

Moving On: An Outcome Evaluation of a Gender-Responsive, Cognitive-Behavioral Program for Female Offenders

Authors

Grant Duwe, Ph.D.
Research Director
Email: grant.duwe@state.mn.us

Valerie Clark, Ph.D.
Research Analyst Specialist
Email: valerie.clark@state.mn.us



1450 Energy Park Drive, Suite 200
St. Paul, Minnesota 55108-5219
651/361-7200
TTY 800/627-3529
www.doc.state.mn.us
July 2015

This information will be made available in alternative format upon request.
Printed on recycled paper with at least 10 percent post-consumer waste

Research Summary

We used a quasi-experimental design to evaluate the effectiveness of Moving On, a gender-responsive, cognitive-behavioral program designed for female offenders. Between 2001 and 2013, there were two distinct periods in which Moving On was administered with, and without, fidelity among female Minnesota prisoners. To determine whether program integrity matters, we examined the performance of Moving On across these two periods. Using multiple comparison groups, we found that Moving On significantly reduced two of the four measures of recidivism when it was implemented with fidelity. The program did not have a significant impact on any of the four recidivism measures, however, when it operated without fidelity.

The growth of the “what works” literature and the emphasis on evidence-based practices have helped foster the notion that correctional systems can improve public safety by reducing recidivism. Given that Moving On’s success hinged on whether it was delivered with integrity, our results show that correctional practitioners can take an effective intervention and make it ineffective. Providing offenders with evidence-based interventions that lack therapeutic integrity not only promotes a false sense of efficacy, but it also squanders the limited supply of programming resources available to correctional agencies. The findings suggest that ensuring program integrity is critical to the efficient use of successful interventions that deliver on the promise of reduced recidivism.

Introduction

Cognitive behavioral treatment (CBT) is one of the most effective correctional tools for reducing recidivism (Allen, MacKenzie, and Hickman, 2001; Lipsey, Chapman, and Landenberger, 2001; Lipsey, Landenberger, and Wilson, 2007; Pearson, Lipton, Cleland, and Yee, 2002; Wilson, Bouffard, and MacKenzie, 2005). CBT includes all programs that address the link between dysfunctional thought processes and harmful behaviors through timely reinforcements and punishments, as well as role-playing and skill-building exercises. These programs aim to improve decision-making and problem-solving skills, and to teach individuals how to manage various forms of outside stimuli. CBT can reduce recidivism by targeting an array of risk factors, including general antisocial cognition and chemical dependency.

While many studies have documented CBT's effectiveness for reducing recidivism, multiple meta-analyses have revealed that the magnitude of this effect can vary widely (e.g., Pearson et al., 2002; Wilson, Bouffard, and MacKenzie, 2005). Researchers have suggested that this variability in effectiveness may be due in part to the implementation fidelity of CBT programs (Lipsey and Cullen, 2007; Lowenkamp, Latessa, and Smith, 2006; Palmer, 1995). That is, CBT programs designed in accordance with established principles of effective correctional interventions that maintain integrity upon implementation should be more effective than the same or similar programs that deviate too far from their original designs and compromise evidence-based program elements (Andrews and Dowden, 2005; Gendreau, Goggin, and Smith, 1999; Lowenkamp et al., 2006). Despite wide acceptance that program integrity is an important piece of effective correctional programs, few studies have examined the link between

program integrity and recidivism. The present study addresses this deficit in the literature with a quasi-experimental design that compares recidivism outcomes of CBT program participants when a program was and was not implemented as designed.

Given that males account for a large majority of all correctional populations, most research on CBT's effectiveness have focused on programs that commonly or exclusively treat males. In addition to examining the link between program integrity and recidivism, the present study makes another contribution to the literature by focusing on a CBT program designed exclusively for women offenders: *Moving On: A Program for At-Risk Women* (Van Dieten, 2010). To date, there is only one outcome evaluation of Moving On, and it has some methodological shortcomings that the present study overcomes.

Effective Interventions and Program Fidelity

The criminal justice system has amassed a library of research on how to deal effectively with crime and the individuals that commit crimes. Criminal justice practitioners can now reference a large body of empirical evidence on best practices in every field from policing, to the courts, to corrections. In corrections, practitioners have increasingly adopted the *principles of effective interventions* outlined by Andrews, Bonta, and Hoge (1990) to design programs and guide facility operations (see also Gendreau, 1996; Gendreau and Andrews, 1990).

In addition to the acceptance of CBT as one of the preferred methods of offender intervention, the *principles* outlined by Andrews et al. (1990) also hold that interventions should be matched to an offender's risk of reoffending, criminogenic needs, and responsivity issues (see Gendreau, French, and Gionet, 2004). This risk-need-responsivity (RNR) model calls for offender risk to be measured using actuarial risk

assessment tools that have been validated and normed (Andrews and Bonta, 2010). The most intensive programs—generally measured by total length and number of hours—should be reserved for individuals rated as high-risk (Sperber, Latessa, and Makarios, 2013). Criminogenic needs are individual characteristics that increase the risk of offending behaviors (Latessa and Lowenkamp, 2005). Static needs (e.g., prior criminal record, age) cannot be changed through interventions, while dynamic needs (e.g., antisocial attitudes, chemical dependency) can and should be targeted for the best recidivism outcomes. The RNR model also dictates that individual characteristics that could affect responsiveness to treatment should be considered when assigning offenders to programs (Andrews and Bonta, 2010; Dowden and Andrews, 1999). Gender is a responsivity issue. While some correctional programs are gender neutral, in that they can be effective for both males and females, some programs target the unique risk factors that affect females more than males, or vice versa.

Well-designed programs that adhere to the RNR model and include many of the other evidence-based intervention strategies outlined by Andrews et al. (1990) can be ineffective if they are not implemented as designed (Matthews, Hubbard, and Latessa, 2001; Van Voorhis and Brown, 1996). By altering an intervention's original design, program administrators risk losing too many of the program components that contribute to its potential effectiveness (Fixsen et al., 2005). Budgetary limitations, staff turnover, time constraints, and many other potential disruptions can erode program integrity (Durlak and DuPre, 2008). Evaluability assessments can be used to measure the degree to which programs maintain integrity upon implementation (Prosavac and Carey, 1992; Trevisan and Huan, 2003).

The Correctional Program Assessment Inventory (CPAI) and the Evidence-Based Correctional Program Checklist (CPC) are two standardized evaluability assessments created specifically to assess the design and implementation of correctional programs (Gendreau and Andrews, 1994; Latessa, 2012). Effective correctional programs can vary in terms of focus and substance, but there are several program elements that contribute to the likelihood that a program will significantly reduce recidivism, including qualified program leadership and staff, evidence-based treatment approaches, and use of risk and need assessments. The CPAI and CPC measure the extent to which these and other elements are present in a program. These tools were developed and validated based on assessments from hundreds of correctional programs. However, very few subsequent studies have examined the relationship between program integrity and recidivism outcomes.

Nesovick (2003) used a condensed version of the CPAI to rate adult and juvenile correctional programs based on 173 recidivism outcome evaluations with 266 effect sizes. Nesovick (2003) did not directly assess programs first-hand using the CPAI. Rather, the author based the assessment on written information about each of the evaluated programs. The average Pearson's r correlation between CPAI scores and phi coefficients derived from the evaluations was 0.46 ($p < 0.05$). The positive correlation coefficient indicates that higher CPAI scores are associated with larger recidivism reduction effects.

Using a more complete, yet still condensed, version of the CPAI, Lowenkamp et al. (2006) examined the relationship between program integrity and effectiveness using data from community-based residential programs ("halfway houses") in Ohio. The

researchers matched more than 3,000 parolees released to halfway houses with a similar set of parolees not released to halfway houses, and rated the halfway house programs using a slightly abbreviated form of the CPAI. Total CPAI scores were positively and significantly associated with new offense reincarcerations, supervision revocations, as well as both of these recidivism measures combined. This positive relationship means that higher program integrity was associated with larger reductions in recidivism for halfway house residents relative to the comparison group.

The present study compares recidivism outcomes from an evidence-based CBT program with a standardized curriculum from when the curriculum was and was not fully implemented. The CPAI and CPC were not used to assess this program at the time of full and partial implementation, but there was documentation about which design elements were lost when the program was altered. This evaluation measures the extent to which the loss of those evidence-based components affected the program's ability to reduce recidivism.

Moving On

Moving On is one of a growing number of standardized CBT programs used to treat correctional populations. Unlike most correctional-based CBT programs, Moving On was designed to treat female, and not male, offenders (Gehring et al. 2010). As female prison populations continue to grow (Carson, 2014), so too has the recognition that female offenders are both similar to and different from male offenders (Brennan et al. 2012; Holtfreter and Wattanaporn, 2013; Makarios, Steiner, and Travis, 2010; Van Voorhis, Wright, Salisbury, and Bauman, 2010; Wright, Van Voorhis, Salisbury, and Bauman, 2012). Male and female offenders share some of the same risk factors and

reentry hardships, including past criminal records, education deficits, and unstable employment histories (e.g., Greiner, Law, and Brown, 2014; Makarios et al., 2010; Smith, Cullen, and Latessa, 2009). However, males and females tend to be incarcerated for different types of offenses (Carson, 2014), and there is evidence that female offenders are more likely to have histories of multiple types of victimization and co-occurring mental health disorders and substance abuse issues (Belknap, 2007; Scroggins and Malley, 2010; Van Voorhis et al. 2010; Wright et a. 2012) .

Moving On is a gender-responsive CBT program that focuses on improving communication skills, building healthy relationships, and expressing emotions in a healthy and constructive manner (Gehring, Van Voorhis, and Bell, 2010; Van Dieten, 2010). The program is delivered in 26 sessions via group and one-on-one discussions, self-assessments, writing exercises, and role-playing and modeling activities. The women are encouraged to set goals for the future and assess their own personal strengths and weaknesses. Each session is designed to last one and a half to two hours (Gehring et al., 2010).

Moving On was initially offered to female offenders in the Minnesota Correctional Facility (MCF)-Shakopee during the fall of 2001 by trained facilitators. Up through 2010, it was generally offered to offenders on a quarterly basis. Participation in the program was voluntary, and offenders often entered the program during the last half of their confinement period. The program lasted a total of 12 weeks, participants were in class four hours per week for a total of 48 hours, and class sizes were relatively small (between 5 and 10 participants).

In 2011, however, a decision was made to begin offering Moving On to offenders shortly after their admission to the MCF-Shakopee. Due to concerns that scheduling offenders for Moving On often seemed to conflict with prison work assignments or participation in other institutional programs, Moving On began to be offered to offenders at the time of intake, or what is referred to as R&O (reception and orientation) at the MCF-Shakopee. Modifying the point at which offenders entered Moving On brought about several substantive changes to the way the programming was delivered. Because R&O generally lasts three weeks, the length of Moving On was trimmed from 12 weeks to 3 weeks. Offenders participated two hours each day, five days per week, for a total of 30 hours.

While some curriculum was cut in reducing overall classroom time from 48 hours to 30 hours, the main program changes involved the elimination of role-playing, skill-building, and homework exercises. The removal of these exercises was due not only to the condensed amount of time over which the programming was offered, but also to the fact that class sizes had greatly expanded to approximately 40-50 offenders per class. The loss of these components also led to the loss of timely reinforcements for each participant's contributions to the group (i.e., recognition, small material rewards), as well as consequences (i.e., redirection, failure to complete the program). The growth in class sizes was attributable, in no small part, to the fact that participation was no longer voluntary; rather, all offenders admitted to the MCF-Shakopee were required to participate in the "watered down" version of Moving On.

In the fall of 2013, a decision was made to return Moving On to the way it had operated prior to 2011. Currently, the full program (i.e., 48 hours of classroom time over

a 12-week period) is being offered on a quarterly basis, participation is voluntary, and class sizes are relatively small (less than 10). The one notable difference compared to how it operated prior to 2011 is that risk assessments are now being used to target which offenders should participate in Moving On. In April 2013, the Minnesota Department of Corrections (MnDOC) implemented the Minnesota Screening Tool Assessing Recidivism Risk (MnSTARR), a risk assessment instrument that has been validated on Minnesota prisoners (Duwe, 2014a). Consistent with the risk principle, offenders with a higher recidivism risk, per the MnSTARR, are being prioritized for participation in Moving On.

To date, there has been only one outcome evaluation of Moving On. Using a sample of female probationers in Iowa, Gehring et al. (2010) compared 190 Moving On participants to 190 similar female probationers that did not participate in any CBT during their probation periods. The treatment and comparison groups were matched on a limited number of characteristics, including judicial district, race, age, risk assessment scores, and probationary period start times. By comparing rates of four recidivism outcomes after 12 to 30 months of follow-up time, Gehring et al. (2010) found that Moving On participants had significantly lower rates of rearrest and new convictions than the comparison group of probationers. Moving On participants and comparison group members did not have significantly different rates of incarceration, but Moving On participants did have significantly higher rates of technical violations. When limiting the sample to Moving On program completers (N = 111) and the same number of matched probationers, Gehring et al. (2010) found that completers had significantly lower rates of rearrest, new convictions, and incarcerations than the comparison group. The difference in rates of technical violations was not significant between the groups.

The results of Gehring et al.'s (2010) analysis are encouraging for Moving On's effectiveness, but this study suffers from two key methodological shortcomings. First, the authors should have used more probationer characteristics to match treatment and control group members, and they could have conducted a more rigorous matching process to ensure balanced treatment and control groups. Second, the authors did not conduct any multivariate analyses to control for the effect of other potential variables on the recidivism outcomes.

Given that Gehring et al.'s (2010) study is the only evaluation of Moving On's effect on recidivism outcomes, there is limited evidence that Moving On works for women correctional populations. However, Moving On's original design and implementation at MCF-Shakopee included multiple elements that contribute to program effectiveness. Table 1 lists 10 evidence-based program characteristics and implementation strategies and whether or not these elements were present during the two phases of implementation that are compared in the ensuing analyses. The early phase of implementation (covering years 2001 to 2010) is referred to as the "High Fidelity" phase, and the latter phase (covering years 2011 to 2013) is referred to as the "Low Fidelity" phase. The elements listed in this table are loosely based on the CPC, as well as general knowledge based on the "what works" literature (Gendreau and Andrews, 1994; Gendreau et al. 1999; Latessa, 2012). Although only 10 items are listed in Table 1, they relate to at least 25 scoring items on the CPC and are the ones most applicable to group treatment programs.

Table 1. Description of Program Components and Strategies Present and Absent During Implementation Phases at MCF-Shakopee

Moving On Components and Implementation Strategies	Implementation Phase	
	High Fidelity (2001 to 2010)	Low Fidelity (2011 to 2013)
<u>1. Facilitator Qualifications:</u> Facilitators have at least a bachelor’s degree or higher in a helping profession, at least two years of prior experience working with offender populations, and completed the 5-day Moving On facilitator training	Yes	Yes
<u>2. Risk, Need, and Responsivity Assessed:</u> Offender risk, need, and responsivity assessed using validated and normed actuarial assessments; participants matched to programming based on assessment results	No	No
<u>3. Target Higher Risk Offenders:</u> At least two-thirds of the participants are rated as medium- or high-risk to reoffend	No	No
<u>4. Criminogenic Targets:</u> Program targets several criminogenic needs with treatment (e.g., antisocial cognition, unhealthy peer and family relationships, harmful emotional expressions)	Yes	Yes
<u>5. Treatment Approach:</u> Program uses an evidence-based treatment approach, including radical behavioral therapy or cognitive-behavioral therapy	Yes	No
<u>6. Treatment Length:</u> The average program length is between 3 and 9 months, and not longer than 1 year	Yes	No
<u>7. Group size:</u> Group sizes range from 5 to 10 participants; no more than 10 participants per facilitator	Yes	No
<u>8. Use of Reinforcers and Punishers:</u> A range of social or tangible rewards are used to acknowledge progress and accomplishments in the program; sanctions are imposed for antisocial or disruptive behaviors observed during the program	Yes	No
<u>9. Completion Criteria and Completion Rate:</u> Completion of the program requires progress and acquisition of new skills; completion not based solely on attendance; a majority of the participants are able to complete the program, but not every participant necessarily completes the program	Yes	No
<u>10. Skill Modeling and Training with Graduated Practice:</u> Facilitators model skills to use in response to adverse stimuli; participants practice skills using simulations and role-playing exercises; new skills are practiced in gradually more difficult situations	Yes	No
<u>Total</u> Percent of evidence-based components and implementation strategies present during phase of implementation	80%	20%

Risk, need, and responsivity assessment instruments were not widely or consistently used within MnDOC during most of the time period covered in this study, so participants were not matched to Moving On based on the results of such assessments (reference item 2 in Table 1). Because risk scores are not available for many of the women included in this study, it is not known whether or not most of the program participants were medium- or high-risk to reoffend (reference item 3 in Table 1).

With the exception of qualified facilitators and the targeting of multiple criminogenic needs, Moving On's second phase of implementation at MCF-Shakopee (the Low Fidelity phase) lost many of the evidence-based elements present during the first phase of implementation (the High Fidelity phase), including ideal program length and group size and the use of skill modeling and training with increasing difficulty (reference items 5 through 10 in Table 1). Overall, the High Fidelity phase of implementation included 80 percent of these items, while the Low Fidelity phase included only 20 percent of these items. In addition to providing an evaluation of Moving On's effectiveness at reducing recidivism by overcoming the methodological shortcomings of the previous study, the present study assesses what effect, if any, the loss of program integrity has on recidivism outcomes.

Data and Methodology

The population for this study consisted of 4,101 female offenders released from prison in Minnesota between 2003 and 2013. Of these offenders, 216 participated in Moving On prior to 2011 when it was run with integrity. Another 864 offenders participated in the program during the 2011-2013 period when it did not operate with fidelity. The remaining 3,021 inmates did not participate in either version of Moving On.

To determine whether participation in Moving On and, more generally, program integrity had an impact on recidivism outcomes, we used a retrospective quasi-experimental design with three separate sets of comparisons. Our first comparison assessed the effects of participating in Moving On prior to 2011 on recidivism. Therefore, our treatment group for this comparison included the 216 offenders released during the 2003-2013 period who participated in Moving On before 2011. The pool for our comparison group, meanwhile, contained 2,972 female offenders released between 2003 and 2013 who did not participate in Moving On.

Our second comparison examined the impact of the Moving On program offered during the 2011-2013 period on recidivism. The treatment group consisted of the 864 offenders who participated in this version of Moving On and were released prior to 2014. Nearly all of the female offenders who were admitted to prison between 2011 and 2013 participated in Moving On. In fact, given that there were only 49 who did not participate, mainly due to very brief lengths of stay in prison, it was not possible to construct a contemporaneous comparison group of non-participants. As a result, we relied on a historical comparison group pool that contained the same 2,972 non-participants used for the first comparison.

For the third comparison, we assessed the effects of participating in Moving On both before 2011 and during the 2011-2013 period on recidivism. More specifically, we compared the 216 pre-2011 Moving On participants with the 864 offenders who participated in the program between 2011 and 2013. In our analyses, the pre-2011 participants comprised the treatment group, while the 2011-2013 participants made up the comparison group.

In an effort to control for observable selection bias, we used propensity score matching (PSM), which we discuss below in more detail, so as to create equivalent comparison groups for all three comparisons. The use of multiple comparisons enables us to draw inferences about the effects of both Moving On and program integrity on recidivism. For example, if Moving On works but program integrity is irrelevant, we should expect to see better outcomes from participants in the first two comparisons but no difference between groups for the third comparison. If integrity matters, however, we should expect to see better recidivism outcomes from the pre-2011 participants in the first and third comparisons. But if Moving On is ineffective and program integrity does not matter, we should not expect to see improved recidivism outcomes in any of the three comparisons.

Dependent Variable

Because there is no single best measure of recidivism, we used multiple measures in this study. We operationalized recidivism as a: 1) rearrest, 2) reconviction, 3) reincarceration for a new offense, or 4) revocation for a technical violation. Among the first three measures, which strictly quantify new criminal offenses, rearrest provides the most sensitive measure of reoffending since not all rearrests result in a reconviction. New offense reincarceration, on the other hand, offers the most conservative reoffending measure given that offenders who are rearrested and reconvicted for a new offense may receive a probation sentence, for example, rather than a prison sentence. Compared to the three reoffense measures, technical violation revocations (the fourth measure) represent a broader measure of rule-breaking behavior. Offenders can have their post-release supervision (i.e., parole) revoked for violating the conditions of their supervised release.

Because these violations can include activity that may not be criminal in nature (e.g., use of alcohol, failing a community-based treatment program, failure to maintain agent contact, failure to follow curfew, etc.), technical violation revocations do not necessarily measure reoffending.

Recidivism data were collected on offenders through June 30, 2014. Because the offenders in this study were released between January 2003 and December 2013, the follow-up time ranged from six months to more than 11 years. Data on arrests and convictions were obtained from the Minnesota Bureau of Criminal Apprehension, while reincarceration and revocation data were derived from the Correctional Operations Management System (COMS)—the MnDOC’s database. Because these data measure only arrests, convictions or incarcerations that took place in Minnesota, the findings presented later likely underestimate the true recidivism rates for the offenders included in this study. We anticipate, however, that the amount of non-Minnesota recidivism will be similar across all treatment and comparison groups.

To accurately measure the total amount of time offenders were actually at risk to reoffend (i.e., “street time”), we accounted for supervised release revocations in the recidivism analyses. For the three recidivism variables that strictly measure new criminal offenses (rearrest, reconviction, and new offense reincarceration), it was necessary to deduct the amount of time they spent in prison for technical violation revocations from their total follow-up period. Failure to deduct time spent in prison as a supervised release violator would artificially increase the length of the at-risk periods for these offenders. Therefore, to achieve a more accurate measure of “street time”, the time an offender spent in prison as a supervised release violator was subtracted from her follow-up period,

but only if it preceded a reoffense or if the offender did not recidivate prior to July 1, 2014. Similarly, to accurately measure “street time” for the technical violation revocation measure, we accounted for the time an offender spent in prison for a new felony offense, which was deducted from the follow-up period as long as it preceded a revocation or if the offender had not been revoked by the end of June 2014.

Independent Variables

Participation in Moving On is the key variable of interest in this evaluation. Offenders who participated in Moving On were assigned a value of “1”, whereas the offenders in the comparison group were given a value of “0”. In the comparison between pre-2011 and 2011-2013 Moving On participants, the former were given a value of “1” while the latter received a value of “0”. The independent, or control, variables included in the statistical models were those that were not only available in COMS but also might have an impact on recidivism and Moving On program selection (see Table 2).

We included a number of measures commonly associated with recidivism risk, such as the offender’s race, age, number of prior supervision failures, number of prior convictions, number of felony convictions and institutional misconduct. Previous research on Minnesota prisoners has shown that suicidal history increases an offender’s risk for recidivism (Duwe, 2014a). We also accounted for admission type (new commit), offense type, commitment county (metro), and length of stay because prior studies have indicated these variables are significant predictors of recidivism for Minnesota prisoners (Duwe, 2010; Duwe and Clark, 2013).

In addition to including factors that increase the likelihood of recidivism, we accounted for factors that have been shown to lower recidivism risk, such as prison visits

(Duwe and Clark, 2013), participation in the Challenge Incarceration Program (CIP)—a correctional boot camp program (Duwe and Kerschner, 2008), and involvement in programming relating to chemical dependency treatment (Duwe, 2010), education (Duwe and Clark, 2014), employment (Duwe, 2012), and work release (Duwe, 2014b).

Combined, the covariates we used tap into a number of risk factors such as antisocial history (prior supervision failures, prior criminal history, prison misconduct), social support (prison visits), antisocial cognition (chemical dependency treatment and CIP are delivered within a cognitive-behavioral framework), education/employment (educational programming, employment programming, and participation in work release), and substance abuse (chemical dependency treatment).

Propensity Score Matching

PSM is a method that estimates the conditional probability of selection to a particular treatment or group given a vector of observed covariates (Rosenbaum & Rubin, 1985). The predicted probability of selection, or propensity score, is typically generated by estimating a logistic regression model in which selection (0 = no selection; 1 = selection) is the dependent variable while the predictor variables consist of those that theoretically have an impact on the selection process. Once estimated, the propensity scores are then used to match individuals who participated in an intervention with those who did not. In matching offenders who entered Moving On with those who did not on the conditional probability of selection into the program, the main advantage with using PSM is that it can simultaneously “balance” multiple covariates on the basis of a single composite score. In doing so, PSM helps create a counterfactual estimate of what would

have likely happened to the offenders in the Moving On group had they not participated in the program.

Despite its growing popularity as a matching technique, PSM has several limitations that are worth noting. First, and most important, because propensity scores are based on observed covariates, PSM cannot control for “hidden bias” from unmeasured variables that are associated with both the assignment to treatment and the outcome variable. Second, in order for PSM to be effective, there must be substantial overlap among propensity scores between the treatment and comparison groups (Shadish, Cook & Campbell, 2002). If the overlap is insufficient, the matching process will yield incomplete or inexact matches. Finally, PSM is generally more effective with larger samples (Rubin, 1997).

In addition to using a large sample (N = 4,101), we tried to address the “hidden bias” limitation, to the extent possible, by including a relatively lengthy list of theoretically-relevant covariates in our statistical models. Moreover, the matching for the first two comparisons was largely successful, which reflects the fact that the overlap in propensity scores was sufficient. Achieving complete and exact matches for the third comparison was more difficult, however, due to the greater separation in propensity scores between the two groups of Moving On participants. As discussed below in more detail, we used multiple matching methods along with covariate and propensity score adjusted Cox regression models.

Matching for Moving On Selection

For each of the three sets of comparisons, we calculated propensity scores by estimating a logistic regression model in which the dependent variable was participation

in Moving On. The variables included in a propensity score estimation model should consist of those related to the outcome—even if it is a weak association—that affect treatment selection and are not caused by the treatment (Shadish et al., 2002). As we described above, the point at which offenders entered Moving On during their confinement varied between the pre-2011 and 2011-2013 periods. More specifically, because Moving On participants from 2011-2013 entered the program toward the beginning of their incarceration, most of the covariates pertaining to participation in programming (e.g., chemical dependency treatment, EMPLOY, etc.) and post-release supervision (e.g., intensive supervised release, discharge, etc.) do not temporally precede their involvement in Moving On. Although Moving On participation is not one of the criteria MnDOC staff consider in making programming and supervision level decisions for female offenders (e.g., whether an offender is placed on intensive supervised release at the time of release is not caused by participation in Moving On), it is possible that Moving On may have affected measures such as institutional misconduct (i.e., discipline convictions).

We therefore estimated a propensity score estimation model that contained only the covariates that would be known at the time of intake and, thus, would precede potential selection into Moving On across both time periods. Yet, to address the possibility that these covariates may not include all of the variables that affected selection, particularly for the pre-2011 period, we also estimated a propensity score estimation model that included all of the covariates we examined. As we note later on, both approaches yielded similar results regarding Moving On's impact on recidivism.

Consequently, we focus on the results pertaining to the propensity score models that included only the covariates known at the time of intake.

Table 2 describes the covariates used in the propensity score estimation models, and it presents the results from these analyses. The results show a number of factors that predicted selection for each of the three comparisons we examined. For the first comparison, the results reveal that the odds of participating in pre-2011 Moving On were significantly greater for offenders incarcerated for a violent offense and inmates with more felony convictions. The odds were significantly less, however, for offenders with supervision failures and those admitted to prison as a release violator. For the second comparison, the likelihood of participating in Moving On from 2011-2013 was significantly greater for offenders with more total convictions, probation violators, and offenders who entered prison with a secondary degree (i.e., high school degree or GED). The odds of participation were significantly lower, however, for offenders who had shorter sentences, more supervision failures, a greater number of felonies, and were admitted to prison as release violators. For the third comparison, the chances of participating in pre-2011 Moving On were significantly greater for offenders with longer sentences and a larger number of felony convictions. Conversely, the odds were significantly less for probation violators.

Table 2. Logistic Regression Models for Moving On Program Selection

<i>Predictors</i>	<i>Predictor Description</i>	<i>1</i>		<i>2</i>		<i>3</i>	
		<u>OR</u>	<u>SE</u>	<u>OR</u>	<u>SE</u>	<u>OR</u>	<u>SE</u>
Age at Intake (years)	Offender age in years at time of admission to prison	0.987	0.008	0.997	0.005	0.986	0.010
Sentence Length (months)	Sentence Length in Months	1.000	0.001	0.982**	0.003	1.041**	0.005
Minority	Minority = 1; White = 0	1.130	0.159	0.974	0.091	1.155	0.203
Metro Commit	Twin Cities metropolitan area = 1; Greater Minnesota = 0	1.092	0.151	0.888	0.088	1.150	0.193
Supervision Failures	Number of prior revocations while under correctional supervision	0.688*	0.153	0.451**	0.145	1.328	0.228
Total Convictions	Total number of convictions, including index conviction(s)	0.994	0.012	1.052**	0.006	0.974	0.015
Felony Convictions	Total number of felonies, including index conviction(s)	1.074*	0.029	0.819**	0.028	1.208**	0.048
Offense Type	Other offense serves as the reference						
Violent	Violent offense = 1; non-violent offense = 0	3.092**	0.306	1.084	0.188	1.830	0.371
Drugs	Drug offense = 1; non-drug offense = 0	1.184	0.300	1.125	0.163	0.713	0.355
Property	Property offense = 1; non-property offense = 0	1.084	0.317	1.054	0.168	0.889	0.371
Felony DWI	Felony DWI offense = 1; non-Felony DWI offense = 0	2.111	0.401	1.149	0.249	1.129	0.479
Admission Type	New Court Commitment serves as the reference						
Probation Violator	Probation Violator = 1; new commit and release violators = 0	0.681	0.217	3.915**	0.176	0.313**	0.307
Release Violator	Release Violator = 1; new commit and probation violators = 0	0.174**	0.521	0.083**	0.749	1.947	1.047
Secondary Degree	Secondary degree at intake = 1; less than secondary degree = 0	1.398*	0.160	1.374**	0.088	1.012	0.205
Post-secondary Degree	Post-secondary at intake = 1; less than post-secondary degree = 0	2.222	0.654	0.801	0.711	1.549	1.207
Constant		0.081**	0.395	0.378**	0.239	0.130**	0.518
N		3,188		3,866		1,080	
Log-likelihood		1457.105		3527.189		787.710	
Area Under Curve (AUC)		0.708		0.750		0.835	
Nagelkerke R ²		0.097		0.209		0.376	

** $p < .01$

* $p < .05$

Comparison #1: Pre-2011 Moving On vs. contemporaneous comparison group

Comparison #2: 2011-2013 Moving On vs. historical comparison group

Comparison #3: Pre-2011 Moving On vs. 2011-2013 Moving On

After obtaining propensity scores for the three sets of comparisons, a “greedy” matching procedure that utilized a without replacement method was used to match the offenders from the treatment and comparison groups. For the first two comparisons, Moving On participants were individually matched to a comparison group of non-participants who had the closest propensity score (i.e., “nearest neighbor”) within a relatively narrow caliper (i.e., range of propensity scores) of 0.05. We obtained a match rate of 99.5 percent for the treatment group offenders in these two comparisons. For example, of the 216 pre-2011 Moving On participants, we found a comparison group match for all but one of the offenders. For the second comparison, we found matches for 860 of the 864 Moving On participants from 2011-2013.

With the third comparison, however, it was more difficult to produce a high rate of exact matches due to the lack of strong overlap in propensity scores between pre-2011 and 2011-2013 Moving On participants. Indeed, we were able to obtain matches for only 80 percent of the pre-2011 participants using a .05 caliper. To avoid bias due to incomplete matching, we used nearest neighbor matching in which we matched all 216 of the pre-2011 participants with 216 participants from the 2011-2013 period.

In Table 3, we present statistics that measure the degree to which PSM was effective in reducing observable selection bias for the three comparisons. We use a measure (“Bias”) developed by Rosenbaum and Rubin (1985) that quantifies the amount of bias between the treatment and comparison samples (i.e., standardized mean difference

$$\text{Bias} = \frac{100(\bar{X}_t - \bar{X}_c)}{\sqrt{\frac{(S_t^2 + S_c^2)}{2}}}$$

Table 3. Covariate Balance for Moving On Selection

Variable	1			2			3		
	MO #1	Comparison	Bias	MO #2	Comparison	Bias	MO #1	MO #2	Bias
Propensity Score	0.11	0.11	0.09	0.33	0.32	1.40	0.44	0.32	35.92
Age at Intake (years)	32.85	33.15	2.70	32.68	32.72	0.35	32.81	32.47	3.03
Sentence Length (months)	47.50	50.17	2.24	20.47	20.41	0.28	47.40	35.69	30.23
Minority	0.42	0.45	4.94	0.36	0.37	1.70	0.42	0.41	1.65
Metro Commit	0.48	0.46	3.26	0.39	0.46	11.75	0.48	0.47	1.63
Supervision Failures	0.53	0.54	1.14	0.83	0.85	3.23	0.53	0.50	3.45
Total Convictions	8.52	8.92	3.68	8.97	9.08	1.15	8.64	8.83	1.85
Felony Convictions	3.37	3.49	2.79	2.27	2.30	1.37	3.52	2.89	13.67
Violent Offense	0.31	0.30	1.76	0.13	0.12	2.47	0.31	0.24	12.61
Drug Offense	0.32	0.31	1.75	0.43	0.44	1.65	0.32	0.36	6.92
Property Offense	0.23	0.24	1.92	0.32	0.32	0.00	0.24	0.22	3.88
DWI Offense	0.07	0.10	9.24	0.05	0.04	3.88	0.06	0.10	12.34
Probation Violator	0.34	0.34	0.00	0.78	0.78	0.00	0.34	0.38	6.82
Release Violator	0.02	0.03	5.23	0.00	0.00	0.00	0.02	0.01	6.03
Secondary Degree at Intake	0.68	0.71	5.28	0.63	0.66	5.10	0.68	0.66	3.47
Post-Sec. Degree at Intake	0.01	0.00	7.78	0.00	0.00	0.00	0.01	0.00	7.78

#1 N = 430 (Moving On #1 = 215; comparison group = 215)

#2 N = 1,760 (Moving On #2 = 860; comparison group = 860)

#3 N = 432 (Moving On #1 = 216; Moving On #2 = 216)

between samples), where \bar{X}_t and S_t^2 represent the sample mean and variance for the treated offenders and \bar{X}_c and S_c^2 represent the sample mean and variance for the untreated offenders. If the bias value exceeds 20, the covariate is considered to be unbalanced (Rosenbaum & Rubin, 1985).

Prior to matching, there were five imbalanced covariates for the first comparison, four for the second comparison, and five for the third comparison. After matching, the results presented in Table 3 show that all 15 covariates (plus the propensity score) had bias values below 20 for the first two comparisons. But for the third comparison, we see that sentence length (plus the propensity score) had a bias value greater than 20.

Analysis

Given that recidivism is typically operationalized as a binary outcome, multiple logistic regression is a popular technique for recidivism analyses. One key assumption that logistic regression makes in analyzing recidivism is that offenders have follow-up periods that are equal in length. When they vary in length, however, the shortest observed follow-up period must be used in order to meet this assumption. For example, because the follow-up periods in this study ranged from six months to 11 years, we would need to limit the follow-up period to six months for all offenders in order to use logistic regression for our recidivism analyses. In addition to resulting in a significant loss of outcome data, the use of such a brief follow-up period for recidivism would weaken our ability to draw valid conclusions about the effectiveness of Moving On or the importance of program integrity.

Because survival analysis models are designed to handle censored observations, they can accommodate follow-up periods that vary in length. Therefore, we used Cox

regression, a multivariate survival analysis technique, for our recidivism analyses. Cox regression relies on time-dependent data, which are important in determining not only whether offenders recidivate but also when they recidivate. More specifically, it uses both “time” and “status” variables in estimating the impact of the independent variables on recidivism. For the analyses presented here, the “time” variable measures the amount of time (in days) from the date of release until the date of first rearrest, reconviction, new offense reincarceration, technical violation revocation, or June 30, 2014, for those who did not recidivate. The “status” variable, meanwhile, measures whether an offender recidivated (rearrest, reconviction, new offense reincarceration, and technical violation revocation) during the period in which she was at risk to recidivate.

Results

In Table 4, we present the recidivism rates for the offenders in the three comparisons we analyzed. In the first comparison, which contains 215 pre-2011 Moving On participants and a contemporaneous comparison group of 215 non-participants, we see that offenders who participated in the program had lower rates for all four recidivism measures, especially rearrest and reconviction. For example, through the end of June 2014, 49 percent of Moving On participants had been rearrested versus 63 percent of those in the comparison group. Likewise, 48 percent of comparison group offenders had been reconvicted compared to 35 percent of Moving On participants.

Table 4. Recidivism Rates for Moving On Participants and Comparison Group Offenders

	<i>Rearrest</i>	<i>Reconviction</i>	<i>New Offense Reincarceration</i>	<i>Technical Violation Revocation</i>
<u>Comparison #1 (N = 430)</u>				
Pre-2011 Moving On	<u>Percent</u> 49.3	<u>Percent</u> 34.9	<u>Percent</u> 16.3	<u>Percent</u> 23.7
Comparison Group	62.8	48.4	20.5	27.4
<u>Comparison #2 (N = 1,720)</u>				
Moving On, 2011-2013	42.9	25.1	6.9	20.9
Comparison Group	71.6	55.6	21.7	25.3
Time-adjusted rate	39.2	24.1	4.8	19.9
<u>Comparison #3 (N = 432)</u>				
Pre-2011 Moving On	49.1	34.7	16.2	23.6
Time-adjusted rate	19.4	7.9	2.3	16.2
Moving On, 2011-2013	31.0	16.7	4.2	19.9

Because we used historical comparison groups for the second and third comparisons, simply comparing recidivism rates through June 2014 can be misleading due to the varying lengths of the follow-up periods (i.e., the longer follow-up period, the higher the recidivism rate) between groups. For example, in the third comparison, the average follow-up period length for pre-2011 Moving On participants was 2,445 days (80 months) versus an average of 528 days (17 months) for the 2011-2013 participants. As a result, we also calculated time-adjusted rates for the two groups that had longer follow-up periods (non-participants in comparison #2 and pre-2011 Moving On participants in comparison #3). Given that the matching process was performed on an individual basis, we shortened the follow-up periods for the offenders in these two groups so that it was commensurate with the length of the follow-up period for their 2011-2013 Moving On counterparts.

To illustrate, for the second comparison, let us assume that a non-participant in the comparison group had a follow-up period of 1,825 days (about five years) while her

matched counterpart in the 2011-2013 Moving On group had a follow-up period of 730 days (about two years). For the non-participant in the comparison group, we calculated her recidivism rates based on a 730-day follow-up period. We performed this calculation for all 860 non-participants in the second comparison and all 216 pre-2011 Moving On participants in the third comparison.

For both the second and third comparisons, we see that the groups with the longer follow-up periods (non-participants in the second comparison and pre-2011 Moving On participants in the third comparison) had much higher recidivism rates. When we examine the time-adjusted rates, however, we see little difference in recidivism for the second comparison between the 2011-2013 Moving On participants and the matched comparison group of non-participants. Moving On participants had rates that were slightly higher than the time-adjusted rates for their comparison group counterparts for all four recidivism measures. For the third comparison, we see that the time-adjusted rates for the pre-2011 Moving On participants are lower than their 2011-2013 Moving On counterparts for all four recidivism measures.

The Effects of Moving On and Program Integrity on the Hazard of Recidivism

To determine the effects of Moving On and program integrity on recidivism, we estimated Cox regression models for each recidivism measure across all three comparisons, resulting in 12 models total.¹ Each model contains covariates known to be associated with recidivism that were excluded from the propensity score estimation models because they follow entry into Moving On, at least for the 2011-2013 participants. As indicated earlier, although we obtained complete matches for our third

¹ For all of the models we estimated, we assessed the proportional hazards assumption by including a time-dependent covariate for Moving On participation.

comparison (pre-2011 participants versus 2011-2013 participants), the matches were inexact due to a lack of covariate balance. Therefore, in the third comparison, we estimated models that included the propensity score, which can be conceptualized as a single covariate that approximates adjusting for all of the covariates in the propensity score estimation model since it encapsulates the distribution of these covariates (Austin, 2014). For the second and third comparisons, we estimated additional Cox regression models that utilized the time-adjusted follow-up periods discussed above. We do not present the results from these additional models, however, because they were largely the same as those produced from the models that used the full follow-up period.

As shown in Table 5, the results indicate that, controlling for the effects of the other covariates, participating in Moving On prior to 2011 significantly reduced two of the four recidivism measures in the first comparison models, lowering the risk of reoffending by 31 percent for rearrest and 33 percent for reconviction. The hazard ratios were in the negative direction for new offense reincarceration and technical violation revocations, but neither one was statistically significant at the .05 level in any of the three models.²

² As noted above, we carried out an additional PSM analyses in which we used all of the covariates (except for age at intake, sentence length, and educational degrees at intake) in the propensity score estimation model. For the first comparison, matches were obtained for 215 Moving On participants, and none of the covariates had bias values above 20. The results from a bivariate Cox regression model showed that Moving On significantly reduced the risk of rearrest (31 percent reduction) and reconviction (38 percent reduction). Significant effects were not found for either new offense reincarceration or technical violation revocations.

Table 5. Cox Regression Models: Impact of Moving On Program Participation on the Hazard of First Recidivism Event

Predictors	Hazard Ratios by Comparisons and Type of Recidivism											
	Rearrest			Reconviction			New Offense Reincarceration			Technical Violation Revocation		
	<u>1</u>	<u>2</u>	<u>3</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>1</u>	<u>2</u>	<u>3</u>
Moving On (pre-2011)	0.695**		0.562**	0.671*		0.527**	0.828		0.666	0.945		1.299
Moving On (2011-2013)		1.125			1.026			1.232			0.874	
Propensity Score			1.126			1.283			2.291			0.869
Age at Release (years)	0.970**	0.980**	0.987	0.971**	0.980**	0.993	0.971*	0.961**	0.991	0.980	0.980**	1.001
Length of Stay (Months)	0.992	0.966**	0.987	0.992	0.970**	0.989	0.995	0.984	0.990	0.992	0.987	0.986
Prison Discipline Convictions	1.007	1.018**	1.014**	1.012*	1.012**	1.017**	1.009	1.014*	1.004	1.009*	1.021**	1.013*
Suicidal Tendencies	1.480**	1.119	1.085	1.219	1.029	1.109	1.467	1.461**	1.304	2.440**	1.762**	1.390
Visited in Prison	0.676*	0.882	0.644*	0.535**	0.980	0.498**	0.587*	0.792	0.570	0.905	0.566**	0.421**
Earned Secondary Degree in Prison	1.116	0.998	1.016	0.931	0.939	0.668	0.975	0.895	0.651	0.683	0.635	0.882
Earned Post-Sec. Degree in Prison	0.855	0.659	0.899	0.765	0.667	0.486	1.206	0.697	0.580	0.566	0.865	0.937
Entered Chemical Dependency TX	0.839	1.187	0.724	0.858	1.197	0.591	0.616	1.045	0.630	1.158	0.573**	1.085
Entered EMPLOY Program	0.931	0.849	1.041	0.835	0.783	1.213	0.608	0.625	0.586	0.418	0.745	0.487
Entered Work Release Program	0.964	0.628**	0.631*	1.003	0.501**	0.523*	0.831	0.492**	0.619	1.718*	1.250	1.146
Entered CIP	0.981	0.513**	0.625	0.834	0.322**	0.432	0.616	0.535	0.236	1.775	2.197**	1.473
Placed on ISR	1.081	0.638	0.798	0.634	0.768	0.464	0.740	0.808	0.508	2.858**	1.076	1.997*
Discharge	2.320*	0.931	1.839	2.823*	1.205	3.261	3.173*	2.225	4.294			
Supervised Release Revocations	1.149	1.002	1.664**	1.335*	1.027	1.561*	1.253	1.063	1.346			
New Offense Reincarcerations										0.683	0.852	0.623
N	430	1,720	432	430	1,720	432	430	1,720	432	422	1,709	429

** $p < .01$

* $p < .05$

Comparison #1: Pre-2011 Moving On vs. contemporaneous comparison group

Comparison #2: 2011-2013 Moving On vs. historical comparison group

Comparison #3: Pre-2011 Moving On vs. 2011-2013 Moving On

For the second comparison (2011-2013 Moving On participants vs. a historical comparison group of non-participants), the results from all three models indicated that Moving On participation did not have a significant effect on any of the four recidivism measures. Although the hazard ratio was in the negative direction for technical violation revocations, it was in the positive direction for the other three measures.³

For the third comparison, which compares pre-2011 and 2011-2013 Moving On participants, the results are largely similar to those observed for the first comparison. Compared to 2011-2013 Moving On participants, the risk of rearrest and reconviction was significantly lower for pre-2011 Moving On participants in all three models. More precisely, the hazard of reoffense was 44 percent lower for rearrest and 47 percent lower for reconviction.⁴ As with the first comparison, significant effects were not observed for either new offense reincarceration or technical violation revocations.⁵ While the hazard ratio was in the negative direction for new offense reincarceration, it was in the positive direction for technical violation revocations.⁶

³ In the additional PSM analyses for the second comparison, we obtained matches for 861 Moving On participants and all of the covariates were balanced. The results were very similar, as the 2011-2013 version of Moving On did not have a significant effect on any of the recidivism measures. The only difference is that the direction of the hazard ratio was negative for reconviction and positive for technical violation revocations.

⁴ In the additional PSM analyses for the third comparison, we also used nearest neighbor matching due to incomplete matches resulting from matching with a .05 caliper. Because four of the covariates (probation violator, discipline, length of stay, and visited) had bias values greater than 20, we estimated models with and without the propensity score. Neither reincarceration measure was statistically significant in either model. Participation in pre-2011 Moving On significantly reduced the risk of rearrest, lowering it from 38-42 percent in the two models. Similarly, pre-2011 Moving On participation significantly decreased the hazard of reconviction, reducing it from 49-58 percent in the two models.

⁵ We also estimated models for this comparison in which we excluded the propensity score. The results were virtually the same, with significant effects for rearrest (44 percent reduction in the hazard) and reconviction (47 percent decrease in the hazard) and non-significant findings for both reincarceration measures.

⁶ To avoid biased estimates, unreliable confidence interval coverage, and convergence problems in logistic regression models, Penduzzi, Concato, Kemper, Holford, and Feinstein (1996) recommend a rule of thumb of ten events per variable (EPV) based on the simulation results from their study. In a more recent simulation study by Vittinghoff and McCulloch (2007), they report the EPV standard could likely be cut in half to five predictors per event. Given the modest sample size for our third comparison (N = 432),

The results also showed the hazard ratio was significantly greater for offenders with more institutional discipline convictions (9 of the 12 models), younger offenders (7), offenders with suicidal tendencies (4), and inmates with shorter lengths of stay in prison (2). Offenders placed on intensive supervised release (ISR) had a significantly greater hazard of revocation for the first and third comparisons, while those with supervised release revocations had a greater risk of subsequent reoffending in three of the models. Similarly, offenders who were discharged (i.e., released to no correctional supervision because they completed their sentence) had an increased risk of reoffending in three models. Participation in work release and CIP increased the hazard of revocation in one of the models, but also decreased the risk of recidivism in several of the models. Finally, offenders who received prison visits had a reduced hazard of recidivism in seven of the models, and the risk of revocation was lower for CD treatment participants in one of the models.

Discussion

The results suggest that Moving On was generally effective in reducing recidivism prior to 2011. Although significant effects were not observed for either reincarceration measure, pre-2011 participation in Moving On lowered the risk of rearrest and reconviction. The findings further showed that between 2011 and 2013, Moving On did not have a significant effect on any of the four measures of recidivism. The results from the first two comparisons were confirmed by the third comparison, which indicated that recidivism outcomes—particularly for rearrest and reconviction—were significantly

combined with the relatively low overall rate for new offense reincarceration for this comparison (10%), the EPV was below five for this model. We estimated models in which the EPV was higher than either threshold (5 or 10), but the results were not substantively different than those reported in Table 5.

better for pre-2011 participants in comparison to those who participated in Moving On during the 2011-2013 period.

Overall, the findings suggest that Moving On can be an effective correctional program for female offenders. But the results also imply that its effectiveness hinges on whether it is implemented with fidelity, which provides support for the notion that program integrity matters when it comes to reducing recidivism. Indeed, when the operation of Moving On was largely consistent with how it was designed, the program significantly lowered the risk of rearrest and reconviction. But when parts of the curriculum were cut, the length of the program was shortened, and class sizes were far bigger than recommended during the 2011-2013 period, participation in Moving On ceased to reduce reoffending. As the quality of the intervention was diluted, so was its beneficial impact on recidivism.

Conclusion

There are several limitations with this study that are worth noting. First, we focused on the effectiveness of a specific correctional program among a sample of female offenders who were incarcerated in Minnesota's prison system. As a result, the findings may not be generalizable to other correctional programs, other offender populations (e.g., probationers or male offenders), or offenders from other jurisdictions.

Second, historical comparisons are generally weaker than contemporaneous ones, and we relied—out of necessity—on historical comparison groups for two of the three comparisons we analyzed. Despite our use of multiple comparison groups, it is possible that the results observed in this study may be influenced by factors unique to the offenders in the 2011-2013 Moving On group that we could not control.

Third, although we were able to document the differences in program integrity between the two time periods we examined, we were unable to determine whether some or all of these differences were responsible for the recidivism outcomes we observed. It is possible, for example, that it was the large class sizes for Moving On during the 2011-2013 period rather than the abbreviated curriculum that weakened its impact on recidivism. The virtual absence of role-playing exercises, on the other hand, may have been the culprit for the worse recidivism outcomes. Or the timing as to when the programming was provided to offenders could have made a difference, since it was offered much earlier during an offender's incarceration period (at intake) for the 2011-2013 participants.

Fourth, and perhaps most important, we were unable to control for either offender motivation or whether they volunteered for Moving On. Recall, for example, that participation in the program was voluntary prior to 2011, whereas it was mandatory between 2011 and 2013. To be sure, it is possible that the reason for the better recidivism outcomes before 2011 is due to volunteerism rather than program integrity. Existing research provides mixed evidence, however, on the impact that volunteerism has on treatment effectiveness. Findings from the substance abuse and sex offender treatment literatures suggest that mandatory interventions can be just as effective as voluntary programming (Anglin, Brecht, and Maddahian, 1989; Grady, Edwards, Pettus-Davis, and Abramson, 2012; Knight, Hiller, Broome, and Simpson, 2000; McSweeney, Stevens, Hunt, and Turnbull, 2007; Mitchell, Wilson, and MacKenzie, 2007). On the other hand, the results from the meta-analysis by Parhar, Wormith, Derkzen, and Beauregard (2008) indicated that voluntary correctional programs produced better recidivism outcomes than

those which were mandatory or coercive. As Parhar et al. (2008) acknowledged, however, their study did not control for factors such as program integrity, treatment intensity, or offender recidivism risk.

Implications for Correctional Policy and Practice

Given these limitations, we cannot definitively conclude that the better recidivism outcomes for the pre-2011 participants were due entirely to greater program integrity. At the same time, however, this study is one of the first to closely examine the relationship between program fidelity and reoffending. While future research is needed to arrive at stronger conclusions about the importance of program integrity, we believe the findings still carry several important implications for correctional research, policy, and practice.

First, the results provide additional evidence that cognitive-behavioral programming can be effective in reducing recidivism for offenders. More narrowly, given the consistency between our findings and those from the only other evaluation of Moving On (Gehring et al., 2010), the evidence suggests the gender-responsive program can successfully lower recidivism for female offenders.

Second, this study offers additional, albeit qualified, support for the idea that program integrity matters. It has long been true that many correctional programs fail to work because they are not rooted in sound criminological theory and, thus, exemplify “correctional quackery” (Latessa, Cullen and Gendreau, 2002). It is also true, however, that a common reason for the failure of programs, including those with a solid theoretical foundation, is due to a lack of therapeutic integrity (Cullen and Gendreau, 2000).

Scholars have argued that some of the variation in effectiveness observed among meta-

analyses of correctional programs likely stems from a lack of program integrity (Cullen, 2002; Gendreau, 1996).

While our research is a microcosm of this broader point about the association between program integrity and effectiveness, it also highlights the importance of accounting for program integrity when interpreting the results from individual program evaluations. For example, had we focused only on the 2011-2013 period and assumed the program operated with integrity, we would have been left with the erroneous conclusion that Moving On does not work. Although “black box” evaluations serve their purpose by helping identify what works within corrections, it is also important to look inside the box to better understand why programs fail or succeed.

Third, this evaluation provides evidence that correctional practitioners can take an effective intervention and make it ineffective. The change made to Moving On in 2011 helped ease concerns over scheduling offenders for other institutional programming, but it also led to the implementation of an unsuccessful program that was inconsistent with its original design. The reasons why a program lacks integrity, however, may not always be unintentional. Anecdotally, we are aware of instances in which practitioners have purposefully altered or “enhanced” evidence-based programs (i.e., programs that had achieved positive outcomes in prior research). Moreover, faced with tight budgets, correctional agencies are frequently under pressure to do more with less, which may include offering the “light”, shortened version of a program. Yet, cutting corners to reduce costs in the short-term may ultimately be cost inefficient over the long run by producing worse recidivism outcomes. We are not suggesting, however, that local

program innovation does not have a place in corrections. Rather, efforts to improve program performance should be conducted within the context of controlled experiments.

Regardless of why a program lacks integrity, we believe this study should be viewed as a cautionary tale for correctional practitioners who modify an intervention without regard to program integrity considerations. Making changes that compromise program integrity can have an adverse impact on recidivism outcomes, as our research suggests. But there are also other, more subtle consequences. As the rehabilitative ideal has made a comeback over the last several decades (Cullen, 2005), correctional agencies have generally embraced the idea of using evidence-based practices, i.e., interventions that have been shown to be effective. Indeed, evidence on “what works” with offenders led to the development of the principles of effective correctional intervention and, more narrowly, the RNR model, which is arguably the prevailing paradigm used within North American correctional systems today. Under the RNR model, one of the main goals is to direct offenders to effective programming based on assessments of their recidivism risk and criminogenic needs (Andrews and Bonta, 2010). By providing offenders with evidence-based programming that addresses their criminogenic needs, correctional agencies can presumably help increase public safety through a reduction in recidivism.

Although correctional agencies may believe they are lowering recidivism through the use of effective interventions, this reduction is likely to be elusive if the programs are not delivered with integrity. As a result, using evidence-based interventions without verifying whether they have been implemented with fidelity may promote a false sense of effectiveness. But perhaps more important, offering offenders programming that is unlikely to reduce recidivism because it lacks therapeutic integrity is also wasteful of

correctional resources, which are almost always scarce. Therefore, in the interests of operating more cost-efficient interventions that yield public safety benefits, ensuring the integrity of programming should be a key consideration for correctional agencies.

In late 2013, the MnDOC returned Moving On to the way it operated prior to 2011 but with one notable exception. This time, offenders are being selected for the program based on their likelihood of reoffending, which is consistent with the risk principle. The current version of Moving On within the MnDOC will thus provide another opportunity to not only evaluate program integrity, but also to assess whether adherence to the RNR model and, more narrowly, the risk principle matters for recidivism outcomes. Given the relatively scant research on program integrity to date, much more remains to be learned about its relationship with recidivism outcomes. In particular, rigorous evaluations are needed to further clarify the degree to which program fidelity affects recidivism outcomes and identify whether there are any conditions under which a lack of integrity may be more or less harmful.

References

- Allen, Leana C., Doris L. MacKenzie, and Laura J. Hickman. 2001. The Effectiveness of Cognitive Behavioral Treatment for Adult Offenders: A Methodological, Quality-Based Review. *International Journal of Offender Therapy and Comparative Criminology*, 45: 498-514.
- Andrews, Don A., and Craig Dowden. 2005. Managing correctional treatment for reduced recidivism: A meta-analytic review of programme integrity. *Legal and Criminological Psychology*, 10(2), 173-187.
- Andrews, Don A., and James Bonta. 2010. Rehabilitating criminal justice policy and practice. *Psychology, Public Policy, and Law*, 16(1): 39-55.
- Andrews, Don A., James Bonta, and R. D. Hoge. 1990. Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17(1), 19-52.
- Anglin, M. Douglas, Mary-Lynn Brecht, and Ebrahim Maddahian. 1989. Pretreatment characteristics and treatment performance of legally coerced versus voluntary methadone maintenance admissions. *Criminology*, 27: 537-555.
- Austin, Peter C. (2014). Double propensity-score adjustment: A solution to design bias or bias due to incomplete matching. *Statistical Methods in Medical Research*. DOI: 10.1177/0962280214543508.
- Belknap, Joanne. 2007. *The Invisible Woman: Gender, crime, and justice* (3rd ed.). Belmont, CA: Thompson Wadsworth.
- Brennan, Tim, Markus Breitenbach, William Dieterich, Emily J. Salisbury, and Patricia Van Voorhis. 2012. Women's Pathways to Serious and Habitual Crime: A Person-Centered Analysis Incorporating Gender Responsive Factors. *Criminal Justice and Behavior*, 39(11), 1481-1508.
- Carson, E. Ann. 2014. *Prisoners in 2013*. Bulletin. Washington, DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.
- Cullen, Francis T. 2002. Rehabilitation and treatment programs. In J. Q. Wilson and J. Petersilia (eds.), *Crime and Public Policy*. 2d ed. San Francisco, Calif.: ICS Press.
- Cullen, Francis T. 2005. The twelve people who saved rehabilitation: How the science of criminology made a difference. *Criminology*, 43: 1-42.

- Cullen, Francis T., and Paul Gendreau. 2000. Assessing correctional rehabilitation: Policy, practice, and prospects. In Julie Horney, Ruth Peterson, Doris MacKenzie, John Martin, and Dennis Rosenbaum (Eds.), *Policies, Processes, and Decisions of the Criminal Justice System*. (Vol. 3, pp. 109-175). Washington, D.C.: National Institute of Justice.
- Dowden, Craig, and Don A. Andrews. 1999. What works for female offenders: A meta-analytic review. *Crime & Delinquency*, 45(4): 438-452.
- Duwe, Grant (2014b). An outcome evaluation of a work release program: Estimating its effects on recidivism, employment, and cost avoidance. *Criminal Justice Policy Review*. DOI: 10.1177/0887403414524590.
- Duwe, Grant and Valerie Clark. 2013. Blessed be the social tie that binds: The effects of prison visitation on offender recidivism. *Criminal Justice Policy Review*, 24: 271-296.
- Duwe, Grant, and Deborah Kerschner. 2008. Removing a nail from the boot camp coffin: An outcome evaluation of Minnesota's Challenge Incarceration Program. *Crime & Delinquency* 54: 614-643.
- Duwe, Grant, and Valerie Clark. 2014. The effects of prison-based educational programming on recidivism and employment. *The Prison Journal*.
- Duwe, Grant. (2010). Prison-based chemical dependency treatment in Minnesota: An outcome evaluation. *The Journal of Experimental Criminology* 6: 57-81.
- Duwe, Grant. (2012). The benefits of keeping idle hands busy: The impact of a prisoner reentry employment program on post-release employment and offender recidivism. *Crime & Delinquency*. DOI: 10.1177/0011128711421653
- Duwe, Grant. (2014a). The development, validity, and reliability of the Minnesota Screening Tool Assessing Recidivism Risk (MnSTARR). *Criminal Justice Policy Review*, 25, 579-613.
- Gehring, Krista, Patricia Van Voorhis, and Valerie R. Bell. 2010. "What Works" for female probationers?: An evaluation of the Moving On Program. *Women, Girls, and Criminal Justice*, 11(1): 6-10.

- Gendreau, Paul and Don A. Andrews. 1990. Tertiary prevention: What the meta-analyses of the offender treatment literature tell us about what works. *Canadian Journal of Criminology*, 32, 173-184.
- Gendreau, Paul, and Don A. Andrews. 1994. *Correctional Program Assessment Inventory* (4th ed.). St. John, New Brunswick, Canada: University of New Brunswick.
- Gendreau, Paul, Sheila French, and Angela Gionet. 2004. What works (what doesn't work): The principles of effective correctional treatment. *Journal of Community Corrections*, 13(3): 4-6.
- Gendreau, Paul. 1996. The principles of effective interventions with offenders. In A. T. Harland (ed.), *Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply*. Thousand Oaks, Calif.: Sage.
- Grady, Melissa D., Daniel Edwards, Carrie Pettus-Davis, and Jennifer Abramson. 2012. Does volunteering for sex offender treatment matter? Using propensity score analysis to understand the effects of volunteerism and treatment on recidivism. *Sexual Abuse: A Journal of Research and Treatment*, 25: 319-346.
- Greiner, Leigh E., Moira A. Law, and Shelley L. Brown. 2014. Using Dynamic Factors to Predict Recidivism among Women: A Four-Wave Prospective Study. *Criminal Justice and Behavior*. Online First, DOI: 10.1177/0093854814553222
- Holtfreter, Kristy, and Katelyn A. Wattanaporn. 2013. The Transition from Prison to Community Initiative An Examination of Gender Responsiveness for Female Offender Reentry. *Criminal Justice and Behavior*, 41(1): 41-57.
- Knight, Kevin, Matthew L. Hiller, Kirk M. Broome, and Dwayne Simpson. 2000. Legal pressure, treatment readiness, and engagement in long-term residential programs. *Journal of Offender Rehabilitation*, 31: 101-115.
- Latessa, Edward T. 2012. *Evaluating Correctional Programs*. Unpublished Paper. Washington, DC: National Institute of Corrections Library: ID 027239.
- Latessa, Edward T., and Christopher Lowenkamp. 2005. What are criminogenic needs and why are they important. *For the Record*, 4: 15-16.

- Latessa, Edward T., Francis T. Cullen, and Paul Gendreau. 2002. Beyond correctional quackery—Professionalism and the possibility of effective treatment. *Federal Probation*, 66: 43-49.
- Lipsey, Mark W., and Francis T. Cullen. 2007. The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews. *Annual Review of Law and Social Science*, 3: 297-320.
- Lipsey, Mark W., Gabrielle L. Chapman, and Nana A. Landenberger. 2001. Cognitive-Behavioral Programs for Offenders. *The Annals of the American Academy of Political and Social Science*, 578: 144-157.
- Lipsey, Mark W., Nana A. Landenberger, and Sandra J. Wilson. 2007. Effects of cognitive-behavioral programs for criminal offenders. Campbell Systematic Reviews: 6. DOI: 10.4073/csr.2007.6.
- Lowenkamp, Christopher T., Latessa, Edward J., and Smith, Paula. 2006. Does Correctional Program Quality Really Matter? The impact of adhering to the principles of effective intervention. *Criminology & Public Policy*, 5(3): 575-594.
- Makarios, Matthew, Benjamin Steiner, and Lawrence F. Travis. 2010. Examining the Predictors of Recidivism among Men and Women Released from Prison in Ohio. *Criminal Justice and Behavior*, 37: 1377-1391.
- Matthews, Betsy, Dana Jones Hubbard, and Edward Latessa. 2001. Making the next step: Using evaluability assessment to improve correctional programming. *The Prison Journal*, 81(4): 454-472.
- McSweeney, Tim, Alex Stevens, Neil Hunt, and Paul J. Turnbull. 2007. Twisting arms or a helping hand? Assessing the impact of “coerced” and comparable “voluntary” drug treatment options. *British Journal of Criminology*, 47: 470-490.
- Mitchell, Ojmarrh, David B. Wilson, and Doris L. MacKenzie. 2007. Does incarceration-based drug treatment reduce recidivism? A meta-analytic synthesis of the literature. *The Journal of Experimental Criminology*, 3: 353-375.
- Nesovic, Aleksandra. 2003. *Psychometric evaluation of the correctional program assessment inventory*. Unpublished Dissertation. Dissertation Abstracts International, 64(9).

- Palmer, Ted. 1995. Programmatic and Nonprogrammatic Aspects of Successful Intervention: New Directions for Research. *Crime & Delinquency*, 47(1): 100-131.
- Parhar, Karen K., J. Stephen Wormith, Dena M. Derkzen, and Adele M. Beauregard. 2008. Offender coercion in treatment: A meta-analysis of effectiveness. *Criminal Justice and Behavior*, 35: 1109-1135.
- Pearson, Frank S., Douglas S. Lipton, Charles M. Cleland, and Dorline S. Yee. 2002. The effects of behavioral/cognitive-behavioral programs on recidivism. *Crime & Delinquency*, 48: 476-496.
- Penduzzi, P., Cocato, J., Kemper, E., Holford, T.R., & Feinstein, A.R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology* 49: 1,373-1,379.
- Prosavac, Emil J., and Raymond G. Carey. 1992. *Program evaluation: Methods and case studies* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Rosenbaum, Paul R., and Donald B. Rubin. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39: 33-38.
- Rubin, Donald B. 1997. Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine*, 127: 757-763.
- Scroggins, Jennifer R., and Sara Malley. 2010. Reentry and the (Unmet) Needs of Women. *Journal of Offender Rehabilitation*, 49(2): 146-163.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Sperber, Kimberly Gentry, Latessa, Edward J., and Makarios, Matthew D. 2013. Examining the Interaction between Level of Risk and Dosage of Treatment. *Criminal Justice and Behavior*, 40: 338-348.
- Trevisan, Michael S., and Yi Min Huang. 2003. Evaluability assessment: A primer. *Practical Assessment, Research & Evaluation*, 8(20): 2-9.
- Van Dieten, Marilyn. 2010. *Moving On: A Program for At-Risk Women. Modules 1 and 6 Facilitator's Guide*. Center City, MN: Hazeldon.

- Van Voorhis, Patricia, and Kelly Brown. 1996. *Evaluability assessment: A tool for program development in corrections*. Report. Washington, D.C.: National Institute of Corrections.
- Van Voorhis, Patricia, Emily M. Wright, Emily Salisbury, and Ashley Bauman. 2010. Women's Risk Factors and Their Contributions to Existing Risk/Needs Assessment: The current status of gender-responsive supplement. *Criminal Justice and Behavior*, 37(3): 261-288.
- Vittinghoff, E. & McCulloch, C.E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology* 165: 710-718.
- Wilson, David B., Leana Allen Bouffard, and Doris L. MacKenzie. 2005. A Quantitative Review of Structured, Group-Oriented, Cognitive-Behavioral Programs for Offenders. *Criminal Justice and Behavior*, 32(2): 172-204.
- Wright, Emily M., Patricia Van Voorhis, Emily J. Salisbury, and Ashley Bauman. 2012. Gender-Responsive Lessons Learned and Policy Implications for Women in Prison: A Review. *Criminal Justice and Behavior*, 39(12):1612-1632.