Chicago, IL: University of Chicago Press.
- O'Connell, D., Brent, J. J., & Visher, C. A. (2016). Decide your time: A randomized trial of a drug testing and graduated sanctions program for probationers. *Criminology & Public Policy*, 15, 1073-1102.
- Oleson, J. C. (2016). HOPE spring eternal: New evaluations of correctional deterrence. *Criminology & Public Policy*, 15, 1163-1183.
- Paparozzi, M. A., & Gendreau, P. (2005). An intensive supervision program that worked: Service delivery, professional orientation, and organizational supportiveness. *The Prison Journal*, 85, 445-466.
- Paternoster, R. (1989). Decisions to participate in and desist from four types of common delinquency: Deterrence and the rational choice perspective. *Law & Society Review*, 23, 7-40.
- Paternoster, R., & Bachman, R. (2013). Perceptual deterrence theory. In F. T. Cullen & P. Wilcox (Eds.), *The Oxford handbook of criminological theory* (pp. 649-671). New York, NY: Oxford University Press.
- Pearsall, B. (2014). Replicating HOPE: Can others do as well as Hawaii? *National*

Institute of Justice Journal, 273, 36-41.

- Petersilia, J., & Turner, S. (1993). Intensive probation and parole. In M. Tonry (Ed.), *Crime and justice: A review of research* (Vol. 17, pp. 281-335). Chicago, IL: University of Chicago Press.
- Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L. E., & Madensen, T. D. (2006). The empirical status of deterrence theory: A meta-analysis. In F. T. Cullen, J. P. Wright, & K. R. Blevins (Eds.), *Advances in criminological theory, Vol. 15—Taking stock: The status of criminological theory* (pp. 367-395). New Brunswick, NJ: Transaction Publishers.
- Pratt, T. C., & Turanovic, J. J. (in press). Celerity and deterrence. In D. S. Nagin, F. T. Cullen, & C. L. Jonson (Eds.), *Advances in criminological theory—Deterrence, choice, and crime: Contemporary perspectives*. New York, NY: Routledge.
- Rothman, D. J. (1971). *The discovery of the asylum: Social order and disorder in the new republic*. Boston, MA: Little, Brown.
- Shannon, L. M. (2013). *Kentucky SMART Probation Program: Year one report*. Frankfort, KY: Administrative Office of the Courts.
- Shannon, L. M., Hulbig, S. K., Birdwhistell, S., Newell, J., & Neal, C. (2015). Implementation of an enhanced probation program: Evaluating process and preliminary outcomes. *Evaluation and Program Planning*, 49, 50-62.
- Sherman, L. W. (2011). Criminology as invention. In C. Hoyle & L. Bosworth (Eds.), *What is criminology?* (pp. 423-439). Oxford, UK: Oxford University Press.
- Snell, C. (2007). *Fort Bend County Community Supervision and Corrections Special Sanctions Court Program*. Houston, TX: University of Houston Downtown.
Retrieved from http://s3.amazonaws.com/static.texastribune.org/media/documents/SANCTIONS_COURT_FINAL_REPORT.pdf?preview
- Steiner, B., Makarios, M. D., Travis, L. F., III, & Meade, B. (2012). Examining the effects of community-based sanctions on offender recidivism. *Justice Quarterly*, 29, 229-257.
- Stevens-Martin, K. D. (2014). The supervision with immediate enforcement court: Tackling the issue of high-risk offenders and chronic probation violators. *Corrections Today*, 76(5), 74-77.

- Virginia Criminal Sentencing Commission. (2014). *Immediate Sanction Probation Pilot Project*. Richmond, VA: Author. Retrieved from [http://www.vcsc.virginia.gov/2016meetings/ Immediate%20Sanction%20Probation %20Evaluation%20Plan%20\(Apr%204%20 2016\)%20HANDOUT.pdf](http://www.vcsc.virginia.gov/2016meetings/Immediate%20Sanction%20Probation%20Evaluation%20Plan%20(Apr%204%202016)%20HANDOUT.pdf)
- Weinrath, M., Doerksen, M., & Watts, J. (2015). The impact of an intensive supervision program on high-risk offenders: Manitoba's COHROU Program. *Canadian Journal of Criminology and Criminal Justice*, 57, 253-288.
- Wilson, J. Q. (1997). Making justice swifter. *City Journal*, 7(4), 30-35.
- Wright, J. P., & Cullen, F. T. (2001). Parental efficacy and delinquent behavior: Do control and support matter? *Criminology*, 39, 601-629.
- Yamane, M. (2016, August 29). Circuit Court judge, HOPE founder Alm retires, but work isn't over. *KHON 2*. Retrieved from <http://khon2.com/2016/08/29/circuit-court-judge-hope-founder-alm-retires-but-work-isnt-over/>

Author Biographies

Francis T. Cullen (PhD, Columbia University) is Distinguished Research Professor Emeritus and a senior research associate in the School of Criminal Justice at the University of Cincinnati. His recent works include *Communities and Crime: An Enduring American Challenge*, *Criminological Theory: Past to Present*, and *Environmental Corrections: A New Paradigm for Supervising Offenders in the Community*. His current research interests are in correctional policy, theoretical criminology, and the organization of criminological knowledge. He is a past president of both the American Society of Criminology and the Academy of Criminal Justice Sciences.

Travis C. Pratt (PhD, University of Cincinnati) is a fellow with the University of Cincinnati Corrections Institute. He is the author of *Addicted to Incarceration: Corrections Policy and the Politics of Misinformation in the United States*. His research focuses on criminological theory and correctional policy, and his work has been published in a variety of journals, including *Criminology*, *Criminology & Public Policy*, *Journal of Research in Crime and Delinquency*, *Justice Quarterly*, *Journal of Quantitative Criminology*, and *Crime and Justice: A Review of Research*. In 2006, he

received the Ruth Shonle Cavan Young Scholar Award from the American Society of Criminology.

Jillian J. Turanovic (PhD, Arizona State University) is an assistant professor in the College of Criminology and Criminal Justice at Florida State University. Her research focuses on violent victimization and offending, correctional policy, and the collateral consequences of incarceration. Her recent work has appeared in *Criminology*, the *Journal of Quantitative Criminology*, and *Justice Quarterly*. She is currently completing work on *Thinking About Victimization: Context and Consequences*.

Leah Butler (MA, Ohio University) is a PhD student in the School of Criminal Justice at the University of Cincinnati. She has published works on bystander intervention to prevent sexual assaults and on public opinion about capital punishment. Her research interests include ways to improve contemporary criminological theory, correctional effectiveness, and victimology.