

Follow-up Evaluation of Ohio's Community Based Correctional Facility and Halfway House Programs

Program Characteristics Supplemental Report

Edward J. Latessa, Ph.D.
Professor and Director

Lori Brusman Lovins, MSW
Project Director

Paula Smith, Ph.D.
Assistant Professor

Center for Criminal Justice Research
University of Cincinnati
School of Criminal Justice
Cincinnati, Ohio 45221-0389

Matthew Makarios, Ph.D.
Assistant Professor
Criminal Justice Department
University of Wisconsin - Parkside
Kenosha, WI 53141-2000

July 2010

Table of Contents

INTRODUCTION	3
SECTION I: THE PRINCIPLES OF EFFECTIVE INTERVENTION.....	3
SECTION II: METHODOLOGY	8
Sample	9
Procedures for Data Collection.....	10
<i>Offender Level Data</i>	10
<i>Program-level Data</i>	11
<i>Program-level Measures</i>	12
Analysis	13
<i>Weighting</i>	14
SECTION III: RESULTS	15
Program Leadership and Development.....	16
Staff Characteristics	19
Assessment.....	22
Treatment	26
Evaluation	32
Overall Scale.....	35
SECTION IV: DISCUSSION.....	37
REFERENCES	41
APPENDIX A.....	44

INTRODUCTION

The University of Cincinnati, School of Criminal Justice was contracted in 2006 by the Ohio Department of Rehabilitation and Correction (ODRC) to conduct a follow-up evaluation of the state's halfway house (HWH) and Community Based Correctional Facility (CBCF) programs¹. This study is designed as a follow-up to the 2002 study, with the following research questions in mind:

1. *What type of offenders benefit most from programming?*
2. *Which programs are most effective at reducing recidivism?*
3. *What models or program characteristics are most important in reducing recidivism?*

The follow-up HWH/CBCF outcome study was released in March of 2010 (Latessa, Brusman Lovins & Smith, 2010). This report addressed the first two questions posed above. What follows serves as a supplemental report designed to answer the third research question. In doing so, data based upon an in-depth examination of program characteristics is presented to determine what program factors are most important in reducing recidivism.

The current report will be organized into the following subsections: Section I provides a brief literature review of the principles of effective intervention; Section II provides a summary of the methodology used for this study; Section III outlines the results of this portion of the study, specifically the program characteristics associated with reduced rates of recidivism; and Section IV summarizes the findings for this study and identifies limitations of the research.

SECTION I: THE PRINCIPLES OF EFFECTIVE INTERVENTION

Correctional research has shown ample support for rehabilitation as a strategy to decrease offender recidivism (Losel, 1995). As a whole, correctional treatment programs reduce

¹ The views expressed in this report are not necessary those of ODRC.

recidivism by approximately 10 percent (Andrews, Zinger, Hoge, Bonta, Gendreau & Cullen, 1990; Gendreau, French, & Gionet, 2004). However, there is great variation among programs with respect to the degree of effectiveness (Gendreau & Ross, 1987; Lipsey, 1992; Gendreau 1996; Latessa & Lowenkamp, 2002, Pearson, Lipton, Cleland, & Yee 2002; Lowenkamp, Latessa, & Smith, 2006). Programs that make use of evidence-based treatment strategies see effects two to four times greater than the 10 percent average (Gendreau, French & Gionet, 2004). To the contrary, well-intentioned treatment providers have also *increased* the likelihood of recidivism by not utilizing evidence-based strategies (Dowden & Andrews, 1999; Lowenkamp 2004; Lowenkamp, Latessa & Lemke, 2006) or attempting to employ such strategies, but with poor fidelity (Andrews & Dowden, 1999; Barnoski, 2004). Research suggests that certain characteristics are associated with effective treatment programs (Andrews et al., 1990, Gendreau, 1996). From this literature, the “principles of effective intervention” were developed. These principles are perhaps most simply defined as the risk, need and responsivity principles.

The risk principle suggests that higher risk offenders receive supervision and treatment comparable to their risk level. Several primary studies as well as meta-analyses have shown support for this principle. A key finding from both the initial and follow-up Ohio HWH/CBCF study was that higher risk offenders were significantly more likely to benefit from residential intervention than lower risk participants (Lowenkamp & Latessa, 2002; Latessa, Brusman Lovins and Smith, 2010). Also using Ohio data in a study of 97 residential and non-residential correctional programs, Lowenkamp, Latessa and Holsinger (2006) found that programs that provided at least .5 more units of services or referrals for higher risk offenders (increased intensity) and programs that kept higher risk offenders longer (increased duration) were more effective at reducing recidivism. Dowden and Andrews (2000) examined 52 studies to determine

the impact of the risk principle on violent offending; they found that programs that adhered to the risk principle produced an effect size of .09 versus .04 for programs that ignored this principle. A second meta-analytic review conducted by these authors in 2006 also showed moderate support for the risk principle (Dowden & Andrews, 2006).

The need principle asserts that criminogenic needs, or those offender needs related to criminal offending should be the primary treatment targets. Andrews and Bonta (2007) have identified the core criminogenic need areas: antisocial attitudes, antisocial associates, antisocial personality features, family relationships, substance abuse, educational/vocational achievement, and structured leisure time. In the Dowden and Andrews's meta-analysis (2006), support for the need principle was also found; programs that mainly targeted criminogenic needs had an effect size of .20 as compared to .00 for programs that targeted primarily non-criminogenic needs. Furthermore, a meta-analysis conducted by Gendreau, French and Taylor (2002) found that programs that targeted 4 to 6 more criminogenic than non-criminogenic needs had an appreciable effect size ($ES=.31$). Finally, Lowenkamp (2004) found not only that targeting criminogenic need factors was positively correlated with program effectiveness ($r=.39$), but that programs using standardized methods to assess risk and need were also more effective at reducing recidivism ($r=.33$).

The responsivity principle consists of both general and specific responsivity. The specific responsivity principle urges that treatment should vary based upon the individual learning styles of the offenders (see Andrews & Bonta, 2007). General responsivity suggests that most offenders respond to behavioral interventions and correctional treatment programs should therefore use a cognitive-behavioral treatment model. There is an abundance of research supporting the use of a cognitive-behavioral approach to change offender behavior and reduce the likelihood of

recidivism (Gendreau & Ross, 1987; Andrews et al., 1990; Antonowicz & Ross, 1994; Henning & Frueh, 1996; Lipsey & Wilson, 1998; Andrews & Dowden, 1999; Baro, 1999; Lipsey, Chapman & Landenberger, 2001; Lipsey, 2001, 2009; Hanson, Gordon, Harris, Marques, Murphy, Quinsey & Seto, 2002; Pearson et al., 2002, Lowenkamp, 2004; Landenberger & Lipsey, 2005; Wilson, Bouffard & MacKenzie, 2005; French & Gendreau, 2006, Lowenkamp, Hubbard, Makarios, & Latessa, 2009).

Landenberger and Lipsey (2005) conducted an analysis using 58 studies of adults and juveniles, and found that on average, cognitive behavioral programs reduced recidivism by 25 percent. However, when the most effective configurations were amassed (e.g., high completion rate, frequent sessions, combining cognitive-behavioral treatment with other services), programs saw more than a 50 percent reduction in recidivism. Recently, Lowenkamp, Hubbard, Makarios, and Latessa (2009) studied the effects of a “real world” application of Thinking for a Change (TFAC)² in a community corrections setting. They found that offenders that participated in TFAC were rearrested at a significantly lower rate (28%) than a comparison sample (43%), controlling for risk level, age, race, and time at risk.

In addition to a program’s need to meet the risk, need and responsivity principles, to be effective, programs must deliver services with high fidelity. Barnoski (2004) conducted an outcome evaluation for The Washington State Institute on Public Policy of two evidence-based treatment programs: Functional Family Therapy (FFT) and Aggression Replacement Training (ART)³. This study found that when delivered competently, ART reduced recidivism by

² Thinking for a Change is a popular cognitive-behavioral curriculum developed for a correctional population, funded by the National Institute of Corrections and authored by Bush, Glick, and Taymans (1997). The curriculum uses a cognitive-behavioral model and includes a cognitive restructuring, social skill development and problem solving component.

³ Both Functional Family Therapy (Gordan et al., 1998) and Aggression Replacement Training (Goldstein, Glick and Gibbs, 1998) have empirical data to support their effectiveness with criminal justice populations.

approximately 24 percent and FFT by 38 percent relative to the comparison group; however, when these programs were not delivered competently, ART *increased* the rate of recidivism by 10 percent and FFT *increased* recidivism by nearly 17 percent.

Another meta-analysis by Andrews and Dowden (1999) identified several key factors that influenced program integrity: following a specific program model, training workers, supervising workers, using printed manuals, monitoring change, administering an adequate dosage of treatment, and having an involved evaluator. Dowden and Andrews (2004) more closely examined the importance of staff practice in delivering effective correctional treatment via a meta-analysis on core correctional practice (CCP)⁴. They found that core correctional practice variables made an independent contribution to increased effects in human service programs, but only if programs were already adhering to the risk, need and responsivity principles.

Considering the relevance of program characteristics in reducing recidivism, it is not surprising that objective actuarial assessments of treatment integrity have surfaced. In an attempt to provide a standardized measure of a program's adherence to the principles of effective correctional intervention, Gendreau and Andrews (1989) developed the Correctional Program Assessment Inventory⁵. Recent research indicates that measures from the CPAI are directly related to rates of program recidivism. Lowenkamp, Flores, Holsinger, Makarios, and Latessa (2010) administered a modified version of the CPAI to intensive supervision programs in Ohio and found that treatment integrity displayed a correlation of .56 with program effect size. Also,

⁴ Andrews and Kiessling (1980) identified five dimensions of effective correctional practice designed to increase the therapeutic potential of rehabilitation programs. The five dimensions include: effective use of authority, anticriminal modeling and reinforcement, problem solving, use of community resources, and quality of interpersonal relationships between staff and clients.

⁵ The administration of the CPAI involves using information gathered during site visits to score each facility on 65 items measuring 6 substantive areas: program implementation, client pre-service assessment, characteristics of the program, characteristics of practices and staff, program evaluation, and a final category with a variety of miscellaneous items.

Lowenkamp (2004; see also Lowenkamp, Latessa, & Smith 2006) administered an abbreviated version of the CPAI in the original Ohio HWH/CBCF study and found that treatment integrity was able to explain nearly 40 percent of the variation of program effect size.

In sum, research has found that actuarial measures of treatment integrity are strongly related to program effectiveness. It suggests that using a standardized approach to gathering characteristics of multiple programs can be used to determine why some programs are more effective than others in reducing recidivism. Further, the principles of effective intervention (risk, need and responsivity) provide researchers and practitioners with a blueprint for creating effective correctional treatment programs. When these principles are met, and met with high fidelity, programs can experience substantial treatment effects.

What follows is an in-depth examination of how these principles (as well as supporting program characteristics) played a role in recidivism reductions among offenders participating in Ohio HWHs or CBCFs. These results should be understood as providing information as to why different programs may have been effective and why others were not. They are intended to assist facility administrators in designing program elements that increase the likelihood of program effectiveness.

SECTION II: METHODOLOGY

This section of the report will highlight the methods used for data collection and analysis. As a supplemental report, a brief review of the sample and the method used to collect the individual and program level data will be provided⁶. Likewise, the statistical techniques used to analyze the program characteristics will be described.

⁶ A more in-depth review can be found in the primary 2010 outcome study (Latessa, Lovins & Smith, 2010).

Sample

The 2010 CBCF/HWH outcome study incorporated over 20,000 offenders. Separate analyses were conducted by treatment sample (CBCF versus HWH), comparison sample [Intensive Supervision Probation (ISP) versus parole/Post Release Control (PRC)], risk level (low, moderate, high), completion status (successful or unsuccessful) and outcome measure (new felony conviction, new misdemeanor or felony conviction, and incarceration). Since the purpose of this supplemental report is to identify what program characteristics correlate with reduced rates of recidivism, the analyses were simplified. HWH and CBCF data were combined and the CBCF/ISP group was used to represent CBCFs⁷. Furthermore, only successful program completers were examined, and the primary outcome variable used was any new conviction.

Hence, the participants in this study consisted of offenders that successfully completed an Ohio CBCF or HWH program between February 1, 2006 and June 1, 2007⁸. These offenders were then matched to comparison cases as follows: 1) CBCF or HWH *probationers* were matched to offenders placed on ISP 2) *parole/PRC* and *Transitional Control* HWH treatment cases were matched to parolee/PRC offenders not exposed to either HWH or CBCF intervention⁹. Comparison cases were matched on *gender* (male/female), *race* (White/non-White), *sex offender status* (sex offender/non-sex offender), *county* (large, medium and small)

⁷ The 2010 report provided separate analyses for CBCF participants matched to ISP cases and CBCF participants matched to parole/PRC offenders. Since there are a significant amount of duplicate treatment cases between these two groups, one had to be chosen for use in this report. The CBCF/ISP group was chosen over the CBCF/parole group as CBCFs are designed to serve probationers, and matching probationers to probationers was considered a more similar comparison.

⁸ Program level data collection occurred from August 2006 to December 2006. The February 2006 to June 2007 dates represent a one year time from around the beginning and end of the program level data collection. The treatment sample was derived by identifying all offenders participating in the program six months before and after the date of each site visit. However, only successful treatment completers, as identified in the CCIS database, were used for the program level analyses.

⁹ In the 2010 outcome study, the CBCF sample was compared to both a parole/PRC and ISP comparison sample. For purposes of this report, only the CBCF and matched ISP group was examined.

and *risk* (low, moderate, and high).¹⁰ The matching process resulted in a one-for-one match between treatment and comparison cases, using the identified matching variables¹¹.

The CBCF offenders participated in one of 20 Ohio CBCF programs in operation in 2006, and the HWH offenders participated in one of 44 Ohio HWH programs. Only successful treatment completers and their matched comparison cases were used for the program-level analyses. The CBCF/ISP sample consists of 5,692 cases, while the HWH sample has 6,580 combined treatment and comparison cases. Hence, the analyses that follow are based upon the characteristics of 64 programs and 12,272 offenders.

Procedures for Data Collection

Offender Level Data

Individual level offender data for the CBCF, HWH and ISP samples were extracted from the Community Corrections Information System (CCIS) maintained by the ODRC. For the parolee/PRC sample, data came from the Department's Offender Tracking System (DOTS-PORTAL) database, ODRC's main inmate database. These data included demographic characteristics, current offense, offense history, county of conviction, identified needs, services delivered, termination type, and employment. Since there was not a consistent risk measure used across the state at the time of the study, data from these systems were used to create a risk scale. Included in the risk measure was: prior incarcerations, prior conviction, offense level, offense category, substance abuse problem, employment problem and age¹². This scale was used to match offenders by risk.

¹⁰ See the primary 2010 study for a full description of the matching process and risk measure created for the study.

¹¹ Since the same pool of ISP comparison cases was used to match both the HWH and CBCF treatment samples, there are some duplicate ISP comparison cases. In all, only 7.6 percent (938 out of 12,272) cases in the sample were duplicates.

¹² See the Appendix A of the 2010 CBCF/HWH outcome study for a full description of the risk tool.

Recidivism data for both the experimental and comparison groups were collected by University of Cincinnati researchers via the Ohio Law Enforcement Gateway (OHLEG) system. Two years was the follow-up timeframe used to collect the recidivism data. While several recidivism measures were employed in the 2010 HWH/CBCF outcome study, any new misdemeanor or felony conviction was the primary outcome variable used to identify the program characteristics that were related to reductions in recidivism. This outcome variable was chosen over felony conviction or new incarceration since it was deemed a valid measure of whether the offender engaged in any new criminal behavior. Furthermore, the new incarceration variable included offenders incarcerated for either a new crime or for technical violations. This is especially problematic in the current research because offenders in a residential placement are likely to have more stringent conditions of supervision and increased surveillance. As a result, having the offenders in treatment by definition increases the likelihood of technical violations and may confound the results of this research.

Program-level Data

Each of the 64 identified CBCF/HWH facilities was visited by a University of Cincinnati research team between August and December of 2006. Site visits were scheduled on days that treatment groups could be observed. Site visits consisted of structured interviews with program director(s), treatment providers, a sample of custody staff and supervisors, and program participants. Offender files, as well as program materials such as treatment manuals, assessments, handbooks, and policies and procedures were also reviewed during site visits. Finally, treatment groups were observed and facilitation skills coded. At the conclusion of the site visit, the research team compiled all materials from the site visit and collectively completed a program summary form. A database with 1,038 variables was created from the program

summary form that identifies each actuarial measure captured during the site visits from all data collection sources.

Program-level Measures

Program level measures used for the current analyses were primarily ascertained from items found on the Evidence Based Correctional Program Checklist (CPC). This instrument is designed to examine how closely correctional programs meet the known principles of effective intervention¹³. Although the items on this instrument served as the framework for creating the data collection tools used for the current study, a more systematic approach was taken to code in-depth information on all core program elements. The CPC evaluates the following areas: 1) program leadership and development, 2) staff characteristics, 3) assessment, 4) treatment, and 5) quality assurance. Data were collected in these five areas and can be described as follows.

Program leadership and development considers the educational and professional experience of the program director. Further, there are items that address the program director's involvement in the development of the program, as well as the selection of staff and delivery of services. Items related to program funding and sustainability, as well as piloting of programs before full implementation are also considered. The *staff characteristics* domain identifies the educational and professional experience of the treatment staff. In addition, staff training as well as support and attitudes of the staff regarding the programming are assessed. Finally, this domain identifies whether or not there is clinical supervision provided to the staff.

Offender assessment considers whether or not the program is using an actuarial, standardized risk/need assessment that is valid for their target population and is used to identify

¹³ The CPC is modeled after the Correctional Program Assessment Inventory developed by Gendreau and Andrews; however, the CPC includes a number of items not contained in the CPAI. In addition, items that were not found to be positively correlated with recidivism were removed.

appropriate offenders for programming. Likewise, the program should assess a range of key responsivity factors using a validated tool. The assessment section also evaluates whether the program has clear eligibility/exclusionary criteria. The items under the *treatment characteristics* domain examine: (1) whether the primary treatment targets of the program are criminogenic; (2) if the program model is centered around social learning and/or cognitive-behavioral theory; (3) that staff and offenders are appropriately matched to programming based on specific responsivity factors; (4) that dosage is appropriate based on the risk level of the offender; (5) that the types of rewards and punishers given as well as the process for doing so are appropriate; (6) that behavioral strategies are employed to change offender behavior; (7) whether the program trains family members and offers an aftercare component; and 8) the method for determining successful program completion.

Finally, items under the *quality assurance* domain reflect the internal and external review strategies employed by a program to maintain the treatment model, including observation of service delivery and surveying client satisfaction with the program. Additional quality assurance items include whether offenders are reassessed, as well as whether the program has undergone process and/or outcome evaluations, and the results of such assessments.

The analyses that follow are reflective of each of these five domains. Individual items are identified in each of the domains, as well as scales that demonstrate how the items predict outcome when combined. A total scale was also created to ascertain how each of the five sub-domains relates to reductions in recidivism.

Analysis

In order to identify what program characteristics were related to reductions in recidivism, differences in the average recidivism rate between treatment and matched comparison cases were

calculated for each program. These differences can be understood as effect sizes in that they indicate the effect that the treatment had on recidivism. Mean effect sizes were then compared between programs with or without an identified program characteristic and correlations were used to examine the strength of the relationship. Individual program characteristic items that were theoretically important and found to have a significant relationship with program effectiveness were included in scales used to measure each of the five CPC domains listed above. Also included in the scales were some factors that did not reach statistical significance, based in part on the small sample size ($n=64$). Hence, non-significant items that approached a correlation of .05 *and* were theoretically relevant were also included as they provided a substantive contribution to the scale.

A correlation coefficient, or r , was used to identify the strength of the relationship between the program characteristic and program effect size because it is widely understood and easy to interpret. The correlation coefficient r ranges between -1 and 1. Negative values are associated with increases in the likelihood of recidivism, whereas positive values denote that the program characteristic is associated with a reduction in recidivism. A value at or near zero indicates that the item has little to no relationship with changes in recidivism.

Weighting

When examining effect size estimates that come from programs with different sample sizes, it is important to weight estimates so that programs with more cases are given greater weight (Rosenthal, 1991; Wolf, 1986). This is based on the assumption that larger sample sizes produce more reliable estimates and should thus be given greater weight. One method of doing

this is to weight effect size estimates by the number of cases in each sample.¹⁴ An important critique of this approach is that weighting by sample size can dramatically reduce the standard error of the estimates since N becomes the total number of participants (12,272) instead of the number of actual programs (64). In order to address this issue, the current research first standardizes the weights so that the mean of the weights is 1 and that N equals 64. This is done by taking the total number of programs and dividing by the total sum of the unstandardized weights. In essence this produces unbiased estimates while at the same time making the N for the analyses equal to the total number of programs. In the current study, effect sizes are weighted using standardized weights before analyses are conducted that examine differences in mean effect sizes and when correlations between program characteristics and effect size estimates are examined.

SECTION III: RESULTS

The results section of this report outlines how various program characteristics correlate with program effectiveness. As mentioned in the methods section, program effectiveness is measured by comparing differences in the rates of recidivism between treatment and comparison cases for each program. The results will be subdivided into five sections: Program Leadership and Development; Staff Characteristics; Assessment; Treatment; and Evaluation/Quality Assurance. Statistics from thirty-four distinct program characteristics will be displayed, as well as scales for each subsection that demonstrate the cumulative effects of adhering to multiple characteristics that correlate with reductions in recidivism. A table is presented for each of the five sub-domains along with a corresponding figure for the cumulative sub-domain scale. The average percentage change in recidivism between treatment and comparison cases is identified

¹⁴ A common method of doing this is weighting each effect size estimate by $n - 3$ where n equals the number of cases in each program. Weighting each effect size estimate (64) by $n-3$ reduced the sample size to 12,080.

for each respective program characteristic. Likewise, the amount of variation on each item is identified by examining the number of programs (N) that met each program characteristic. Furthermore, the strength of the relationship between the item and outcome measure is measured using Pearson's Correlation Coefficient (r).

Program Leadership and Development

Table 1 displays the Program Leadership and Development variables. The first item in this domain is age of the program. The data suggest that well established programs are more likely to reduce recidivism than newer programs. Program age 0-15 showed a 2.6% reduction in recidivism versus programs age 16 to 29 (3.9% reduction) and programs over 29 years old (6.3% reduction). The correlation between program age and new conviction is .15. Table 1 also shows that a significant reduction in recidivism is associated with programs whose directors regularly allocate time toward direct supervision of staff. Program directors that conduct fewer than five hours of structured supervision per month had less than a 1 percent (.4%) reduction in recidivism versus a 5.3 percent reduction for programs with a director spending at least 5 hours per month conducting structured supervision. Structured supervision might entail regularly attending treatment team meetings and offering feedback on cases being reviewed, scheduling supervisory meetings with staff responsible for overseeing clients, or scheduling individual supervision meetings with staff.

Table 1 also examines how the gender of the clientele served by the program correlates with outcome. For this item, a coed scale was created. Programs that were not coed were scored

Table 1: Program Leadership and Development Items

Items	% Change in Recidivism	N	r (weighted)
Age			0.15
0-15	2.6	40	
16-29	3.9	13	
30+	6.3	11	
PD Supervision*			0.29
< 5 hours/mo	0.4	23	
5+ hours/mo	5.3	41	
Coed Scale*			0.17
-1	1.6	14	
0	3.0	34	
1	6.1	14	
PD Qualified			0.07
No	2.7	33	
Yes	3.9	31	
Literature Review			0.06
No	3.3	63	
Yes	6.9	1	
Stable Funding			0.08
No	2.3	19	
Yes	3.8	45	
Leadership Categories**			0.41
0 to 1	-3.8	15	
2 to 3	4.6	37	
4 to 5	7.3	12	

* significant difference at the .05 level
 ** significant difference at the .01 level

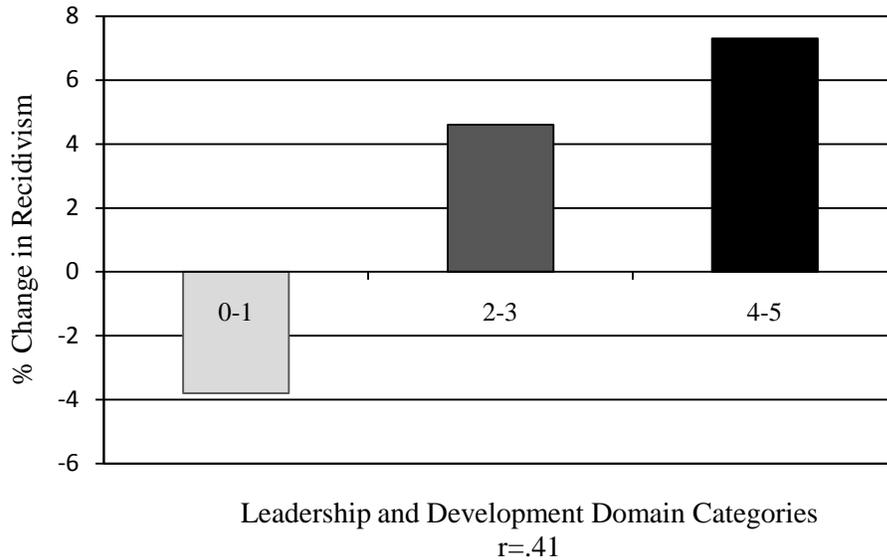
as 0. Programs earned a point for treating both males and females, as long as living space was not shared (meaning occupied at the same time). However, a point was deducted if the program was coed and living space was shared. Coed programs that shared living space had a reduction in recidivism of 1.6 percent. Programs that served only males or only females had an average 3

percent reduction in recidivism. Finally, programs that served both genders and maintain separate living spaces within the facility reduced recidivism by 6 percent on average. This difference was significant and the item had a .17 correlation with recidivism. This suggests that those programs that treat both genders and successfully provide separate accommodations for them have better outcomes.

Other items in this subcategory include the qualifications of the program director. Programs operated by directors that had at least a bachelor's degree in a helping profession were slightly more effective at reducing the number of convictions (3.9% versus 2.7% reduction in recidivism). Likewise, although variation on this item was limited, programs that conducted a thorough review of the literature when selecting program materials and regularly disseminated current literature on evidence based practices to program staff had better outcomes (3.3% difference versus 6.9% difference). Finally, programs whose funding was stable also produced a better outcome (3.8% reduction) as opposed to programs that felt unstable funding affected the ability of the program to operate smoothly (2.3% reduction in recidivism). Note that none of these last three items showed significant differences in outcome and the corresponding correlation between these program characteristics and outcome was relatively weak (.06 to .08).

Figure 1 demonstrates the cumulative effects of meeting multiple items within the Program Leadership and Development domain. Programs that scored between 0-1 points in this section *increased* the likelihood of recidivism by 3.8 percent (N=15). Programs scoring 2-3 points decreased recidivism by 4.6 percent (N=37). Finally, programs scoring 4-5 points in this section decreased recidivism by 7.3 percent (N=12). These differences are significant and the overall correlation between this categorical scale and outcome is .41.

Figure 1: Program Leadership and Development Categories



Staff Characteristics

Table 2 displays the Staff Characteristics variables. The first item in this domain measures characteristics that programs look for when hiring new staff. Six desired staff characteristics were associated with improved outcome: 1) assertive/directive; 2) firm but fair; 3) won't get walked on by offenders; 4) problem solving skills; 5) paperwork skills; and 6) computer skills. The staff skills scale is coded so that a point is awarded for each of the six items. Programs that looked for at least 4 of these characteristics when hiring new staff had better outcomes (a 15% reduction in recidivism) over programs that looked for between 1 and 3 of these characteristics (4.6% recidivism reduction) or none of them (.3% increase in recidivism). This item was significantly related to outcome with a correlation of .28. This finding suggests that looking for a range of appropriate staff skills/values such as coping skills (problem solving ability), interpersonal skills (assertiveness); philosophy in corrections (firm but fair) and job skills (paperwork/computer skills) is important.

Table 2: Staff Characteristics Items

Items	% Change in Recidivism	N	r (weighted)
Staff Skills Scale**			0.28
0	-0.3	16	
1 to 3	4.6	46	
4+	15.0	2	
Clinical Meetings Scale*			0.24
0 to 3	1.3	24	
4 to 5	5.3	40	
Meeting Trainings*			0.24
No	0.9	22	
Yes	5.1	42	
Initial Training--Time			0.20
Above or Below 60-90 Hrs	2.0	37	
60-90 Hrs	5.6	27	
Initial Training--Treatment Oriented*			0.22
Above or Below 20-30%	2.4	46	
20-30%	6.7	18	
Clinical Supervision by Certified Staff**			0.28
No	2.2	49	
Yes	8.4	15	
Staff Characteristics Categories**			0.42
0 to 1	-1.4	7	
2 to 4	2.9	49	
5 or more	12.9	8	

* significant difference at the .05 level

** significant difference at the .01 level

The next Staff item examines the nature of clinical or treatment team meetings held at a facility. A scale was also created for this item where a point was awarded for each of the following meeting characteristics: 1) 4 or more meetings held per month; 2) case files reviewed during meetings; 3) meetings attended by the program director; 4) meetings attended by case

managers; and 5) meetings attended by security staff. Programs with clinical meetings with 3 or fewer of the above characteristics had an average reduction in recidivism of 1.3 percent. Programs meeting 4 or 5 of these characteristics reduced recidivism by an average of 5.3 percent, which does reflect a significant difference. This item also held a relatively high correlation with outcome with an r-value of .24. This finding suggests that having frequent meetings is important as well as using a multi-disciplinary team to review cases.

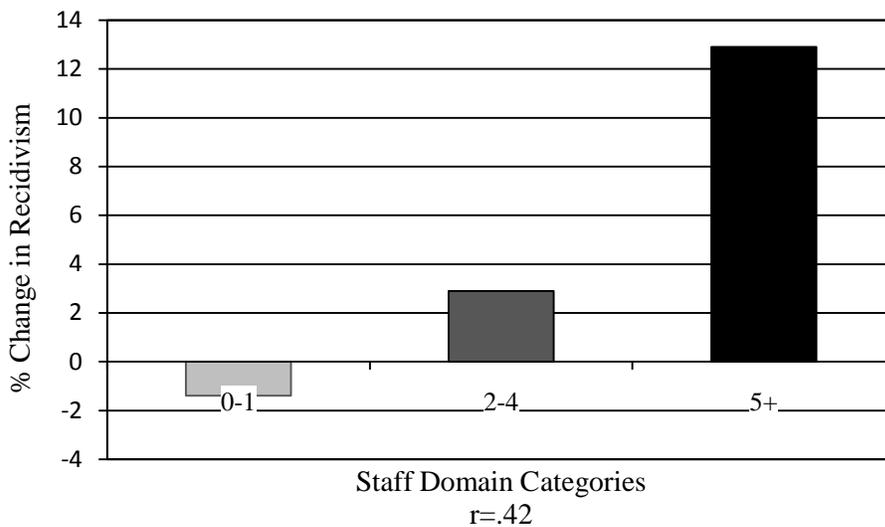
The next three items in Table 2 examine program training. The “Meeting Trainings” item measured whether formal training was incorporated into agency meetings. Programs that offered training in at least one meeting per month had better outcome (5.1% reduction in recidivism) than programs that did not (0.9% recidivism reduction). This item, along with the above “clinical meetings scale” item suggests that periodically structuring meetings with file review and training increased program effectiveness.

The next two items examined initial training efforts by programs. Maximum benefit was obtained from programs that required between 60 and 90 hours of initial training, which includes formal training and on-the-job training (5.6% reduction in recidivism). Programs that required either more or less than this amount had weaker outcomes (2% recidivism reduction). Furthermore, programs appeared to benefit most when 20-30 percent of the initial training was treatment-oriented (6.7 percent reduction in recidivism). This suggests that new staff need training on agency policies and procedures as well as clinically related topics.

The final item in the Staff Characteristics section involves clinical supervision. This item suggests that programs that offer clinical supervision by a certified staff person experience an average recidivism reduction of 8.4 percent versus 2.2 percent for programs that do not. This difference in recidivism is significant, and the r-value is modest at .28.

Figure 2 demonstrates the cumulative effects of meeting multiple items within the Staff Characteristics domain. Programs that scored between 0 and 1 points in this section increased the likelihood of recidivism by 1.4 percent (N=7). Programs scoring 2 to 4 points decreased recidivism by 2.9 percent (N=49). Finally, programs scoring 5 or more points in this section decreased recidivism by 12.9 percent (N=8). The overall correlation between this categorical scale and outcome is .42.

Figure 2: Staff Characteristics Categories



Assessment

Table 3 represents the Assessment variables. The first item in this domain measures program exclusionary criteria. For this item, programs received a point for each of the following: 1) NOT excluding high risk offenders; 2) excluding offenders with both a history of arson and a history of serious or repeated violent offenses¹⁵. Meeting one of these characteristics

¹⁵ “Violent offenses” was defined differently by each agency according to their own exclusionary criteria. However, typically agencies that excluded violent offenders specified that these were offenders with 1) a long history of violent behavior, i.e., repeated assaults or serious violent misconducts, or 2) a serious violent offense such as murder or a recent offense that was particularly heinous. Few programs excluded offenders with *any* history of violent behavior.

was associated with less than a 1 percent reduction in recidivism (0.6%); meeting both of these characteristics was associated with a 7.0 percent reduction in recidivism. This item was significantly correlated with outcome ($r=.37$). An associated item in this section measured whether exclusionary criteria (whatever it might be) was routinely followed by the program. Programs that followed their exclusionary criteria performed better (3.6% recidivism reduction) than those that did not regularly follow their written criteria (0.4% reduction). Furthermore, programs that identified that overall participants were appropriate for the services being offered by the program experienced a slight improvement (3.4% recidivism reduction) over programs that felt more than 20 percent of the offenders served were inappropriate for the program (2.2 percent reduction). This item however, showed limited variation in that 58 of the 64 programs felt at least 80 percent of their clients were appropriate for the services being offered.

The next several items relate to risk and criminogenic need tools used by the program. Programs that assess those risk factors related to criminal recidivism have improved outcomes (3.5% reduction in recidivism) over programs that do not assess risk (2.5% reduction). While this difference is not significant and the correlation for this item is low ($r=.03$), note the lack of variation on this item as just 7 programs failed to systematically assess key risk factors. Related, programs that have incorporated additional specific need assessments, beyond a general risk/need tool (such as the Level of Service Inventory-Revised) also performed better. Programs that incorporated at least one standardized criminogenic need tool (e.g., substance abuse, anger management, or criminal thinking scale) experienced a 3.5 percent reduction in recidivism versus a 2.8 percent reduction for programs that did not. As with the previous item, few programs ($N=8$) failed to meet this standard.

Table 3: Assessment Items

Items		% Change in Recidivism	N	r (weighted)
Exclusionary Criteria Scale**				0.37
	1	0.6	34	
	2	7.0	30	
Exclusionary Criteria Followed				0.10
	No	0.4	7	
	Yes	3.6	57	
Clients Appropriate				0.02
	No	2.3	6	
	Yes	3.4	58	
Risk Assessed				0.03
	No	2.5	7	
	Yes	3.5	57	
Specific Need Assessments				0.03
	No	2.8	8	
	Yes	3.5	56	
Assessment Training Scale				0.08
	0 to 1	2.8	43	
	2	4.2	21	
Copies of Risk Assessment Provided				0.19
	No	0.3	11	
	N/A	3.4	38	
	Yes	5.5	15	
Responsivity Factors Assessed				0.13
	0 to 2 Characteristics	3.0	59	
	3 or More Characteristics	6.7	5	
Validation of Risk Assessment				0.41
	No	1.8	52	
	Yes	11.3	12	
Assessment Categories**				0.38
	0 to 4	-1.7	16	
	5 to 6	2.4	29	
	7 or more	7.9	19	

* significant difference at the .05 level ; ** significant difference at the .01 level

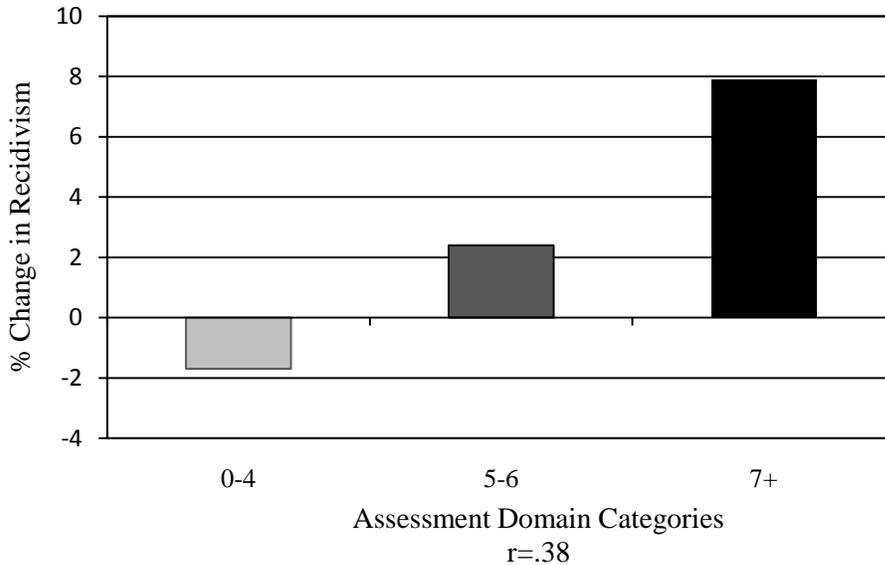
Where there was more variation is in how agencies trained staff on the risk/need assessment used by the program. This item suggests that it is important to provide staff with both formal training on the risk tool used by the program, as well as additional on-the-job training so that staff could see how the instrument was used and become comfortable using the tool. Programs offering both forms of training experienced an average reduction in recidivism of 4.2 percent, while programs offering *either* formal or on-the-job training *or* no training at all had a 2.8 percent average recidivism reduction. For programs that did not assess risk themselves, but relied upon probation/parole or some other entity to assess risk, there was also evidence supporting the importance of the agencies receiving full copies of the risk assessments (beyond just the risk classification or total score). Programs that used an outside assessor and obtained copies of each offender's risk/need tool showed an average reduction in recidivism of 5.5 percent versus just 0.3 percent for those that did not.

Programs that included an array of standardized responsivity assessments also showed improved outcome. For this item, programs were awarded a point for assessing the following responsivity areas: 1) personality; 2) anxiety; 3) mental health; and 4) learning styles. Programs assessing three or more of these areas using a standardized assessment tool experienced an average reduction of recidivism of 6.7 percent; programs that assessed two or fewer areas had a 3.0 percent reduction. Finally, programs that validated their risk tools on a similar population had a higher average rate of recidivism reduction (11.3%) over programs that did not validate/norm their instrument (1.8% reduction).

Figure 3 demonstrates the cumulative effects of meeting multiple items within the Assessment section. Programs that scored between 0 and 4 points in this domain *increased* the likelihood of recidivism by 1.7 percent (N=16). Programs scoring 5 to 6 points decreased

recidivism by 2.4 percent (N=29). Finally, programs scoring 7 or more points in this section decreased recidivism by 7.9 percent (N=19). These differences are significant and the overall correlation between this cumulative assessment scale and outcome is .38.

Figure 3: Assessment Categories



Treatment

Table 4 represents the Treatment variables. The first item in this domain measures a program's primary treatment targets. For this item, programs received a point for targeting each of the following characteristics: 1) criminal attitudes; 2) interpersonal relationship skills; 3) relationships with significant others (spouse/partner); 4) vocational skills; 5) self-control; and 6) criminal personality (e.g. lying, cheating, and stealing). This item was assessed by asking staff what needs they tended to spend time targeting as well as reviewing program group topics. Programs that targeted 1 or 2 of these needs experienced less than a 1 percent reduction in recidivism (.6%), while programs that targeted 3 to 4 of these characteristics saw an average reduction of 3.5 percent. Finally, programs that targeted 5 or 6 of these characteristics had a 10.4

percent reduction in recidivism. There was a significant difference in these recidivism rates and this item had a respectable correlation with outcome .30. This suggests the importance of programs targeting a range of criminogenic need areas versus focusing on one or two areas, such as employment or substance abuse.

The next four items measure the types of groups provided by the program, as well as the qualities of the groups. Positive attributes for groups were identified, consisting of factors such as percentage of time role play occurred, number of hours of programming per week, staff training on curriculum, and number of total sessions. These items were coded from program schedules, group observation as well as staff and resident interviews. Table 4 shows that programs offering a cognitive behavioral group such as Thinking for a Change were more effective at reducing recidivism, but only if the program had positive attributes. This item was coded as follows: 1) a point was deducted if a cognitive behavioral group was offered but it had no positive attributes; 2) no point was awarded if there was not a cognitive behavioral group offered; or 3) 1 point was awarded if a cognitive behavioral group was offered 4 or more hours per week OR the group allocated at least 50 percent of the time to role playing activities. Programs that scored -1 on this item had an average *increase* in recidivism of 1.4 percent; programs that scored 0 showed an average 4.8 percent reduction in recidivism; programs scoring 1 on this item had an average reduction in recidivism of 6.3 percent. This suggests improved outcomes by incorporating a cognitive-behavioral curriculum, but only when the curriculum delivered had positive qualities.

Programs that offer domestic violence intervention must also adhere to quality programming to achieve an adequate reduction in recidivism. For this intervention, positive attributes consisted of offering at least 10 sessions, incorporating some role play, and ensuring

Table 4: Treatment Items

Items	% Change in Recidivism	N	r (weighted)
Treatment Targets**			0.30
1 to 2	0.6	22	
3 to 4	3.5	37	
5 to 6	10.4	5	
Cognitive-behavioral Group Scale**			0.36
-1	-1.4	23	
0	4.8	20	
1	6.3	21	
Domestic Violence Group Scale			0.17
-1	-1.3	4	
0	3.9	55	
1	4.6	4	
2	6.9	1	
Gender-specific Group Scale*			0.48
0	2.8	25	
1	9.5	2	
2	14.2	3	
Dual Diagnosis Group			0.07
No	3.3	60	
Yes	6.3	4	
Punishment Process*			0.24
Follows 1 Guideline	-0.3	4	
Follows 2 to 4 Guidelines	2.5	48	
Follows 5 Guidelines	7.1	12	
Appropriate Types of Punishers			0.14
No	0.0	8	
Yes	3.8	56	

Table 4: Treatment Items Continued

Items	% Change in Recidivism	N	r (weighted)
Use of Graduated Practice			0.08
No	3.3	63	
Yes	6.9	1	
Group Size			0.19
Ratio greater than 10/1	2.3	45	
No More than 10/1	5.7	19	
Treatment Categories**			0.55
0 to 2	-1.9	24	
3 to 4	4.2	23	
5 or more	8.9	17	

* significant difference at the .05 level

** significant difference at the .01 level

that staff was trained on the curriculum. The domestic violence scale was coded as follows: 1) a point was deducted for programs that offer domestic violence group with none of the positive attributes mentioned for this item; 2) 0 points were awarded if no domestic violence group was offered; 3) 1 point was awarded when this group was offered with one or two positive attributes; and 3) 2 points were awarded when this group was offered with all three positive attributes. Table 4 shows that the bulk of programs (55) did not offer this group, thereby scoring 0 on this item. However, programs scoring -1 on this item had an average 1.3 percent *increase* in recidivism, while programs scoring 1 (4.6% reduction) or 2 (6.9% reduction) showed improved outcomes on average.

Of the programs that serve females, gender specific groups were assessed. These included groups offered only to women that were designed to address female-specific issues (e.g., Sista-to-Sista or Moving On). Positive attributes for this item included use of a curriculum, any role play, and at least 4 sessions. This item was coded as follows: 1) 0 points awarded if no

gender-specific group was offered; 2) 1 point was awarded for programs offering a gender-specific group; or 3) 2 points were awarded if a gender specific intervention was offered with at least one positive attribute. Data from Table 4 suggest that programs offering no gender specific interventions had an average reduction in recidivism of 2.8 percent. Programs with a gender specific group with no positive attributes reduced recidivism by 9.5 percent while programs with at least one positive attribute reduced recidivism by an average of 14.2 percent. While this finding is based upon just a handful of programs, the gender specific program scale was highly correlated with recidivism ($r=.48$), suggesting that offering gender responsive services increases program effectiveness, particularly when groups have the positive attributes listed above.

The final group-specific item concerns dual-diagnosis programming¹⁶. Note that there is limited variation for this item in that just four programs were coded as offering dual-diagnosis groups. However, of those providing this intervention, they averaged a 6.3 percent reduction in recidivism versus a 3.3 percent average reduction for programs that did not offer this group.

With regard to effective sanctioning within programs, Table 4 shows that programs that follow effective guidelines for issuing punishers are more effective. Effective guidelines identified by the data include: 1) punishment is based on exhibiting an antisocial behavior; 2) an explanation is provided for the punishment; 3) punishments are individualized (i.e. group punishment is avoided); 4) punishment is undesirable to the offender; and 5) there is written policy on the use of punishers/sanction. Programs following just one of the above guidelines had

¹⁶ Substance abuse groups were also examined, but found to have little impact at reducing recidivism, and were therefore not included in the treatment scale. However, consistent with other group data, the use of modeling/role play activities (at least 25% of the time) within substance abuse groups was correlated with improved effect sizes for programs. Yet, just 6 of the 64 programs regularly used behavioral techniques, which may be why this intervention had little overall effect on recidivism. The recommendation therefore is that substance abuse treatment (like other interventions) includes structured skill building and other behavioral practice techniques.

an average *increase* in recidivism of .2 percent. Programs consistently following 2 to 4 of the above guidelines reduced recidivism by an average of 2.5 percent, and programs following all five guidelines had a 7.1 percent reduction. These differences were significant with a correlation of .24. Not only was the process for punishment important, but so was whether the types of punishers were considered appropriate. Examples of inappropriate punishment types include use of shaming techniques or using treatment or treatment activities (i.e. thinking reports) as the punishment¹⁷. The majority of programs (N=56) used appropriate types of punishers and experienced an average reduction in recidivism of 3.8 percent. Programs using inappropriate punishers essentially showed no effect (-.01 percent increase).

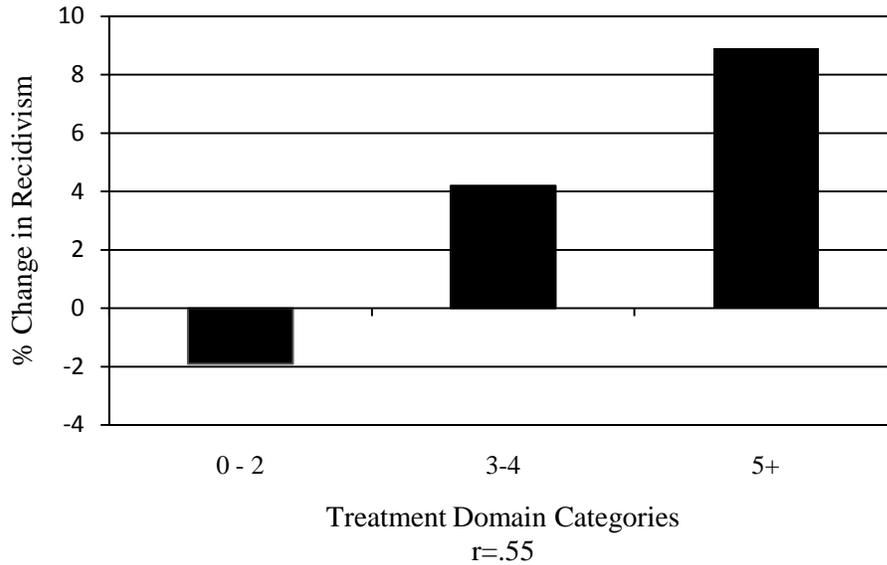
The treatment findings above related to group intervention demonstrated the importance of incorporating practice or role play as a key intervention technique. Unfortunately, just one program was coded as offering regular graduated practice of skills in increasingly difficult situations. The program that offered this intervention showed an average decrease in recidivism of 6.9 percent versus a 3.3 percent reduction for the remaining programs. Finally, programs that consistently offered treatment groups of an appropriate size (10 or fewer participants per facilitator) performed better than programs that exceeded this ratio. Specifically, programs with an appropriate group size had an average 5.7 percent reduction in recidivism (versus a 2.3 percent reduction).

Figure 4 demonstrates the cumulative effects of meeting multiple items within the Treatment section. Programs that scored between 0 and 2 points in this domain *increased* the likelihood of recidivism by 1.9 percent (N=24). Programs scoring 3 to 4 points decreased recidivism by 4.2 percent (N=23). Finally, programs scoring 5 or more points in this section

¹⁷ While programs are encouraged to use thinking reports or other treatment activities to help offenders correct thinking/antisocial behaviors, they should not be used as the punishment itself as treatment should not be framed as an aversive activity to offenders.

decreased recidivism by 8.9 percent (N=17). These differences are significant and the overall correlation between this cumulative assessment scale and outcome is .55.

Figure 4: Treatment Categories



Evaluation

Table 5 represents the Evaluation/Quality Assurance variables. Note that this table includes r-values for both weighted and unweighted variables to ascertain whether weighting the program by size affected the outcome. Weighted or not, this section produced no items that were significantly related to outcome. However, there were a few items that showed some substantive difference when quality assurance factors were adhered to. The first item in this domain measures the frequency with which groups are observed for the purpose of quality assurance. In programs where group observation took place less than three times a year, the average reduction in recidivism was 2.8 percent. Programs that observed group three or more times per year saw an average reduction of 3.8 percent. Note that when cases were weighted, the correlation for this item becomes negative.

The next item measures factors staff are evaluated on, typically in an annual performance evaluation. For this item, programs received a point for assessing each of the following characteristics: 1) receptive to supervision; 2) ability to handle crisis situations/de-escalate residents; 3) avoid negative interactions with residents; and 4) empathetic with residents. Programs that assessed staff on one or fewer of these factors saw an average reduction in recidivism of 2.3 percent. Programs that incorporated two or more of these four factors into a performance evaluation showed an average reduction in recidivism of 5.9 percent.

Another quality assurance measure that produced a slight positive program effect was related to file audits. Programs that conduct structured internal CCIS audits showed an average reduction in recidivism of 4.1 percent, versus programs that did not (3.2% reduction). Furthermore, there was some evidence that programs that relied upon external agencies to provide services to residents, and had quality assurance practices aimed at monitoring the quality of such services (i.e., reviewing progress notes and periodically observing the intervention), had improved outcome (6.0 % reduction in recidivism versus a 3.6% reduction). Note however, that when this item is weighed by program size, the correlation, while weak, becomes negative.

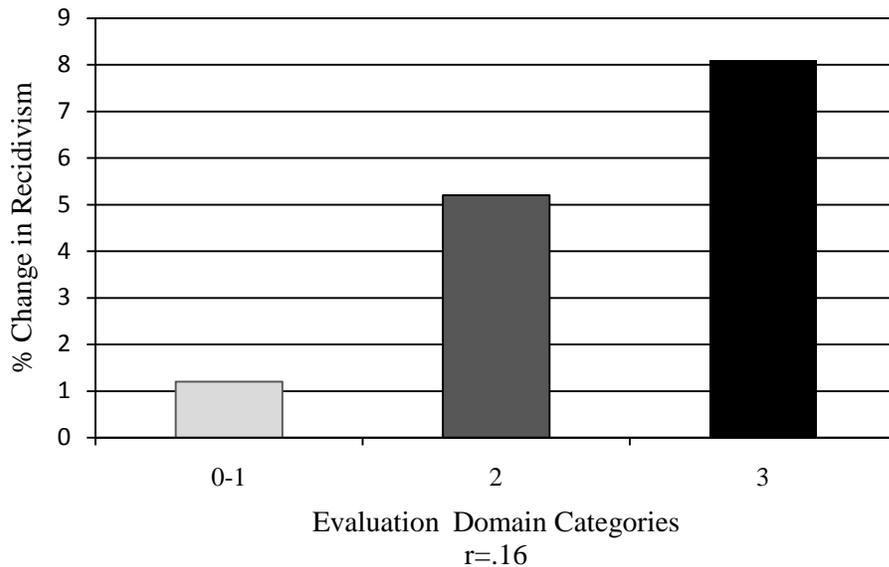
Figure 5 demonstrates the cumulative effects of meeting multiple items within the Evaluation/Quality Assurance section. Programs that scored between 0 and 1 points in this domain decreased the likelihood of recidivism by 1.2 percent (N=35). Programs scoring 2 points decreased recidivism by 5.2 percent (N=19). Finally, programs scoring 3 or more points in this section decreased recidivism by 8.1 percent (N=10). Differences for the unweighted scale only are significant with the overall correlation between this categorical Evaluation scale and outcome being .29 unweighted (and .16 weighted).

Table 5: Evaluation/Quality Assurance Items

Items	% Change in Recidivism	N	r (unweighted)	r (weighted)
Group Observation			0.05	-0.11
Less than 3X per year	2.8	20		
3 or more times per year	3.8	44		
Staff Evaluation Scale			0.19	0.19
0 to 1	2.3	43		
2 or more	5.9	21		
Internal CCIS Audits			0.05	0.07
No	3.2	43		
Yes	4.1	21		
External Quality Assurance			0.11	-0.03
No	3.6	55		
Yes	6.0	9		
Evaluation Categories			0.29*	0.16
0 to 1	1.2	35		
2	5.2	19		
3	8.1	10		

* significant difference at the .05 level; ** significant difference at the .01 level

Figure 5: Evaluation/Quality Assurance Categories



Overall Scale

Table 6 presents data on an overall cumulative scale for each of the five domains examined (Program Leadership/Development, Staff Characteristics, Assessment, Treatment, and Evaluation/Quality assurance). This table shows the distribution of programs within each category along with the associated correlation coefficient. The overall correlation with outcome was relatively high at .72. While this correlation was similar when only new felony convictions was used as the outcome measure ($r=.62$), the correlation did drop when new incarceration was used at the outcome ($r=.33$)¹⁸. This decrease in r value is likely due to the confounding nature of using new incarceration as an outcome. As mentioned in the methods section, the current measure of incarceration includes incarceration for technical violations. This is problematic because offenders in treatment programs are likely to have more conditions of supervision and closer monitoring of adherence to those conditions. As a result, offenders in treatment programs are at increased risk of being incarcerated, not because the treatment is ineffective, but because they are more likely to get into trouble for technical violations. Support for this notion is shown when new incarcerations are restricted to only those offenders who received a felony conviction. When using this outcome, the correlation between the total scale and new incarceration again becomes strong ($r = .59$)

Figure 6 offers a graphical description of the overall score categories. Programs with a total score ranging between 0 and 13 had an average *increase* in recidivism of 2.9 percent (N=22). Programs with cumulative scores ranging between 14 and 20 reduced recidivism by a rate of 5.8 percent (N=36). Finally, programs scoring 21 or more total points reduced recidivism by an average rate of 17.8 percent over comparison samples (N=6). The progressive increase in

¹⁸ This includes incarceration for any reason (new crime or technical violation).

effect sizes as the overall score categories increase suggests the overall score is related to a reduction in recidivism.

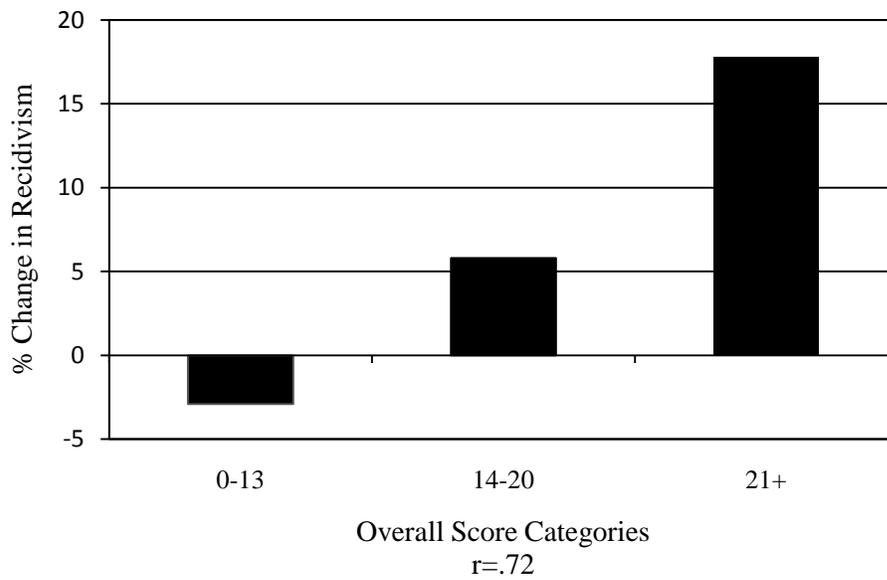
Table 6: Overall Cumulative Score

Items	% Change in Recidivism	N	r (weighted)
Overall Score Categories**			0.72
0 to 13	-2.9	22	
14 to 20	5.8	36	
21 or more	17.8	6	

* significant difference at the .05 level

** significant difference at the .01 level

Figure 6: Overall Cumulative Score Categories



SECTION IV: DISCUSSION

This report served as a supplement to the 2010 Ohio HWH/CBCF study designed to assess the effectiveness of 44 HWH and 20 CBCF programs. This research sought to evaluate what program characteristics within these 64 programs were most important in reducing offender recidivism. To answer this research question, program characteristics within five sub-domains were evaluated: *Program Leadership/Development*; *Staff Characteristics*; *Assessment*; *Treatment*; and *Evaluation/Quality Assurance*. Scales were created using individual program characteristics that had a significant and/or substantive relationship with program effectiveness. The correlations of the sub-domain scales ranged from .16 for *Evaluation/Quality Assurance* to .55 for *Treatment*. When all scales were combined, the correlation between the cumulative categorical score and outcome was respectable ($r=.72$).

Within the *Program Leadership/Development* domain, regular supervision by the program director ($r=.29$) and operating a co-ed facility with separate living spaces ($r=.17$) were individual items that emerged as significant in reducing program recidivism. Harbor Light, Pathfinder and CATS were Ohio programs that met several of the indicators in the Program Leadership/Development domain, indicating that these are well-established and stable programs with an involved and qualified leader. Within the *Staff Characteristics* domain, all six individual factors had a correlation coefficient ranging between .20 and .28. Of particular importance was holding multi-disciplinary clinical meetings that included periodic training for staff. Likewise, clinical supervision by a certified clinician was important, as was seeking staff for programs that have a range of skills and values. Several of the Oriana House sites scored well in this section, suggesting it is an organization that does a good job of selecting appropriate staff for its sites, as well as training and supervising the staff.

Both the *Assessment* and *Treatment* domain showed evidence in support of the principles of effective intervention. In terms of the *Assessment* domain, support for the risk principle was shown as programs that did not exclude high risk offenders showed better effects. Validating the risk/need assessment used by the program was also important. Within the *Treatment* domain, support of the need principle was clear as programs that targeted a range of criminogenic needs (5-6) reduced recidivism by over 10 percent versus those that targeted just 1-2 criminogenic needs (.6% recidivism reduction). Likewise, support was shown for the general responsivity principle, as well as the importance of program fidelity by examining group data. Improved effect sizes were evident when cognitive-behavioral, gender-specific and domestic violence programs had “positive attributes”. Positive attributes consisted of strategies such as role playing, having an ample dosage of treatment, and training staff on the curriculum. Use of role play, to varying degrees, increased effects in all three of these group interventions. Cognitive-behavioral groups showed maximum effects when at least 50% of group time was allocated to role play. Interestingly, there was support for use of gender-specific interventions, but again, particularly when such interventions had cognitive-behavioral elements, such as role playing.

The Community Corrections Association (CCA) programs scored well with regard to assessment. They met several indicators on this scale, implying that offenders are appropriately selected for their programs, and are assessed using a range of tools. In the treatment domain, Lorain-Medina, NEOCAP, Oriana Crossweah and Cliff Skeen scored particularly well among the CBCFs; Harbor Light again scored well as a HWH program. This suggests that the treatment targets at these facilities were likely appropriate, as well as the primary treatment model and the degree to which the programming incorporated positive program attributes.

The *Evaluation/Quality Assurance* domain showed limited ability in predicting effective outcomes. One reason for this might be the homogeneity of several of the items assessed in this domain. For example, all but a handful of the programs assessed had undergone a previous outcome evaluation (the original 2002 Ohio HWH/CBCF study). Likewise, most of the programs were accredited by the American Correctional Association (ACA) and overseen by ODRC, and therefore had to meet certain audit standards relative to the evaluation domain, such as regular file review. Nonetheless, CCA as well as some of the Oriana and Alvis House sites scored well in the Evaluation domain, suggesting they have worked to implement effective quality assurance practices.

Also of interest are factors one might have expected to see as predictors of reduced rates of recidivism, but did not. Items related to effective reinforcement, having quality aftercare, providing family intervention, and matching offenders based upon key responsivity issues did not emerge as important predictors of program effectiveness. Like with the evaluation domain items, many of these (and related) factors lacked variation. In other words, when few (or nearly all) programs are coded as having a particular program attribute, it becomes difficult to predict outcome. Of the items listed above, few programs had quality aftercare, matched based upon responsivity factors, or provided structured family intervention. On the other hand, most programs had reinforcers in place, although the process used for reinforcing offenders did vary across programs.

Hence, one limitation of the study is that many items coded did not emerge as significant predictors of reduced recidivism, not because the factors are not important for programs to incorporate, but because the data for that item lacked variation. Additionally, analyses were ultimately based upon a sample size of 64 programs. Adding additional programs, as well as

broadening the types and locations of the programs (as all facilities were residential programs serving adult offenders in Ohio) would improve the study's generalizability. Finally, program characteristics were coded by several trained UC researchers. Although steps were taken to ensure that program data was accurately and reliably collected and coded, a certain amount of subjectivity is inevitable when conducting process evaluations.

Despite these limitations, recommendations can be derived from the findings that should assist programs to enhance treatment services. In general, findings support the principles of effective intervention, which Ohio programs have been encouraged to follow for some time now. Other more specific findings also emerged that will help administrators design effective programs. Examples include program directors conducting at least 5 hours of structured supervision per month, or adapting staff and unit schedules so that security staff can participate in treatment teams. When treatment curricula are selected, administrators should ensure that an appropriate percentage of time is spent on practicing skills. Furthermore, programs should be designed so that at least 5 criminogenic needs are being targeted. This study's findings, along with the outcome study results which strongly supports application of the risk principle, should be taken under consideration when designing programs or making program adaptations. Of equal importance, however, is the process by which programs make changes. Programs should be thoughtful about what areas to address, and not attempt to address all deficiencies at once. Instead programs should prioritize need areas and then plan for and pilot an intervention designed to address the identified need. This helps to ensure that program changes are appropriate, that program staff and managers do not get overwhelmed, and that changes can be monitored and adjusted as needed. Attempting to implement effective practices and doing so in a manner that promotes sustained improvements will likely lead to increased program effects.

REFERENCES

- Andrews, D. and Bonta, J. (2007). *The Psychology of Criminal Conduct (4th ed.)*. Cincinnati, OH: Anderson Publishing Co.
- Andrews, D., I. Zinger, R. Hoge, J. Bonta, P. Gendreau, and F. Cullen (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology* 28(3): 369-404.
- Andrews, D. and Dowden, C. (1999). What Works in Young Offender Treatment: A Meta Analysis. *Forum on Correctional Research*, 11(2).
- Antonowicz, D. and Ross, R. (1994). Essential components of successful rehabilitation programs for offenders. *International Journal of Offender Therapy and Comparative Criminology* 38: 97-104.
- Barnoski, R. (2004). *Outcome evaluation of Washington State's research based programs for juvenile offenders*. Olympia Wash: Washington State Institute for Public Policy.
- Baro, A. (1999). Effects of a cognitive restructuring program on inmate institutional behavior. *Criminal Justice and Behavior*, 26(4): 466-484.
- Dowden, C. and Andrews, D. (1999b). What works in young offender treatment: A meta-analysis. *Forum on Correctional Research*, 11: 21-24.
- Dowden, C. and Andrews, D. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology*, 449-467.
- Dowden, C. and Andrews, D. (2004). The importance of staff practice in delivering effective correctional treatment: a meta-analytic review of core correctional practice. *International Journal of Offender Therapy and Comparative Criminology*, 48: 203-14.
- Dowden, C. and Andrews, D. (2006). The importance of staff practice in delivering effective correctional treatment: A meta-analytic review of core correctional practice. *International Journal of Offender Therapy and Comparative Criminology*, 48: 203-14.
- French, S. and Gendreau, P. (2006). Reducing Prison Misconducts. *Criminal Justice and Behavior*, 33(2), 185-218.
- Gendreau, P. (1996). "The Principle of Effective Intervention with Offenders." In A.T. Harland (ed.), *Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply*. Thousand Oaks, CA: Sage Publications.

Gendreau, P. and Andrews, D. (1989). The Correctional Program Assessment Inventory.

Gendreau, P., French, S., and Gionet, A. (2004). What works (what doesn't work): The principles of effective correctional treatment." *Journal of Community Corrections*, 13: 4-30.

Gendreau, P., French, S., and Taylor, A. (2002). What Works (What Doesn't Work) Revised 2002. *Invited Submission to the International Community Corrections Association Monograph Series Project*.

Gendreau, P. and Ross, R. (1987). "Revivification of Rehabilitation: Evidence from the 1980s." *Justice Quarterly*, 4:349-407.

Hanson, C., Gordon, A. Harris, A, Marques, J., Murphy, W. Qunisey, V. and Seto, M. (2002). First Report of the Collaborative Outcome Data Project on the Effectiveness of Psychological Treatment for Sex Offenders. *Sexual Abuse: A Journal of Research and Treatment*, 14(2): 169-194.

Henning, K. and Frueh, C. (1996). Cognitive behavioral treatment of incarcerated offenders: An evaluation of the Vermont Department of Corrections' Cognitive Self-Change program. *Criminal Justice and Behavior*, 23(4): 532-541.

Landenberger, N. and Lipsey, M. (2005). The positive effects of cognitive-behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1(4): 451-476.

Latessa, E., Brusman Lovins, L., and Smith, P. (2010). *Follow-up Evaluation of Ohio's Community Based Correctional Facility and Halfway House Programs—Outcome Study*. Technical Report. Center for Criminal Justice Research, University of Cincinnati, Cincinnati, OH.

Latessa, E. and Lowenkamp, C. (2002). *Evaluation of Ohio's Community Based Correctional Facilities and Halfway House Programs*. Technical Report. Center for Criminal Justice Research, University of Cincinnati, Cincinnati, OH.

Lipsey, M (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. Pp. 83-127 in T. Cook, H. Cooper, D. Cordray, H. Hartmann, L. Hedges, R. Light, T. Louis, and F. Mosteller (eds.), *Meta-Analysis for Explanation*. New York: Russell Sage Foundation.

Lipsey, M., Chapman, G., and Landenberger, N. (2001). Cognitive behavioral programs for offenders. *The Annals for the American Academy of Political and Social Science*, 578: 144-157.

Losel, F. (1995). "Increasing Consensus in the Evaluation of Offender Rehabilitation? Lessons from Recent Research Synthesis." *Psychology, Crime, & Law* 2: 19-39.

- Lowenkamp, C. T. (2004). *Correctional Program Integrity and Treatment Effectiveness: A Multi-site, Program-Level Analysis*. Doctoral Dissertation. University of Cincinnati.
- Lowenkamp, C. T., Flores, A. W., Holsinger, A. M., Makarios, M. D. & Latessa, E. J. (2010). Intensive Supervision Programs: Does Program Philosophy and the Principles of Effective Intervention Matter? *Journal of Criminal Justice*. Doi:10.1016/j.jcrimjus.2010.04.004
- Lowenkamp, C.T., Hubbard, D., Makarios, M. and Latessa, E. (2009). A Quasi Experimental Evaluation of Thinking for a Change a "Real World" Application. *Criminal Justice and Behavior*, 36(2): 137-146.
- Lowenkamp, C.T., Latessa, E., & Holsinger, A. (2006). The risk principle in action: What we have learned from 13,676 offenders and 97 correctional programs. *Crime and Delinquency*, 52(1): 1-17.
- Lowenkamp, C.T., Latessa, E.J., & Lemke, R.J. (2006). *Evaluation of Ohio's RECLAIM Funded Programs, Community Corrections Facility, and DYS Facilities: FY 2002*. Technical Report. Center for Criminal Justice Research, University of Cincinnati, Cincinnati, OH.
- Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does correctional program quality really matter? The impact of adhering to the principles of effective intervention. *Criminology & Public Policy*, 5: 201-220.
- Lowenkamp, C.T., Makarios, Latessa, Lemke & Smith (2010). Community Corrections Facilities for Juvenile Offenders in Ohio: An Examination of Treatment Integrity and Recidivism. *Criminal Justice and Behavior*, 37(6): 695-708.
- Pearson, F., Lipton, D., Cleland, C. and Yee, D. (2002). The effects of behavioral/cognitive-behavioral programs on recidivism. *Crime and Delinquency*, 48 (3): 476-496.
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Newbury Park, CA: Sage.
- Wilson, D., Bouffard, L.A. and MacKenzie, D. (2005). A quantitative review of structured, group-oriented, cognitive-behavioral programs for offenders. *Criminal Justice and Behavior*, 32(2): 172-204.
- Wolf, F. M. (1986). *Meta-Analysis: Quantitative Methods for Research Synthesis*. Newbury Park, CA: Sage.

APPENDIX A

Table 7: Program Characteristic Domains and Overall Score based upon *All Program Participants*

Domain		% Change in Recidivism	N	r (any conviction)
Leadership**				0.40
	0 to 1	-7.2	15	
	2 to 3	-0.6	37	
	4 to 5	3.4	12	
Staff Characteristics*				0.31
	0 to 1	-4.3	7	
	2 to 4	-1.8	49	
	5 or more	5.2	8	
Assessment**				0.37
	0 to 4	-5.4	16	
	5 to 6	-2.3	29	
	7 or more	2.9	19	
Treatment**				0.49
	0 to 2	-5.9	24	
	3 to 4	-0.1	23	
	5 or more	3.2	17	
Evaluation				0.09
	0 to 1	-2.2	35	
	2	0.3	19	
	3	-1.3	10	
Overall Score**				0.58
	0 to 13	-6.6	22	
	14 to 20	0.5	36	
	21 or more	8.5	6	

* significant difference at the .05 level
 ** significant difference at the .01 level
 correlations (r values) based upon weighted cases

Data in the body of the report examines the correlations between domain scores and outcome (any conviction) for successful treatment completers only. Table 7 examines the percent change in recidivism and corresponding r value for both successful and unsuccessful treatment completers. Correlations for the entire sample were either similar or lower as compared to the correlations when only successful completers were examined. For the leadership domain, the correlation between outcome and leadership

items for successful completers was .41 versus .40 for all participants. For staff characteristics, a correlation of .56 was reported; this dropped to .31 for all participants. For assessment, the r-value using the successful completer only sample was .36 compared to .37 for all participants. In the treatment domain, the correlation between outcome and treatment items dropped ($r=.60$ versus .49). The evaluation domain dropped just slightly ($r=.12$ versus $r=.09$). Finally, the correlation between the overall score was .72 for successful completers, and dropped to .58 when all participants were used as the sample.

Hence, the correlations in two of the domains as well as the overall score showed a drop when both completers and non-completers were examined compared to successful completers only. However, this is not surprising given that this report examined what program characteristics were associated with stronger treatment effects. When including the entire population of offenders *exposed* to an intervention (versus successfully completing the intervention), the full effects of the intervention cannot be examined. Aside from the evaluation domain, using all participants as the sample still produced modest correlations with recidivism (above .30 in each domain, and .58 for the overall score). Furthermore, in all but the evaluation domain (which generally had limited predictive ability), scales could still be developed that showed reductions in recidivism with higher domain scores when non-completers were added to the sample.

Table 8: Correlation Coefficients Using Multiple Outcome Measures for *All Participants*

	r (felony conviction)	r (felony incarceration)	r (any incarceration)
Leadership	0.36**	0.31**	0.09**
Staff Characteristics	0.22*	-0.23	0.03
Assessment	0.31**	0.21**	-0.16
Treatment	0.41**	0.36**	-0.09
Evaluation	0.14	0.21	-0.08
Overall Score	0.47**	-0.22	0.33**

* significant difference at the .05 level
 ** significant difference at the .01 level
 correlations (r values) based upon weighted cases

The outcome variable used to conduct the analyses in the body of the report was *any conviction*. Table 8 presents correlations on the sample of both completers and non-completers using alternative measures of recidivism (felony conviction, felony incarceration, and any incarceration--technical or new crime). When other outcome variables are examined, not surprisingly, the *any conviction* and *felony conviction* correlations look fairly similar. Likewise, in general the incarceration outcome variables were more weakly correlated with the program characteristic domains, particularly when both technical and new crimes are used as the outcome. This finding is not entirely unexpected given that many of the non-completers included in the sample may have been returned or sent to prison on a technical violation related to their failure to complete the treatment program.