An Empirical Examination of Variation in Effective Correctional Program Characteristics by Gender

A Dissertation Submitted to the:

Division of Research and Advanced Studies
Of the University of Cincinnati

In Partial Fulfillment of the
Requirements for the Degree of

DOCTORATE OF PHILOSOPHY (Ph.D.)

In the School of Criminal Justice
Of the College of Education, Criminal Justice, and Human Services

August, 2012

By

Lori Brusman Lovins

B.A., University of Cincinnati, 1993
M.S.W., University of Cincinnati, 1996

Dissertation Committee: Edward J. Latessa, Ph.D. (Chair)
Paula Smith, Ph.D.
Patricia Van Voorhis, Ph.D.
Christopher T. Lowenkamp, Ph.D.

ABSTRACT

Males strongly dominate those served in corrections. As such, correctional interventions have largely been designed without attention to gender differences. Yet, with growing numbers of females served in the correctional system, gender-responsive programming has increasingly drawn the interest of researchers and practitioners. Feminist scholars argue for gender-specific programming that focuses on the needs of women, such as economic hardship, healthy relationships, mental illness, addiction and victimization. The body of literature that supports this perspective is primarily qualitative. Other researchers have empirically examined how core principles tied to recidivism reduction, such as the use of risk assessment, targeting of criminogenic needs and use of a cognitive-behavioral treatment model applies to female offenders. Based on both theoretical and empirical grounds, these researchers argue that core interventions should be employed regardless of an offender's gender. The current study empirically explores these contrasting perspectives by examining 138 programs serving male and female populations to determine how the predictors of program effectiveness vary by gender. Findings outline both similarities and differences in effective program traits for male versus female programs. Results can be used to help develop gender responsive programming for both women and men, but also suggests the need for continued empirical examination of variation in effective programming based on gender.

ACKNOWLEDGEMENTS

The opportunity I have been given to undergo the process of completing a doctoral program would not have been possible without the support of many people. I would like to begin by thanking Dr. Latessa. He has not only mentored my academic and professional growth, he has laid the foundation to my career by providing me with endless opportunities. His genuine drive toward improving the field of corrections was an inspiration and gave purpose to the long work hours. More important than professional development, he took my family under his wing and always made sure we were taken care of—I am eternally grateful for this. I would also like to thank Dr. Smith. She is one of the smartest people I know and has been a great source of knowledge, but more importantly she's encouraged me to follow my dreams.

I must also acknowledge those dear friends who have offered continuous reassurance through this process. This includes my UCCI colleagues, as well as my friends inside and outside the program who never stopped believing I would *one day* finish.

I also owe a great debt to my family, especially my parents, Gary and Evelyn Brusman and my husband's parents, Margie and Chuck Lovins. I truly would not have been able to complete this program and pursue my career goals without their support—both emotional and practical. They not only offered unending encouragement, they spent endless hours caring for our babies while we worked and completed the requirements of our degree. They set our mind and heart at ease knowing that the time we spent away from our children was with people who love them as much as we do. They supported me in nurturing both my professional and personal life. I would definitely not be where I am if not for their support.

Finally, I want to thank my immediate family. Sam, Henry and Walt have had to endure much time away from their "ma" so that I could complete my goal of a Ph.D. and follow my

passion of working with correctional programs. They have been my source of laughter, inspiration and strength, and they are my greatest accomplishments. Finally, I want to thank my husband. It has been a long 10 years with many ups and downs—there were some really tough downs that we made it through, and more ups than we probably deserve. I have the utmost respect for you as a person and a professional. You've taught me that less worry and more fun makes for a happier life. I am grateful we took this path together and am so thankful for your friendship and love along the way.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

CHAPTER 1
STATEMENT OF THE PROBLEM

Introduction

A growing number of offenders are incarcerated each year. At yearend 2009, 1.6 million prisoners were held in federal or state prisons in the United States; this is a slight increase from the year prior (BJS, 2010). Rising incarceration rates mean that increasing numbers of inmates are reentering the community. In fact, in the United States in 2009, one in every 32 adults was under some form of correctional supervision (BJS, 2010). Furthermore, among state parole discharges in 2006, only 44 percent of offenders successfully completed their supervision term (BJS, 2007). This suggests the need for increased attention on successful reentry from prison to the community.

Along with the release of more prisoners to the community, the criminal justice system is also experiencing an influx of female offenders. Between 1990 and 2009, the number of women in corrections has grown by nearly 50 percent, while the male corrections population has grown at only half this rate (BJS, 2010). In 1990, men accounted for 86 percent of the correctional population; this decreased to 82 percent in 2009 (BJS, 2010). Between 1995 and 2005 the number of women in prison increased an average of 4.6 percent per year, compared to less than a three percent increase for men (Hartney, 2007). Yet, women are still more likely to be placed on community supervision then men. At yearend 2009, about two-thirds of all men in corrections were under community supervision compared to 85 percent of females (BJS, 2010). Hence, while men still make up the preponderance of offenders served in the correctional system, the increasing rate of female offenders, particularly in the community corrections system, demands attention.

The ever increasing rate of both men and women involved in the correctional system indicates a clear need for more effective correctional treatment strategies. Much research has been conducted on effective interventions with offenders (see MacKenzie, 2000 for a review). However, less empirical research has focused on gender responsive strategies to treating offenders. This dissertation is designed to add to the growing body of literature on effective correctional interventions, but with a specific focus on whether effective treatment strategies vary by gender. That is, data will be examined to determine what treatment characteristics are most important for decreasing recidivism for males versus females in community correctional facilities, as well as identifying those treatment characteristics that are important, despite an offender's gender. The hope is that findings from this dissertation will aid in designing empirically-driven gender responsive programming.

Community Based Correctional Interventions

Research suggests that community-based interventions are oftentimes more effective at rehabilitating offenders than incarceration programs (Andrews, Zinger, Hoge, Bonta, Gendreau and Cullen, 1990; Lipsey and Wilson, 1998, Gendreau, French and Taylor, 2002). One proposed reason for the enhanced effectiveness is that offenders are being treated in the environment in which they live (in vivo); therefore, the prosocial skills they learn in treatment can more easily transfer and be maintained in the offender's life (Gordon, Arbuthnot, Gustafson and McGreen, 1988, Henggeler, Melton, Brondino, Scherer and Hanley, 1997). Others have argued the inhumane nature of prisons prohibit rehabilitative efforts (Toch, Adams and Grant, 1989). Thus, when considering the increased effectiveness and decreased cost of most community-based strategies relative to incarceration, there is an amplified focus on community-based interventions in addressing criminal behavior.

Residential Community Correctional Programs

One example of a community-based intervention used across the nation is the residential community correctional program, commonly known as the half-way house. Half-way houses first began in England and Ireland in the early 17th Century as transitional programming for inmates (Latessa and Allen, 1982). While a handful of halfway houses emerged in the U.S. in the mid-1800s, the need for this correctional strategy was not fully realized until the mid-twentieth century. The development of the International Halfway House Association in 1964 as a professional organization for halfway house administrators and personnel (now called the International Community Corrections Association) was pivotal in the growth of these residential community correctional programs (Wilson, 1985).

Halfway houses were primarily designed to provide a step-down for offenders transitioning from prison to the community (Latessa, Travis and Lowenkamp, 2005). Yet, residential community correctional programs are also used as an alternative to imprisonment (Latessa and Travis, 1991). This provides a more secure setting for offenders not appropriate for probation, without subjecting them to the negative effects of imprisonment. Whether used for the purpose of transition or as a last resort before prison, such places have the potential to assist offenders in obtaining employment and stable housing (Latessa and Travis, 1991). Well-developed residential community correctional facilities also offer programming to address the needs that impact an offender's ability to refrain from criminal behaviors such as substance abuse, criminal thinking, lack of education and employability (Lowenkamp and Latessa, 2004). There is, however, wide variation in the strategies residential programs use to assist offenders to reintegrate with the community. While some provide only temporary housing for offenders,

others use multiple approaches to aid individuals with community reentry (Latessa and Travis, 1991).

The effectiveness of residential community correctional programs at reducing recidivism varies widely (Latessa and Travis, 1991, Lowenkamp and Latessa, 2004, Lowenkamp and Latessa 2005, and Lowenkamp, Smith, and Latessa, 2007). Latessa (1998) noted several common shortcomings of halfway house programs, including inadequate assessment, low qualifications and high turnover among staff, and lack of theoretically based treatment models. While many halfway house programs are plagued with these deficiencies, others are not and have been effective at reducing recidivism. Lowenkamp and Latessa (2002) found in a study of residential community correctional facilities that halfway houses could produce sizeable effect sizes depending on who they served, what they targeted, and how such needs were addressed.

Research on Effective Correctional Treatment

There is currently a plethora of research related to effective correctional interventions. As suggested above, there is great variability in the ability of correctional programs to reduce recidivism (Gendreau and Ross, 1987; Andrews et al., 1990; Lipsey, 1992; Gendreau 1996; Latessa, Cullen and Gendreau, 2002; Pearson, Lipton, Cleland and Yee, 2002). Research suggests that effective treatment programs are associated with particular program traits (Andrews et al., 1990, Gendreau, 1996). The "principles of effective intervention" define such traits. These principles provide researchers and practitioners with a template for creating effective correctional treatment programs.

The principles of effective intervention can be most succinctly defined as the risk, need and responsivity principles (Andrews and Bonta, 2010). The risk principle asserts that higher levels of supervision and treatment should be applied to those offenders most likely to recidivate.

The need principle asserts that client needs associated with criminal behavior (i.e. criminogenic needs), should be the primary focus of treatment efforts. Finally, the responsivity principle guides practitioners as to how interventions should be applied. It consists of both general and specific responsivity. General responsivity suggests that most offenders respond to behavioral interventions, thus correctional treatment programs should adhere to cognitive-behavioral models. The specific responsivity principle urges that treatment should be individualized to meet the unique needs and learning styles of the offenders (Andrews et al, 1990).

There is however, debate in the criminal justice field as to the extent that these principles apply to females. Since males make up the vast majority of the correctional population, most quantitative data represents male offenders. Feminists in the field are particularly concerned about the application of these "male driven" principles to females. Feminists argue that women have different pathways to criminal behavior; as such their needs differ from that of men (Daly, 1994; Bloom, Owen and Covington, 2003; Reisig, Holtfreter and Morash, 2006). Given this, they reject that the same treatment strategies used to treat males are equally effective for women and girls.

To the contrary, there are researchers that have investigated the empirical relationship between the principles of effective intervention and recidivism using samples of female offenders. They argue that there are few gender differences in the correlates of criminal behavior and that these principles apply to females just as they do males (Simourd and Andrews, 1994, Dowden and Andrews, 1999a, Lowenkamp, Latessa and Smith, 2007). Furthermore, the theoretical foundation that forms the basis of the general responsivity principle, namely, behavioral and social learning theories, is gender neutral. This suggests there should be no distinction between men and women regarding the application of cognitive-behavioral treatment

strategies. Hence, on both empirical and theoretical grounds, these researchers argue the principles of effective intervention pertain to both males and females.

Until recently, little data on gender-responsive treatment strategies for offenders has been made available. While meta-analyses have helped to determine the applicability of the principles of effective treatment to female populations, this study methodology has limitations. Meta-analyses are restricted in their ability to measure these principles, as the inner workings of programs can oftentimes not be ascertained from what is published in the original studies. This dissertation will therefore serve as an important contribution to this limited body of research as a comprehensive assessment of program characteristics was used to identify what treatment characteristics are particularly important for men versus women.

Research Questions

The purpose of this dissertation is to identify both unique and shared indicators of effective treatment for male and female offenders. Specifically, this dissertation seeks to explore the following research questions:

1. What program characteristics are important for *both* adult male and female offenders?
2. What program characteristics are *more* important for adult male offenders?
3. What program characteristics are *only* important for adult male offenders?
4. What program characteristics are *more* important for adult female offenders?
5. What program characteristics are *only* important for adult female offenders?
6. What program characteristics are important for *neither* adult male nor female offenders?

To address these questions, this study will explore key characteristics of 125 programs, 75 of which are exclusively male and 50 of which include female offenders. The program level data collected at each of these facilities is detailed enough to empirically test each of the

principles of effective intervention. The program characteristics associated with a reduction in recidivism over respective control groups will be compared to identify important factors for both male and female programming. This study is important as it can be used to assist correctional programs to develop gender responsive strategies, using empirical findings to guide decisions about the implementation of effective programming. Assisting correctional programs to develop gender-responsive programs that are effective at reducing offender recidivism will lead to better allocation of limited resources as well as enhanced public safety.

Summary

Given that an ever-increasing number of both men and women are being incarcerated, examining effective community-based interventions for reentry purposes is prudent. This chapter has briefly reviewed data finding that female offenders are among the fastest growing population in corrections, suggesting the need to explore effective correctional interventions specific to women as well as men. A brief history on the half-way house as a community-based strategy was presented, along with how employing effective correctional principles in half-way house interventions is imperative. This dissertation will examine application of the principles of effective programming to both men and women, using data collected on 125 residential community correctional programs. The next chapter will review literature that considers whether such principles are equally applicable to male and female offenders. Specifically, Chapter Two will explore the literature on effective correctional strategies, outlining the principles of effective intervention and exploring the two primary schools of thought regarding how to effectively treat and manage female offenders.

CHAPTER 2
WHAT WORKS?

The Nothing Works Debate

There is a large body of literature on "what works" in the rehabilitation of offenders. Much of this literature grew from the pivotal 1974 Martinson article wherein he declared that "nothing works" to rehabilitate offenders. This conclusion stemmed from his review of 231 evaluation studies of treatment programs conducted between 1945 and 1967. Based on this review he concluded that while he may have found "isolated" instances at success or partial success, there is little hope of a sure way to reduce recidivism through rehabilitation.

This study, however, was followed by a series of rebuttals, wherein other researchers who conducted narrative reviews of the literature found evidence of effectiveness of rehabilitation (Palmer, 1975; Gendreau and Ross, 1979, Gendreau and Ross, 1987; Cullen and Gendreau, 1989). In one of the first critiques of the Martinson study, Palmer (1975) found that Martinson and his colleagues overlooked many positive instances of success. He suggested the better question was "which methods work best for which type of offender under what conditions" (p. 150). Gendreau and Ross (1979) further argued that many treatment programs failed not because treatment is ineffective for offenders, but because treatment lacked therapeutic integrity. After conducting lengthy narrative reviews, both Gendreau and Ross (1987) and Cullen and Gendreau (1989) began to outline some of the program characteristics that appeared to be associated with reductions in recidivism[1]. These narrative reviews were followed by a more parsimonious way

[1] These characteristics will be reviewed in the next section of this dissertation.

to summarize the literature and determine the most effective strategies for changing offender behavior.

Meta-Analysis

The early studies on effectiveness of correctional treatment were conducted using a narrative review of the literature or a ballot box approach[2]. By the late 1970s, a more sophisticated statistical technique emerged that provided a quantitative rather than qualitative review of the correctional treatment literature (Smith and Glass, 1977). This technique, known as meta-analysis, began to replace the traditional ways of summarizing a body of literature (Cullen and Gendreau, 1999). The meta-analysis involves collecting relevant studies, coding the studies based upon a pre-determined set of variables, creating summary statistics for each study and analyzing the aggregated data quantitatively (Lipsey and Wilson, 2001, Izzo and Ross, 1990).

Despite the utility of this technique in summarizing a body of literature, there are several limitations to the meta-analytic approach that should be noted. The first is the selection bias/file drawer problem. This problem suggests there may be bias in the selection of studies to be included in the analysis, particularly since studies with significant findings are more likely to be published than studies with null findings (Rosenthal, 1991). Hence, conducting a meta-analysis that includes only the more easily accessible published studies can skew the meta-analytic findings to favor treatment effectiveness. Consequently, meta-analysts must attempt to incorporate both published and unpublished studies in the analyses.

A second popular argument against the use of meta-analysis is the "apples and oranges" argument. Since meta-analyses summarize a body of literature wherein factors such as sampling

[2] In the correctional evaluation literature, the ballot box approach involves collecting studies of interest and tallying the effective versus ineffective evaluation findings (Cullen, 2002)

units, independent variables and dependent variables vary, some argue that calculating average effect sizes is misleading. Lipsey and Wilson (2001) however argue that increasingly, the primary question of interest involves finding the sources of differences in study findings rather than aggregating to a grand average. Hence, subcategories of study findings can be analyzed separately, creating a distribution of effect sizes and related statistics to be compared.

Third, it is argued that meta-analysis does not allow one to capture the qualitative distinctions between studies. This is related to the need for systematic coding of study characteristics. Yet, this limitation also holds true for the less sophisticated ballot box approach, and plagues virtually any quantitative method of study. Like any other study design, it is possible to combine a qualitative and quantitative approach to capture both a synopsis of findings as well as nuances that might be missed in summarizing a body of literature (Slavin, 1995).

A final key argument is that "blemished" studies are often included in meta-analyses. Since a meta-analysis is summarizing study findings, when study designs in the original works are flawed, that affects the meta-analytic findings. Unfortunately, aside from obvious cannons, there is little agreement in the field as to what constitutes methodological quality (Lipsey and Wilson, 2001). Being too stringent on the required methodological quality results in few studies meeting criteria, which limits the generalizability of findings. On the other hand, too much flexibility will compromise the study findings. To address this issue, researchers have either adhered to strict methodological criteria or treated methodological variation as an empirical matter to be investigated in the study (Lipsey and Wilson, 2001). If there is a significant relationship between quality of the studies and effect size, then differential weighting of studies can be employed (Rosenthal, 1991).

Despite these limitations, the meta-analytic technique is viewed by many as the superior method of summarizing a body of literature. Lipsey and Wilson (2001) suggest that this technique provides an organized way of deciphering information from a large number of studies. Likewise, the systematic coding of study characteristics permits more precise examination between study features and findings than a qualitative review. Simply put, Gendreau, French and Goinet (2004) assert that meta-analysis both limit biases inherent in narrative reviews and offer an easily interpretable conclusion. This is especially valuable when trying to convey research results to practitioners that are not statisticians, but want information on evidence-based practices.

Rehabilitative Strategies[3]

Many meta-analyses have endorsed the narrative reviews that predicated them by 10 to 15 years (Gendreau and Andrews, 1990). That is, meta-analyses have largely demonstrated that treatment is effective at decreasing recidivism among offenders (Davidson, Gottschalk, Gensheimer and Mayer,1984; Garrett, 1985; Mayer, Gensheimer, Davidson and Gottschalk, 1986; Andrews et al., 1990; Izzo and Ross, 1990; Lipsey, 1992; Lipsey and Wilson, 1998; Dowden and Andrews, 1999a; Dowden and Andrews, 1999b; Gendreau, Goggin, Cullen and Andrews, 2000; MacKenzie, 2000; Pearson et al., 2002; Landenberger and Lipsey, 2005; Wilson, Bouffard and MacKenzie, 2005; French and Gendreau, 2006).

The earliest correctional meta-analyses focused on treatment effects for juvenile justice youth. Garret (1985) found that overall, adjudicated delinquents responded positively to treatment. Davidson et al. (1984) reported a similar finding, although he concluded that due to a modest overall effect size, the null hypothesis could technically not be rejected. More recently,

[3] The effectiveness of offender treatment will be discussed in detail throughout the second chapter of this dissertation. This section provides a brief introduction to the effectiveness of rehabilitation as a correctional strategy.

Lipsey has conducted a series of meta-analyses on general preventative and rehabilitative strategies of juvenile delinquents. Studies include examination of specific treatment approaches or intervention types for juveniles (Lipsey and Wilson, 1993; Landenberger and Lipsey, 2005; Lipsey, 2009), treatment of serious juvenile delinquents (Lipsey and Wilson, 1998, Lipsey, 1999), predictors of violence and serious delinquency among juveniles (Lipsey and Derzon, 1998), and variation in treatment effectiveness for juvenile offenders by race (Wilson, Lipsey and Soydan, 2003). Overall, using the meta-analytic approach, Lipsey has found much support for the rehabilitation of youthful offenders, particularly when strategies deemed effective are utilized as the treatment approach.

As with juveniles, rehabilitative strategies have also been found to be effective with non-juvenile populations. Losel (1995) summarized the findings of 13 meta-analyses on offender treatment. He found a general, positive mean effect size for offender treatment programs. Likewise, Dowden and Andrews (1999a) found treatment to be effective with females, while Hanson and Bussiere (1999) found positive treatment effects for sex offender populations. Andrew et al. (1990) conducted a key meta-analysis that served as one of the original tests of the principles of effective interventions[4]. This study found an overall effect size of .10 for treatment programs. Yet when studies were categorized as "appropriate", "inappropriate" and "unspecified", based upon adherence to the principles, the respective effect sizes were .30, -.07 and .13. Hence, there was substantial variability on the effect of treatment based upon the program's application, or lack thereof, of the principles of effective intervention. This early meta-analysis appears to have laid the groundwork for many future studies aimed at exploring the types of intervention most effective at changing offender behavior[5]. As a whole, the use of

[4] The Principles of Effective Intervention will be reviewed in detail later in this dissertation.
[5] Additional meta-analyses will be sited throughout the second chapter that explores offender rehabilitation.

meta-analysis has served as an important strategy for both appraising the correctional treatment literature, and disseminating interpretable findings to providers of correctional interventions.

Deterrence-Based Strategies

Certainly treatment programs are not the only strategy employed to help reduce recidivism among offenders. The correctional system also utilizes incarceration (e.g., jails, detention centers, prisons, juvenile correctional facilities, community residential centers), supervision/monitoring (e.g., probation, parole, intensive supervision probation, electronic monitoring), and criminal sanctions (e.g., fines, restitution, community service) to assist in decreasing the rate of criminal behavior. Incarceration strategies are based, in part, on the theory of incapacitation. This theory represents a strategy for crime control that involves the physical removal of unlawful individuals from society in order to abate crime (Visher, 1987). The remaining strategies, however, in conjunction with incarceration, are based largely upon deterrence theory. Deterrence theory has two components: general and specific deterrence. Specific deterrence suggests that if an individual who engages in unlawful behavior is exposed to a sanction, he or she will be deterred from engaging in that behavior in the future[6]. General deterrence suggests that individuals who observe the negative consequences of others' unlawful behavior will be deterred from engaging in the same behavior (Patternoster, 1987).

Cullen, Pratt, Micelli and Moon (2002) suggest that deterrence models are based upon a faulty theory, namely, rational choice theory. This theory purports that people consciously try to avoid pain and seek pleasure; therefore, in corrections the "cost" of engaging in unlawful

[6] Deterrence theory suggests that the "sanction" must meet certain specifications to be effective: 1) certainty of punishment; 2) severity, in that the punishment must be serve enough to deter the behavior; and 3) celerity or swiftness of punishment (Patternoster, 1987).

behavior must be increased in order to decrease crime (Cullen et al., 2002). Not only does this theory support the courts' use of harsher criminal sanctions, but correctional "programs" have been developed based on specific deterrence strategies, i.e. boot camps, scared straight programs, and drug prevention programs (e.g. DARE). Programs such as these, in their traditional form, are not designed to teach offenders the skills needed to maintain a crime-free lifestyle; rather, these programs are designed to either create an unpleasant environment where structure, discipline, and drill are emphasized (e.g. boot camps) or to attempt to scare offenders out of engaging in an unlawful behavior (e.g. scared straight and DARE). Finally, supervision and monitoring programs are also based upon a specific deterrence model, in that, programs or interventions such as intensive supervision probation, electronic monitoring, or drug testing are assumed to deter individuals from engaging in criminal behavior by increasing the probability of such behavior being detected (MacKenzie, 2000).

Programs such as restitution, fines and community service are based on both deterrence theory (specific and general) and a restorative justice model. Restorative justice suggests that crime causes harm to the victim, the community and the offender, and that the goal of sentencing should be to reverse the harm caused by the criminal act. This can be accomplished by creating sentencing and correctional practices that seek to restore the victims, the community and the offender to their original state (Braithwaite, 2002). Llewellyn and Howse (1998) suggest that there are core tenants to a restorative justice program: a process that involves voluntary participation, truth telling, and a face-to-face meeting. Hence, restitution, fines and community service are not traditional restorative justice programs; rather they pull from both restorative justice and deterrence philosophies in responding to offending behavior. For example, requiring an offender to pay restitution for a criminal act is not only unpleasant for the offender, thereby

deterring future criminal acts, but it also attempts to diminish the harm caused to the victim(s) of the act.

With regard to effectiveness of restorative justice approaches, Bonta, Wallace-Capretta, and Rooney (1998) conducted a meta-analysis looking at the model's impact on recidivism. The results revealed that restorative justice models had a modest ability to reduce recidivism (ES=.08). In this study, a very broad definition of restorative justice was used, one that encompassed court ordered restitution and community service programs. Latimer, Dowden and Muise (2005) conducted a later meta-analysis narrowing program inclusion to those that met a more true definition of restorative justice. Specifically, they studied conferences and victim-offender mediations which were voluntary, community-based, and sought to bring together the offender, victim and community. Furthermore, they expanded the outcome variables to include victim and offender satisfaction, restitution compliance, and recidivism. Latimer et al. (2005) found that the more traditional restorative justice programs were significantly more effective than non-restorative approaches to criminal justice. Yet, the researchers caution that there is an inherent self-selection bias related to the criterion that offenders must volunteer to participate in the program. Furthermore, the average effect size for the restorative justice approach (.07) is still substantially lower that what has been found for evidence-based rehabilitation programs.

Evidence of Effectiveness of Deterrence-Based Strategies

As a whole, there is significantly less empirical support for incapacitation and deterrence-based correctional strategies as compared to human service and treatment strategies (Dowden and Andrews, 1999b, Gendreau et al., 2000, MacKenzie, 2000, Gendreau et al., 2004, Lipsey, 2009). Gendreau et al. (2000) conducted a meta-analysis that included 140 studies of community sanctions and 325 studies of incarceration. With regard to incarceration, offenders

with lengthier sentences demonstrated a slight *increase* in recidivism (3%), which offers little support that harsher sentences are effective at deterring criminal behavior. They found that among the community sanctions (intensive supervision, arrest, fines, restitution, boot camps, scared straight programs, drug testing and electronic monitoring) the overall effect size was .00. Consistent with Bonta et al. (1998), restitution and fines were the only categories showing a hint of reduction in recidivism. Notwithstanding, the introduction of treatment to these programs increased the effect size to .10. The authors therefore concluded that the effectiveness of intermediate sanctions is mediated solely through the provision of treatment (Gendreau et al., 2000). Lipsey's more recent 2009 meta-analysis supports these conclusions; he found that interventions using a "therapeutic" philosophy decreased recidivism up to 12 percent, while philosophies based on control, such as deterrence or discipline programs increased recidivism up to 8 percent.

Petersilia and Turner (1993) conducted one of the largest social science studies to date that used an experimental design to randomly select offenders for intensive supervision probation (ISP) or general probation. They found that while the ISP offenders had higher rates of technical violations (likely due to closer monitoring) there were not significant differences in the rates of rearrest or reconviction for a new crime between the two groups (Petersilia and Turner, 1993). Hence, ISP was no more effective than general probation at reducing the rate of criminal behavior. Yet, like Gendreau et al. (2000), the researchers found that higher levels participation in treatment programs were associated with a 10-20 percent reduction in recidivism (Petersilia and Turner, 1993). In her review of the research of intermediate sanctions, Petersilia (1998) concluded that intermediate sanctions are desirable for offenders for whom incarceration is

unnecessarily severe but ordinary probation is too lenient. However, without a rehabilitative component, she suggests that reductions in recidivism are elusive (Petersilia, 1998).

A recent study on Intensive Supervision Programs support Petersilia's earlier findings. Lowenkamp, Flores, Holsinger, Makarios and Latessa (2010) examined fifty-eight ISPs to explore the role of both the program's philosophy and treatment integrity in producing desirable effects. They found that as a whole, human service oriented programs had a mean effect size of .06 while deterrence oriented programs showed negative program effects (-.11). The use of effective treatment characteristics also significantly improved the programs' outcome; however, the study found that while programs that adhered to effective treatment strategies were more effective, a deterrence-based program philosophy eliminated these positive treatment effects. Hence, similar to Lipsey (2009), they found that a human service (rather than deterrence-oriented) philosophy was essential in producing positive treatment effects.

With regard to correctional programs using a deterrence-based model, MacKenzie, Wilson and Kider (2001) conducted a meta-analysis of 29 studies on the effectiveness of boot camps at decreasing recidivism. She and her colleagues found there were no overall significant differences in recidivism rates between boot camps and comparison samples. Hence, boot camps were not effective at decreasing recidivism among participants.

In synopses of the literature on what works and what does not work to reduce recidivism, MacKenzie (2000) concluded that programs that emphasize specific deterrence (e.g., shock probation, scared straight), programs designed to increase control and surveillance in the community (e.g., intensive supervision probation, home confinement/electronic monitoring, residential programs, urine testing) and programs emphasizing structure, discipline and challenge (e.g., boot camps, wilderness programs) do not work to reduce recidivism. Similarly, Gendreau

et al. (2004) suggest that intermediate sanctions showed no appreciable reductions in recidivism, and were therefore classified among the "What Does Not Work" strategies (p. 28).

In conclusion, criminal sanctions and deterrence-oriented programs without a core treatment component are largely ineffective at reducing criminal behavior. This deduction leads to the next section of this dissertation. Here, attention will turn to specific details regarding effective correctional strategies, specifically, what characteristics differentiate effective and ineffective treatment programs. Certainly, the above section highlighted the importance of the incorporation of treatment into correctional interventions. The literature also suggests that not all treatment is the same; therefore a detailed review of the principles of effective intervention will be outlined in what follows.

The Principles of Effective Intervention

Overview of the Principles

As a whole, correctional treatment programs reduce recidivism by approximately 10 percent (Andrews et al., 1990; Gendreau et al., 2004). However, as previously mentioned, there is great variability among programs with respect to the degree of effectiveness. Programs that opt for evidence-based treatment strategies see two to four times greater effects than the 10 percent average (Gendreau et al., 2004). On the other hand, treatment providers have also inadvertently *increased* the likelihood of recidivism by applying treatment without using evidence-based strategies (Dowden and Andrews, 1999b; Lowenkamp 2004; Lowenkamp, Latessa and Lemke, 2006) or employing such strategies with poor fidelity (Andrews and Dowden, 1999; Barnoski, 2004).

The following section will review the core evidence-based correctional treatment strategies. These strategies will be summarized into 4 principles of effective intervention: 1) the

risk principle, 2) the need principle, 3) the responsivity principle, and 4) the fidelity principle. Key Canadian researchers (Andrews et al., 1990; Gendreau, 1996; Andrews and Bonta, 2010) have contributed to outlining the principles of effective intervention and providing empirical support for these principles. The definition and empirical status of each of the four principles will be discussed, followed by a review of the method used to measure programmatic application of the principles of effective intervention.

The Risk Principle

Andrews, Bonta and Hoge (1990) outline two key aspects of the risk principle: 1) classification of offenders based upon likelihood of recidivism; and 2) matching level of service to level of risk. Stated simply, the risk principle suggests that offenders be classified by risk and intensive correctional interventions be reserved for higher risk offenders (Andrews and Bonta, 2010). Examples of correctional interventions include incapacitation and supervision strategies such as incarceration, half-way house placement, intensive probation, general probation, electronic monitoring or day reporting. Correctional interventions also include treatment programs, such as community correctional facilities, group homes, intensive outpatient treatment, group treatment, family interventions or case management. The risk principle again suggests that more intensive interventions, particularly those that involve incarceration or other forms of out of home placement as well as community-based intensive supervision or treatment be reserved for moderate to high risk offenders (Lowenkamp and Latessa, 2002, Andrews and Bonta, 2010). In order to identify such offenders, a validated risk assessment must be employed.

Risk Assessment

There is a plethora of research supporting one's ability to predict risk using actuarial tools (Hoffman and Beck, 1980; Shields and Simourd, 1991; Harris, Rice and Quincy, 1993; Harris,

1994; Hoge and Andrews, 1996; Lowenkamp, Holsinger and Latessa, 2001; Barbaree, Seto, Langton and Peacock, 2001; Gendreau, Goggin, and Smith, 2002). Bonta (2002) outlines ten guidelines for selection and use of correctional risk assessments. Such guidelines include using a theoretically-based, actuarial and validated tool, which includes multiple criminogenic need[7] domains and limits incorporation of non-criminogenic factors. Gendreau, Little and Goggin (1996) support the argument that risk assessments should assess dynamic as well as static factors. Not only are dynamic factors important for identifying targets for change, but these researchers found that the dynamic predictors performed at least as well as the static predictors (Gendreau et al. 1996). Tools such as the Level of Service Inventory (both the adult and youth version) have multiple studies supporting its ability to accurately predict risk (Shields and Simourd, 1991; Gendreau, Little and Goggin, 1996; Gendreau, Goggin and Smith, 2002; Lowenkamp, Holsinger, Brusman-Lovins and Latessa, 2004; Flores, Lowenkamp, Smith and Latessa, 2006; Bechtel, Lowenkamp and Latessa, 2007), including prediction with special populations (Lowenkamp, Latessa and Holsinger, 2001; Holsinger, Lowenkamp and Latessa, 2006) and in predicting institutional misconduct (Holsinger, Lowenkamp and Latessa, 2006a). As such, correctional entities have access to a ranges of predictive tools which allows them the accurately classify offenders according to risk.

Empirical Findings on the Risk Principle

Based upon the empirically informed assumption that one can accurately predict risk by using a validated, actuarial tool, how this tool is incorporated into practice speaks to the second component of the risk principle: matching offenders to interventions based on risk level. Andrews et al. (1990) was one of the first studies to support application of the risk principle.

[7] Criminogenic needs will be defined and discussed in the Need Principle section.

This meta-analysis was designed as an original test of the risk, need and responsivity principles. The researchers found, as reviewed earlier, that interventions classified as "appropriate" produced effect sizes that were 36 percentage points higher than interventions classified as "inappropriate". Appropriate interventions were defined, in part, by the targeting of higher risk cases. Unfortunately, the effects of application of the risk principle could not be disaggregated from application of the need and responsivity principles.

Andrews and Dowden have since conducted a series of meta-analyses testing application of the risk principle on various populations (Dowden and Andrews, 1999a; Andrews and Dowden, 1999; Dowden and Andrews, 2000; Dowden and Andrews, 2006). One such meta-analysis used 229 studies to examine the treatment of juvenile offenders (Dowden and Andrews, 1999b). They found that juvenile programs that adhered to the risk principle produced an effect size of .13 compared to an effect size of .03 for programs failing to adhere to this principle. Dowden and Andrews (1999a) conducted a meta-analysis of 26 studies to explore application of the risk, need and responsivity principles to female offenders. Here too they found that programs that primarily targeted higher risk women produced an effect size of .19, while programs targeting lower risk women increased likelihood of recidivism (ES = -.04).

Dowden and Andrews (2000) studied the impact of the risk principle on violent offending. Although findings for this study were non-significant, they found that programs that folled to the risk principle produced an effect size of .09 compared to .04. Furthermore, Dowden and Andrews (2006) conducted an exhaustive meta-analytic review of the risk principle where they found moderate support, particularly for females and young offenders in programs where the need and responsivity principles were met. In a more recent meta-analysis with juvenile offenders, Lipsey (2009) found that risk level of participants was a robust indicator of program

success. He found that interventions applied to high risk juveniles were substantially more effective, although the effect was offset somewhat by juveniles with histories of violent behavior.

Other key researchers investigating application of the risk principle to various populations are Lowenkamp and Latessa (Lowenkamp and Latessa, 2002; Lowenkamp, Latessa and Lemke, 2006, Lowenkamp, Latessa and Holsinger, 2006). Rather than using meta-analyses, these researchers have relied primarily on large statewide data sets that include individual and program-level data. This has arguably allowed for a more accurate assessment of risk since meta-analyses are limited in their ability to fully investigate the relationship between risk level and recidivism due to poor reporting practices of level of risk in the primary studies (Dowden and Andrews, 2006).

Lowenkamp and Latessa (2002) conducted a study that examined the relationship between treatment effectiveness, risk level and recidivism rates for over 15,000 adult offenders placed in 54 half-way houses and community based correctional facilities in Ohio. This study found that high risk offenders placed in these programs averaged an 8 percent reduction in recidivism, relative to a comparison sample; to the contrary, low risk offenders placed in these same facilities *increased* their likelihood of recidivism by an average of 4 percent. A follow-up study conducted in 2010[8] that examined over 20,000 offenders reached the same conclusion--residential treatment options should be reserved for higher risk offenders. Among successful halfway house completers, low risk offenders showed a 5 percent *increase* in recidivism over matched comparison cases while high risk offenders showed nearly a 15 percent reduction in recidivism (Latessa, Lovins and Smith, 2010).

[8] Data for the current dissertation comes from this follow-up study.

A study by Lowenkamp, Latessa, and Lemke (2006) also found support for the risk principle with juvenile offenders. They found that the average effect size for juvenile residential programs that adhered to the risk principle (as defined by accepting no more that 25% low risk offenders) was .14; the effect size for juvenile residential facilities that failed to employ the risk principle (more than 25% of residents were low risk) was -.1. Finally, Lowenkamp, Latessa and Holsinger (2006) in a study of 97 residential and non-residential correctional programs 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.

Although not a direct study of the risk principle, Burgeon and Armstrong (2005) conducted a study that examined the effect of treatment dosage on recidivism. Their study included 620 incarcerated male offenders, 482 of which were assigned to a 5, 10 or 15-week prison-based treatment program and 138 of which were assigned to no treatment. The study required that staff conduct a series of assessments and then assign offenders to the groups based upon risk, need, and specific responsivity considerations. Results indicated that the prison-based program was effective at reducing recidivism by approximately 10 percent, and that dosage played a mediating role in the effectiveness of treatment. The authors concluded that moderate risk offenders with few needs could be sufficiently treated with 100 hours of service, offenders who are high risk *or* high need required 200 hours of treatment, and high risk offenders with multiple needs may require in excess of 300 hours of treatment. Notably, risk level and criminogenic needs were the primary considerations regarding dosage decisions. Overall, this study not only supports the risk principle by finding that higher risk offenders responded best to

higher level intervention, but it provides preliminary data on what dosage of treatment is needed to produce positive treatment effects.

In conclusion, there is ample evidence to suggest that correctional interventions, including both supervision and treatment programs, should 1) assess risk using a validated risk tool, and 2) use classification information to make decisions about appropriate matching of offenders to services. Findings from many of the aforementioned studies suggest significant harm associated with application of intensive correctional interventions to low risk offenders. As such, it is important that correctional programs develop exclusionary criterion that incorporates risk, and then attempt to differentiate treatment or supervision based upon level of risk so that low risk offenders are not exposed to intensive correctional strategies.

The Need Principle

The need principle suggests that in order to reduce future criminal behavior, programs must target those dynamic risk factors[9] associated with criminality, namely criminogenic needs (Andrews, 1999; Andrews, Bonta, Hoge, 1990; Gendreau, 1996; Andrews and Bonta, 2010). Andrews and Bonta (2010) review the core criminogenic need areas: antisocial attitudes, antisocial associates, antisocial personality features (e.g., impulsivity, risk-taking behavior, lack of empathy, aggression/hostility, poor problem-solving skills), family relationships, substance abuse, educational/vocational achievement, and structured leisure time. Like with the risk principle, application of the need principle begins with appropriate assessment of a range of criminogenic needs. As such, a validated risk/need tool[10], such as the Level of service Inventory (Andrews and Bonta, 1995), should be utilized by treatment and supervision programs to

[9] Dynamic risk factors are amenable to change whereas static risk factors cannot be decreased.
[10] Examples of other validated risk/need tools include the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS), the Hare Psychopathy Checklist-Revised (adult, youth and screening versions—Hare, 2003), the Wisconsin Risk/Need Assessment (adult and juvenile version), and the Youth Assessment and Screening Index (YASI).

effectively identify criminogenic needs. Once identified, intermediate objectives can be developed around the criminogenic needs, which should form the basis of a correctional case plan (Bonta, 2002). Such practices are important to help ensure that the program's primary targets are criminogenic in nature.

Empirical Findings on the Need Principle

There is much empirical evidence supporting the aforementioned criminogenic needs as accurate predictors of risk for criminality (Zamble, 1993; Simourd and Andrews, 1994; Hoge and Andrews, 1996; Gendreau, Little and Goggin, 1996; Lipsey and Derzon, 1998; Jones-Hubbard and Pratt, 2002; Gendreau, Goggin and Smith, 2002). There is also research that identifies a range of *non*-criminogenic factors that are often seen as correctional program targets. Examples of such targets include self-esteem, mental health issues, vague emotional problems, physical activity, bonding among antisocial peers, and fear of punishment. The research suggests that these needs are not predictive of recidivism; therefore targeting these areas fails to decrease an offender's propensity to commit crime (Dowden and Andrews, 1999, Gendreau, Little and Goggin, 1996, Hanson and Bussiere, 1998).

Like with the risk principle, there also exists a body of literature that tests whether programs that adhere to the need principle by targeting primarily criminogenic need factors are more effective at reducing recidivism among participants (Andrews et al., 1990; Dowden and Andrews, 1999; Andrews and Dowden, 1999; Dowden and Andrews, 2000; Gendreau, French and Taylor, 2002; Lowenkamp, 2004). Again, Andrews et al. (1990) began the process of systematically testing the principles of effective intervention, and found support for application of the need principle. Yet, while they found substantial differences between "appropriate" and "inappropriate" programs, measured in part by targeting criminogenic need factors, this study did

not isolate the specific effects of targeting criminogenic needs. Another meta-analysis of 52 studies aiming to determine effective treatment strategies for violent offenders also found support for the need principle (Dowden and Andrews, 2000). In programs where the majority of targets were criminogenic, the effect size averaged .20 compared to .00 for programs targeting mostly non-criminogenic needs. These same researchers examined application of the need principle to female offenders (Dowden and Andrews, 1999a). They also found substantial differences in effect sizes based on targeting primarily criminogenic needs (ES=.26) versus targeting primarily non-criminogenic needs (ES=.04). This study was instrumental in providing empirical evidence of application of the principles to specialized populations.

Rather than simply looking at whether the majority of program targets were criminogenic in nature, Gendreau, French and Taylor (2002) conducted a meta-analysis that looked at the proportion of criminogenic versus non criminogenic targets and how this related to effect sizes. They found an effect size of .31 for programs that targeted 4-6 more criminogenic needs. Not only does this study find support for the need principle, it also provides more detailed information related to the density of criminogenic targets necessary to produce substantial decreases in recidivism. Similarly, a recent meta-analysis by French and Gendreau (2006) examined 68 studies to determine the effect of correctional treatment on reducing prison misconducts. Specific to the need principle, they found that the average effect size was .29 for programs that targeted 3-8 criminogenic needs, .16 for programs that targeted 1-2 criminogenic needs, and .06 for programs that failed to target any criminogenic needs. They also found that programs effective in reducing misconducts were also more effective in decreasing recidivism. Again, this study supported the importance of adequate density of criminogenic targets.

The Lowenkamp (2004) study which examined characteristics of residential community correctional programs found not only that targeting criminogenic need factors was positively correlated with program effectiveness, but that programs using standardized methods to assess risk and need were more effective at reducing recidivism. Specifically, survey results found that the number of criminogenic services offered by the program was a significant predictor of recidivism (r=.39) as was use of a standardized need assessment (r=.33).

Finally, a study conducted by the Pennsylvania Department of Corrections found that key criminogenic needs were important indicators of success on parole. Zajac and Bucklen (2006) conducted a survey of offenders paroled to the community. Successful parolees were identified as those remaining in the community at least three years, whereas unsuccessful parolees were revoked to PADOC. Contrary to the popular assumption that employment and housing are the most important factors for successful reentry, Zajac and Bucklen found three primary themes differentiating parole successes from parole failures: 1) parole violators were more likely to have unrealistic post-release life expectations; 2) parole violators were more likely to maintain antisocial attitudes and beliefs; and 3) parole violators were more likely to lack effective problem-solving/coping skills. Furthermore, other key criminogenic needs also played a role in return to prison, including antisocial peers and drug and alcohol use. They also found evidence for protective factors associated with success on parole, including having a positive relationship with a spouse, having a mentor, and having a positive experience in a community correctional facility or with one's parole officer. Overall, this study supports the need to target the major criminogenic need factors linked to recidivism in order to assist offenders with successful reentry to the community.

In summary, the need principle suggests that programs that focus primarily on criminogenic need factors are more successful at reducing recidivism and increasing success rates among participants. In order to do so, programs must adopt a validated need assessment(s) so that the relevant criminogenic needs can be identified. Recent studies (Gendreau, French and Taylor, 2002; French and Gendreau, 2006) further demonstrate that a higher density of criminogenic targets are more strongly related to program effectiveness. This distinction is important, particularly as the next principle, responsivity, is explored as this principle suggests that programs must also consider non-criminogenic responsivity issues when appropriate. Nonetheless, the evidence is clear that correctional programs must concentrate on targeting and reducing criminogenic need factors if they wish to be effective at reducing recidivism.

The Responsivity Principle

The responsivity principle consists of two key components: *general responsivity* and *specific responsivity*. General responsivity suggests that most offenders respond to a particular model of treatment, namely, cognitive, behavioral and social learning approaches. Specific responsivity asserts that offenders have different learning styles and may possess barriers that impact their ability to succeed in programming/interventions (Cullen, 2002). Hence, as a whole this principle maintains that 1) treatment programs should utilize cognitive-behavioral and social learning approaches, but also 2) programs should individualize services so that individual learning styles are considered and barriers (e.g. transportation, mental health issues, and motivation) are addressed in order to increase the likelihood of treatment success.

General Responsivity

Bandura (1973) outlines two major components to social learning/cognitive-behavioral theory, namely, cognitive restructuring and cognitive skills. According to this model, criminal

behavior, as well as cognitions that prompt, support and reinforce such behavior, are learned from individuals who model antisocial attitudes and behavior. As such, practitioners must assist offenders to restructure maladaptive and antisocial thinking. Likewise, offenders must be taught skills to better manage their thoughts and behaviors.

Spiegler and Guevremont (1998) outline four defining themes of behavior therapy: scientific, active, focused on the present, and teaching model. These themes help to explain why this therapeutic model works well in the treatment of offenders. First, cognitive-behavioral treatment is *scientific*. As such, treatment targets are measurable and providers can determine via the measurement of acquisition of skills when an offender has successfully completed treatment. Likewise, cognitive behavioral treatment is scientific in that it is an empirically validated approach, unlike many of the traditional psychotherapeutic approaches. Second, cognitive behavioral interventions are *active*. The expectation of a cognitive-behavioral model is that offenders learn new skills for managing their lives, as opposed to traditional talk therapies which simply engage offenders through conversations and discussions regarding their issues. The third characteristic is a *focus on the present*. As such, offenders are exposed to an intervention that can take hold immediately as it is designed to address current conditions that maintain problematic behavior. Finally, this model is focused on *teaching*. The goal of the cognitive-behavioral approach is to help offenders identify high risk situations, thoughts, feeling and behaviors, and then teach them more prosocial ways of managing these risk factors.

Empirical Findings on the General Responsivity Principle

There is ample evidence to support the notion that correctional programs experience the greatest effects when they utilize a cognitive-behavioral or social learning approach (Gendreau and Ross, 1987; Garrett, 1985; Mayer et al., 1986; Andrews et al., 1990; Izzo and Ross, 1990;

Antonowicz and Ross, 1994; Palmer, 1995; Henning and Frueh, 1996; Lipsey and Wilson, 1998; Andrews and Dowden, 1999; Baro, 1999; Lipsey, Chapman and Landenberger, 2001; Lipsey, 2001, 2009; Hanson, Gordon, Harris, Marques, Murphy, Quinsey and Seto, 2002; Pearson et al., 2002, Lowenkamp, 2004; Landenberger and Lipsey, 2005; Wilson, Bouffard and MacKenzie, 2005; French and Gendreau, 2006). Both early and more recent narrative reviews (Gendreau and Ross, 1987, Palmer, 1995, MacKenzie, 2000) reveal that programs using a cognitive-behavioral model are more effective at reducing recidivism. Early reviews formed the basis for development of the principles of effective intervention. Later reviews resembled prior meta-analyses and other empirical studies to reach the same conclusion.

Henning and Freuh (1996) examined the effectiveness of Bush's Cognitive Self Change program used by the Vermont Department of Corrections and found a 29 percent reduction in the treatment group over the comparison sample, supporting use of this cognitive-behavioral program. Likewise a study of the "real world" application of Thinking for a Change (TFAC)[11] found that after controlling for risk level, age, race, and time at risk, offenders that participated in TFAC were rearrested 15 percentage points less frequently than a comparison sample (Lowenkamp, Hubbard, Makarios, and Latessa, 2009). They Finally, Baro (1999) studied the effects of randomly assigning 123 offenders to a cognitive restructuring or self-help group on misconducts. He found that while there was no significant differences between the groups in total number of misconducts received, certain types of misconducts decreased for the cognitive restructuring group, namely, disobeying a direct order and assault.

The early meta-analyses, conducted primarily on youth, also find support for the general responsivity principle. Garrett (1985) in his study of juvenile delinquents found the largest effect

[11] The development of Thinking for a Change was funded by the National Institute of Corrections and it was authored by Bush, Glick, and Taymans. The curriculum uses a cognitive-behavioral model and includes a cognitive restructuring, social skill development and problem solving component.

sizes for social learning, family therapy and cognitive approaches, which supported the Davidson et al. (1984) finding that social learning approaches were particularly effective. Izzo and Ross (1990) found that cognitive behavioral programs for juvenile delinquents were more than twice as effective as non-cognitive approaches. Lipsey (1992) found that overall, juvenile treatment programs were successful at reducing recidivism by about 10 percent; however, when programs were multimodal and used a cognitive behavioral design, effect sizes increased to nearly .30. More recently, Lipsey (2009) found that the largest mean effect size for any intervention appeared for cognitive-behavioral programs (.26). Likewise, Dowden and Andrews (2000) found that cognitive-behavioral treatment reduced recidivism in serious, violent offenders 18 percent above all other types of treatment combined.

Meta-analyses on specialized populations also demonstrate the importance of using a cognitive-behavioral model. Dowden and Andrews (1999) found that the principles of effective intervention, including use of a cognitive behavioral model, were applicable to females. Likewise, Hanson et al. (2002) studied treatment effects with sex offenders and found that cognitive behavioral and systemic approaches decreased recidivism for both general and sexual offenses. In studying effective treatment approaches for substance abusers, Taxman (2000) categorized cognitive behavioral interventions as an effective approach while Pearson and Lipton (1999) deemed it a promising approach.

Some meta-analyses have focused solely on social learning and cognitive-behavioral programs. Mayer et al. (1986) reviewed only programs using a social learning model and found that programs were able to significantly impact recidivism as well as attitude and other maladaptive behaviors. Lipsey (1999) conducted a study of cognitive-behavioral juvenile programs and found that treated offenders recidivated at a rate 2/3 that of the comparison group.

A meta-analysis of 58 studies of adults and juvenile programs found that cognitive behavioral programs reduced recidivism by an average of 25 percent (Landenberger and Lipsey, 2005). 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. Finally, Wilson, Bouffard and MacKenzie (2005) conducted a review which identified specific cognitive-behavioral treatment programs that produced positive reductions in recidivism, namely, Moral Reconation Therapy, Reasoning and Rehabilitation (Ross and Fabiano, 1985), and various cognitive restructuring programs.

In summary, there is a plethora of evidence to suggest that following the general responsivity principle, i.e., use of a cognitive-behavioral or social learning treatment modal, is essential for correctional treatment programs to in order to decrease recidivism. There is both a clear theoretical rationale for use of behavioral strategies to modify offender behavior, as well as empirical support for application of this model. The next section will review the second component of the responsivity principle, namely specific responsivity.

Specific Responsivity

The specific responsivity principle acknowledges that not all individuals respond to treatment the same way; therefore, it is imperative for programs to address individual learning styles and barriers to treatment while delivering effective interventions (Andrews and Bonta, 2010). Some offender needs that could be responsivity factors include lack of motivation, transportation issues, mental health issues, language barriers, ethnicity or cultural issues and housing. None of these needs are strongly correlated with future offending. As such, if a program only targets these factors and ignores the criminogenic needs of offenders, the program will have little impact on the future criminal behavior of participants (Andrews and Bonta,

2010). In contrast, if a program ignores important responsivity issues (i.e. acute mental illness), the likelihood of the respective offender completing treatment services is low, which also makes it unlikely that the program will be successful in decreasing criminality.

Like with risk and need, it is important that responsivity factors be measured using standardized and validated tools (Van Voorhis, Braswell and Lester, 2007). Examples of assessments include the Culture Fair IQ Test (Cattell, 1973), the Minnesota Mutiphasic Personality Inventory (Megargee, Carbonnell, Bohn and Sliger, 2001), the Interpersonal Maturity Level (Warren, 1966), the Jesness Inventory (Jesness and Wedge, 1983), and the University of Rhode Island Change Assessment (McConnaughy, Prochaska, and Velicer, 1983). Once assessed, offenders should be matched to interventions which will increase their likelihood of success in the program. Jesness (1988) argues that it is important to distinguish among offenders based upon personality for both management and treatment purposes.

Empirical Findings on the Specific Responsivity Principle

There is less empirical data on the specific responsivity principle when compared to the risk, need and general responsivity principles. One explanation for this is that adherence to the specific responsivity principle is difficult to measure, as there are multiple responsivity variables to consider. Nonetheless, studies have evaluated the impact of specific responsivity factors on recidivism. Andrews et al. (1990) in their study of the principles of effective intervention did include specific responsivity considerations as a necessary component within the "appropriate" services category, which as specified earlier, produced significantly higher effects then the other service categories (ineffective, unspecified and criminal sanctions). Yet, again, this study does not represent a direct analysis of the specific responsivity principle.

Van Voorhis, Spruance, Ritchie, Listwan, Seabrook, and Pealer, (2002) however, conducted a study directly measuring the impact of specific responsivity. This study examined the effect of a cognitive skills program on reincarceration when specific personality subtypes are taken into consideration. They used the Jesness Inventory (1996) to classify offenders into four personality subtypes: asocial aggressive; neurotic; dependent; and situational. Findings indicated that the cognitive skills program was most appropriate for dependent and situational offenders. Neurotic offenders, on the other hand, responded poorly to the program, with recidivism rates nearly 20 percent higher than the comparison group. This study clearly demonstrates the mediating effect responsivity factors, such as personality, can have on otherwise effective correctional interventions. Furthermore, Van Voorhis, (1987) expressed concern that the failure to incorporate specific responsivity into correctional treatment or program evaluations may be "masking" treatment effects.

Sperber (2003) also studied differences in personality types using the Jesness inventory within a group of child molesters. In a sample of 85 child molesters, she found significant differences among offenders based on classification by personality type (neurotic, dependent, situational, and aggressive). She found that neurotic molesters scored highest on an emotional and intellectual empathy scale, but lowest on self-esteem. She found that dependent molesters were more likely to deny blame. Finally, situational child molesters scored lowest on emotional empathy. She concludes that addressing such responsivity characteristics can increase program effectiveness by individualizing treatment targets.

Wilson, Lipsey and Soydan (2003) conducted a meta-analysis on the effectiveness of mainstream service programs for minority juvenile justice youth as compared to non-minority (White) youth. They were interested in whether mainstream interventions needed to be

culturally tailored in order to address delinquency issues. Hence, differences in several outcome measures (including recidivism, academic performance, peer relations, behavior problems, psychological adjustment and attitudes) between minority and non-minority youth were compared. The results showed that overall, services were equally effective for minority and White youth. Wilson et al. (2003) concluded that mainstream intervention programs could be effectively used with minority youth without culturally tailoring the programs. While these results offer limited support of the specific responsivity principle as it relates to ethnicity/race, they do support the need principle in that treatment targets should reflect those dynamic needs associated with recidivism, which do not appear to vary by race. These findings are also consistent with Lipsey's recent 2009 meta-analysis where he concluded that with few exceptions, correctional interventions for juveniles were neither more nor less effective for minorities.

In summary, the specific responsivity principle requires correctional providers to consider individual learning styles as well as characteristics that may serve as barriers to success in treatment. While this principle has not attracted the empirical attention that the other principles of effective intervention have, there is still notable evidence of the importance of considering the impact of specific responsivity issues. Together, general and specific responsivity assist programs in selecting an appropriate treatment model, namely, cognitive-behavioral/social learning and then modify programming to meet the individual learning styles and needs of offenders.

The Fidelity Principle

As indicated in this review of the principles of effective intervention, there is ample empirical support for each of the risk, need and responsivity principles. As such, it is clear from the literature that programs that use a cognitive-behavioral approach to target the criminogenic

needs of higher risk offenders are most effective at reducing recidivism (Gendreau, French and Gionet, 2004). The correctional field has therefore largely directed its attention on *how* to effectively implement these principles. As such, fidelity, or the ability of programs to implement these strategies appropriately, has become a principal concern. The fidelity principle asserts that for evidence based programs to be effective, they must be implemented as designed. In an extensive narrative review of the literature, Gendreau and Ross (1987) also concluded that many programs fail due to lack of therapeutic integrity.

Empirical Findings on the Fidelity Principle

When otherwise effective programs fail to be implemented correctly, treatment effects are impacted. An outcome study was conducted on two evidence-based treatment programs used with juvenile justice youth, namely Functional Family Therapy (FFT) and Aggression Replacement Training (ART)[12] (Barnoski, 2004). This study found that competent delivery of ART and FFT resulted in significant reductions in recidivism (24% and 38% respectively). Yet, when these same programs were incompetently administered, both programs resulted in an *increased* rate of recidivism among participants (10% with ART and 17% with FFT). Barnoski (2004) further examined the effect of therapist competency ratings on recidivism and found that offenders assigned to therapists rated as "highly competent" and "competent" recidivated at a rate of 14 percent and 17 percent respectively, compared to the control group who reoffended at a rate of 22 percent. Therapists rated as "marginal" and "not competent" recidivated at a rate of 25 percent and 29 percent respectively. This study clearly demonstrates the effect of poor fidelity.

[12] Aggression Replacement Training (Goldstein, Glick and Gibbs, 1998) and Functional Family Therapy (Gordon, Arbuthnot, Gustafson, and McGreen, 1998) have evidence of effectiveness with juvenile justice populations.

Many factors impact proper implementation of a program. Studies have found that programs that have an involved researcher or evaluator experience higher treatment effects (Andrews and Dowden, 1999; Lipsey, 1999, 2009). Lipsey (2009) found in his meta-analysis of juvenile correctional interventions that interventions implemented with high fidelity were more effective; high fidelity was measured, in part, by having an involved researcher in the program implementation. Another study identified several key influences on program integrity (Andrews and Dowden, 1999). Factors included such things as training and supervising workers, using structured manuals, following a specific program model, monitoring program change, administering an adequate dosage of the intervention, and like the previously mentioned study, having an involved evaluator. Hence, they found additional factors that affect contribute to evidence-based programming (Andrews and Dowden, 1999).

A closer examination of the importance of staff practice in delivering effective correctional treatment was conducted by Dowden and Andrews (2004). They found that for programs adhering to the risk, need and responsivity principles, additional core correctional practice (CCP) [13] variables independently contributed to increased program effects. Examples of such variables include relationship establishment, effective reinforcement and disapproval, problem solving, structured skill building, effective modeling and effective use of authority. Only advocacy/brokerage failed to reach significance. They concluded that staff characteristics and training in core skills must be addressed to ensure maximum therapeutic effect.

French and Gendreau, (2006) studied the effects of various program characteristics on misconducts. They found that the average effect size for high quality programs was .38, relative

[13] Andrews and Kiessling (1980) originally identified five dimensions of effective correctional practice, i.e. core correctional practices, which helped to increase the therapeutic potential of rehabilitation programs, namely effective use of authority, anticriminal modeling and reinforcement, problem solving, use of community resources, and quality of interpersonal relationships between staff and clients.

to a .20 effect size for programs rated as "moderate" and .13 for those rated as "low". Thus, not only does program quality affect recidivism rates among participants, but it also affects the rate of misconducts within a program. Furthermore, Lowenkamp, Flores, Holsinger, Makarios and Latessa (in press) examined how adherence to treatment integrity impacted the effectiveness of fifty-eight Intensive Supervision Programs (ISPs). They found that treatment integrity measures that were consistent with the principles of effective intervention increased the effect size of the ISP (from -.09 for programs with a low treatment rating and .14 for programs with a high treatment rating). This study is important in that it expands application of program fidelity to community supervision practices.

As with the other principles, in order to accurately measure program quality, standardized, validated tools must be employed. The next section will explore the role of program evaluation in measuring a program's fidelity to evidence based practices.

The Role of Program Evaluation in Adhering to the Fidelity Principle

Gendreau (1996) in his description of the principles made reference to the importance of evaluating the effectiveness of services so as to appropriately brokerage for clients. The Correctional Program Assessment Inventory or CPAI (Gendreau and Andrews, 1989, 2001) is a standardized assessment tool designed to measure how closely a correctional program adheres to the principles of effective intervention. This tool has undergone predictive validity studies. Lowenkamp (2004) reported a strong correlation between the CPAI total score and recidivism (r=.41), while Nesovic (2004) reported a similar correlation (r=.46). Likewise, French and Gendreau (2006) found programs classified as higher quality produced lower misconduct rates. As a process evaluation tool, the CPAI serves a valuable purpose of providing insight into the

"black box" of a program to determine *why* a program may or may not be achieving an effective outcome (French and Gendreau, 2006).

Several researchers have noted the importance of standardized evaluation of correctional programs (Van Voorhis, 1987; Van Voorhis and Brown, 1996; Latessa and Holsinger, 1998; Gendreau, Goggin and Smith, 1999; Matthews, Hubbard and Latessa, 2001; Lowenkamp, Latessa and Smith, 2006). Latessa and Holsinger (1989) and Matthews et al. (2001) outlined the key issues surrounding program effectiveness and use of the CPAI to assess programs. Gendreau, Googin and Smith (1999) provided guidelines for the implementation of effective correctional programming. The 32 guidelines were divided into four general categories: general organizational factors, program factors, change agent activities, and staffing. They recommended use of the CPAI for programs concerned with program implementation.

Lowenkamp, Latessa and Smith (2006) not only emphasize the importance of evaluation research, they empirically demonstrate the importance of program quality. These researchers examined program characteristics and treatment integrity data from 38 community-based residential programs. Findings indicated a fairly strong correlation between treatment integrity (as measured with the CPAI) and reduction in recidivism. Specifically, they found that two-thirds of the programs were classified as "unsatisfactory, and that the better a program scored, the lower its recidivism rate. This suggests 1) that programs struggle with implementation of effective programming and 2) that program quality does matter.

In reaction essays to the Lowenkamp, Smith and Latessa (2006) study, Welsh (2006) argues that evaluative research must become a formative part of program development, and calls for monitoring of program design, content and delivery. Andrews (2006) acknowledges the barriers programs face in implementing the principles of effective intervention. He therefore

outlines some general "rules" for programs to follow in order to increase likelihood of success in correctional programming. In addition to following the risk, need and responsivity principles, he suggests that program supervisors attend to the relationship and program structuring skills of program staff, including clinical supervision and modeling, reinforcement and monitoring of such skills.

In summary, the final principle of effective intervention, program fidelity, suggests that if a program uses a cognitive-behavioral model and targets criminogenic need factors for higher risk offenders, but does not train staff, provide clinical supervision and provide quality assurance around effective implementation of the model, it is much less likely to be effective in reducing recidivism (Andrews, 2006). Programmatic evaluation tools, such as the CPAI can be used to assist programs to determine whether they are effectively adhering to the risk, need, and responsivity principles as well as and related effective practice and quality assurance issues.

"What Works" for Females

Treatment Philosophies for Female Offenders

There is an abundance of empirical evidence supporting the importance of rehabilitation in corrections (see Gendreau et al., 2000). Furthermore, there exists specific evidence concerning the factors associated with effective correctional interventions (see Gendreau, 1996). Nonetheless, some researchers refute application of the "what works" literature to particular populations. Specifically, two distinct camps exist with differing philosophies regarding the treatment of female offenders (Hubbard and Matthews, 2007, Daigle, Cullen and Wright, 2007). Those that adhere to a *generality approach* argue that based upon both theoretical and empirical grounds, the principles of effective intervention in fact apply to women (see Dowden and

Andrews 1999). Others who rely on a *specificity approach* assert that empirical literature in corrections is based primarily on males; therefore gender-specific approaches to the treatment of girls and women are needed (see Covington and Bloom, 2007, Hannah-Moffat, 2004, Ward and Brown, 2004). Review of the literature also exposes a third group of scholars who are able to successfully blend concepts and philosophies from both perspectives (see Blanchette and Brown, 2006). Since a primary goal of this dissertation is to determine the role gender plays in evidence-based treatment, this section of the dissertation will review these various approaches on the treatment of female offenders.

A Generality Approach

Two of the key arguments from generalists regarding relevancy of the principles of effective intervention to the treatment of female offenders are 1) the gender-neutral nature of the theories that form the basis of effective correctional practices, namely cognitive, behavioral and social learning theories (Bandura, 1979; Spiegler and Guevremont, 2003); and 2) the existence of empirical data regarding application of the principles of effective intervention to female populations (Simourd and Andrews, 1994; Dowden and Andrews, 1999a; Lovins, Lowenkamp, Latessa and Smith, 2007). Both of these arguments will be reviewed in what follows as part of the generalist perspective to the treatment of female offenders.

Theoretical Rationale for a Generalist Perspective

Gendreau and Ross (1979) argued that criminal behavior is learned. This simple insight had important implications for the development of the general responsivity principle, which suggests a primary goal of rehabilitation is to teach offenders prosocial ways to manage their environment. Cognitive theory, behaviorism, and social learning theory are all considered

psychological theories of human behavior; key concepts from these theories have been merged to formulate cognitive behavioral treatment (Speigler and Guevremont, 2010).

Cognitive theory asserts that thinking affects behavior and that thinking can be influenced or changed. Meichenbaum (1977) suggested that individuals have cognitive templates they carry from situation to situation. For example, if one believes that it is wrong to assault a peer, one is likely to also believe that it is wrong to assault a stranger. Cognitions are believed to play a major role in criminal behavior in the form of irrational thinking (Yochelson and Samenow, 1976; Samenow, 1984), internalization of antisocial values (Kohlberg, 1976), and having limited cognitive skills (Ross and Fabiano, 1985). Furthermore, antisocial thinking is identified as a key risk factor for criminal recidivism (Gendreau, Little and Goggin, 1996), suggesting it plays a key role in influencing antisocial behavior. One commonly used therapeutic model grounded in cognitive theory is Rational Emotive Therapy (Ellis and Greiger, 1977). The primary goal of RET is to identify irrational thinking, and then restructure maladaptive thoughts to more rational ways of thinking.

Behaviorism or radical behavioral therapies, conceptualized in the early 1900's by Pavlov, Watson, Skinner and others, rely on two primary mechanisms for modifying behavior: classical and operant conditioning. Classical conditioning, conceptualized by Pavlov (1927), involves the association between a stimulus and a response, and use of the stimulus to modify behavior. In the treatment of offenders, classical conditioning is perhaps most readily seen by way of aversive therapies (Van Voorhis, Braswell and Lester, 2007, Laws and Marshall, 2003). Such techniques involve the overt or covert pairing of an aversive with a behavior one wishes to decelerate. For example, a sex offender's inappropriate fantasies may be paired with mild shocks (an aversive) in order make the target behavior more unpleasant, thereby decreasing it.

Operant conditioning, on the other hand, is a form of direct learning that involves the use of reinforcement to accelerate a target behavior and the use of punishment to decelerate a target behavior (Skinner, 1953). Effective use of reinforcers and punishers within a correctional treatment model is key for increasing prosocial and decreasing antisocial behaviors (Lowenkamp, 2004, Dowden and Andrews, 2004; Lowenkamp, Latessa and Smith, 2006; French and Gendreau, 2006).

Social learning theory generally asserts that the primary mechanism of learning comes from observation (Bandura, 1977). Key factors in the social learning model are the use of role models and the process of modeling behaviors that clients can imitate (Bandura, 1977). Within a correctional context, it is important that staff have prosocial skills and values consistent with the rehabilitative ideal (Gendreau and Andrews, 2001). Andrews and Bonta (2006) outline the therapeutic skills needed to effectively approve and disapprove of behaviors and to model good self-management skills so that program participants can learn by example.

In summary, cognitive, behavioral and social learning theories together make up what is commonly referred to as cognitive-behavioral treatment (Andrews and Bonta, 2010). A principle argument of generalists is that each contributing theory within cognitive-behavioral treatment is a general theory of human behavior, applicable to all people despite gender, ethnicity, culture or age (Andrews and Bonta, 2010). Furthermore, there is substantial evidence that cognitive-behavioral therapy is a superior model for the treatment of offenders (Andrews et al., 1990; Izzo and Ross, 1990; Lipsey, 1992; Dowden and Andrews, 1999a). Hence, there is no theoretical rationale for believing that females require a treatment model different from that found to be most effective at modifying antisocial behavior. Furthermore, while there are competing theories to the treatment of female offenders (e.g. feminist, relational, holistic models), there is a dearth

of empirical data demonstrating that these approaches are more effective at reducing recidivism among female offenders than cognitive-behavioral therapies (Worell, 2001, Blanchette and Brown, 2006).

Empirical Rationale for a Generalist Perspective

Researchers and practitioners in the generality camp rely heavily on what sound empirical studies find to be effective at reducing recidivism. Key empiricists also emphasize the importance of using a theoretical framework to both guide research studies and ground empirical findings (Gendreau and Ross, 1979; Bonta, 2002; Andrews and Bonta, 2010, Andrews et al., 1990). With that said, researchers in the "generalist camp" recognize that females are oftentimes underrepresented in correctional research because they make up a small portion (15%) of the total population of offenders (Greenfeld and Snell, 1999). As such, research specific to females has been conducted to test whether the principles of effective intervention, or aspects related to the principles, apply to women.

Use of Risk Assessment with Female Offenders

There has long been criticism about the application of risk assessments to girls and women. Many feminist scholars question the validity of actuarial tools used to classify women (Funk, 1999; Bloom, 2000; Reisig, Holtfreter and Morash 2006; Holtfretter and Cup, 2007) or express concern about misclassification of females (VanVoorhis and Presser, 2001). Consequently, researchers have explored what risk factors apply to females as well as the validity of risk assessments on women.

Numerous studies have explored whether key factors in predicting risk are similar for males and females. As mentioned previously, the criminogenic needs that are consistently found to be the most powerful at predicting criminal behavior include antisocial attitudes/cognitions,

antisocial associates, and antisocial personality (e.g., anger, impulsivity, risk-seeking, lack of empathy and lack of remorse). Needs with moderate predictive strength include education, employment, dysfunctional family environment, use of leisure time, and substance abuse. Lastly, lower class origins, reduced verbal intelligence, and indicators of personal distress (e.g., depression, anxiety, and self-esteem) are considered needs with limited ability to predict criminal behavior (Andrews, Bonta and Hoge, 1990; Gendreau, Little, and Goggin, 1996).

In a meta-analysis of the correlates of delinquency by gender, Simourd and Andrews (1994) found that lower class origins and personal distress/psychopathology were weak predictors, educational/vocational achievement and family factors were moderate predictors, and antisocial personality, attitudes and associates were the strongest predictors for both males and females. These results were consistent with the extant literature on predictors of risk for offenders (Gendreau, Andrews, Goggin and Chanteloupe, 1992).

Loucks and Zamble (1999) studied male and female offenders in a Canadian federal facility also to explore whether risk factors were the same for both. They found that psychopathy was the strongest predictor for both males and females, and that history of abuse was a weak predictor. In contrast, Hubbard and Pratt (2002) conducted a meta-analysis on factors associated with female delinquency. They were especially interested in examining factors not readily examined in prediction studies (e.g., self-esteem and sexual abuse). Like previous studies on correlates to crime, they found that antisocial peers (ES=.51) and history of antisocial behavior (ES=.48) were the strongest predictors of risk. School relationships (ES=.25), personality (ES=.21), physical or sexual abuse (ES=.21), antisocial attitudes/beliefs (ES=.18) and family relationships (ES=.17) were moderate predictors of risk. Finally, self-image (ES=.13), anxiety (ES=.06), and socio-economic status (ES=.03), were the weakest

predictors. Thus, aside from abuse history, which had a moderate correlation with delinquency, risk factors looked fairly similar between the genders and were consistent with the current literature on predictors of criminality.

With regard to studies on specific risk prediction instruments, Lowenkamp, Holsinger and Latessa (2001) explored the validity of the Level of Service Inventory-Revised (LSI-R) for men and women, as well as the role of childhood abuse in predicting risk. Findings indicated that the LSI-R accurately predicted recidivism for both males and females. Furthermore, in contrast with Hubbard and Pratt (2002), they found that although women were more likely to report abuse, history of physical or sexual abuse failed to predict recidivism, once risk was controlled using the LSI-R. In all, while most predictors look similar between males and females, data appears equivocal regarding the role of abuse history in predicting risk for female offenders.

Smith, Cullen, and Latessa (2009) conducted a recent meta-analysis with 14,737 female offenders on the relationship between the LSI-R and re-offending. Findings from this study using 27 effect sizes yielded an average correlation of .35 with a 95% CI of .34 to .36 for female offenders. Moreover, this correlation was statistically similar to that of the male comparison sample. Overall, a correlation of .35 for women is similar or greater than average LSI-R effect sizes reported in other research with male dominant populations (Gendreau, Little, and Goggin, 1996; Gendreau, Goggin, and Smith, 2002).

While the above studies explore risk factors for females and use of classification tools with women, these next studies specifically examine application of the principles of effective intervention to females. Dowden and Andrews (1999a) found evidence that the risk, need and responsivity principles do in fact apply to women as well as men. In their meta-analysis of 26

studies, they found that programs that primarily targeted higher risk women produced an effect size of .19, while programs targeting lower risk women produced an effect size of -.04. Interestingly, for programs that solely served women (versus mostly female participants) the effect size for programs applying the risk principle rose to .24. In addition, they found substantial differences in effect sizes between programs that targeted primarily criminogenic needs (ES=.26) and those that targeted mainly non-criminogenic needs (ES=.04). Finally, Dowden and Andrews (1999a) found that female programs using a cognitive behavioral approach were more effective at reducing recidivism among participants.

Lovins, Lowenkamp, Latessa and Smith (2007) conducted a primary study on the application of the risk principle to women offenders. They too found support for the risk principle with women. Specifically, high risk women who participated in residential programming had an adjusted probability of re-arrest of .46 relative to .66 for the comparison sample receiving community supervision. To the contrary, low risk women participating in residential programming had a re-arrest probability of .18 relative to just .06 for the control group. Thus, higher risk women who participated in residential treatment showed lower probability of recidivism relative to a comparison group while lower risk women increased in likelihood of re-arrest after exposure to the same treatment.

Gender as a Responsivity Factor

Scholars who embrace the generalist perspective recognize that the needs of females differ from that of males. Generalists recognize that issues such as low self-esteem, mental illness, victimization and poverty are very common among female offenders. These needs, however, tend to be classified by generalists as *specific responsivity* issues rather than criminogenic needs (Andrews and Bonta, 2010). Thus, problems such as lack of adequate

housing or childcare should be addressed, but addressed as a barrier to treatment success rather than as a primary treatment target aimed at reducing recidivism.

Feminists, on the other hand, argue that issues such as victimization or poverty should be addressed as a central program element as it is related directly to criminality among women (Bloom, Owen and Covington, 2003). Thus, a primary difference between the perspectives is that generalists assert that effective correctional programming must differentiate offenders by risk, focus on criminogenic needs, and use a cognitive-behavioral approach, despite participant gender (Gendreau et al., 2004). Conversely, those arguing for specificity feel gender-specific issues are key and should be the focus of treatment for women in the criminal justice system (Covington and Bloom, 2007, Hannah-Moffett, 2004).

A Specificity Approach

Gender-specific Corrections

Feminist scholars argue that the correctional system would look much different if interventions were designed with women in mind, rather than simply applying correctional strategies used for males to females (Chesney-Lind, 2000; Bloom, Owen and Covington, 2003, Salisbury, Van Voorhis and Spiropoulos, 2009). Bloom, Owen and Covington (2003) examined the existing body of knowledge on effective practices with female offenders and developed a set of guiding principles for the management and treatment of criminal justice involved women.

The first guideline offered by Bloom, Owen and Covington (2003) is the acknowledgement that gender is an important consideration in criminal justice, including how men and women's pathways to crime as well as their offending patterns differ. As such, women-centered programming and oversight should be created, and systems should hire staff that are both qualified and motivated to work with women. Second, safety, respect and dignity must be

in place in correctional environments for women; this is necessary for women to respond to behavioral intervention. Third, the role of relationships in women's lives must be a central focus, with regard to policies, practice and programming. Relationships include those with partners, children and community support systems. Forth, the field must offer comprehensive treatment that addresses substance abuse, trauma and mental health. This requires comprehensive staff training in these areas as well as training in cultural sensitivity. Fifth, intervention should incorporate strategies to improve women's socioeconomic status so that they can support themselves and their children. This includes providing educational and vocational training for female offenders and enhancing services that support the economic needs of women. Finally, community supervision and reentry should be a comprehensive system of care, such that correctional programs collaborate with community services to meet the diverse needs of women. These guidelines developed by Bloom, Owen and Covington (2003) can be used to determine whether a correctional systems adheres to a gender-responsive philosophy.

The Pathways Model

It is clear from crime statistics that there are substantial differences between male and female criminal behavior. Perhaps the most obvious difference is the gender gap related to the base rates of offending (Steffensmeier and Allan, 1996). This applies to both overall rates of offending, with men offending at much higher rates than women, as well as males offending at significantly higher rates in each of the major crime categories. There are particular gender discrepancies noted between rates of violent offending (Seffensmeier and Allan, 1996).

Aside from the obvious differential offending patterns between men and women, feminists assert that the road to criminal justice involvement for females is also unique (Daly 1992, 1994, Morash, 2006). Chesney-Lind has long argued that the pathway to criminal justice

involvement for girls differs from that of boys. She asserts that girls' exposure to the criminal justice system is oftentimes a result of running away from an abusive situation, which leads to basic survival strategies involving the commission of crime, as well as longer term depression and drug use (Chesney-Lind 1997, Chesney-Lind and Shelden 2004). Chesney-Lind (2000) argues that status offenses, such as runaway, are more likely to be criminalized for girls versus boys and that that high rate of drug abuse among female offenders is often a result of self-medicating abusive pasts. As such, more girls and women are incarcerated than is necessary, and once incarcerated, most fail to receive services to meet their unique needs.

The pathways model is a key theory of female criminality used to argue for a different course of treatment for women offenders. Daly (1992) analyzed the presentence investigations of 40 women for a federal court to develop a framework for women's pathways to criminal involvement. Her framework pinpointed the role of childhood abuse, domestic violence, drug addiction and economic marginalization in female criminality. Specifically, Daly (1992) identified five unique pathways to illegal behavior:

1. *Street women* were individuals who had escaped abusive childhood homes. They became drug-addicted and then homeless. These women relied on criminal behavior to survive on the streets and contend with the addiction. Their criminal records were extensive and their crimes were related to their drug use, including drug distribution, prostitution and stealing;
2. *Harmed and harming women* were also abused as children, but responded to the abuse by acting out or engaging in aggressive behavior. Many of these women had co-occurring mental illness and alcohol or drug addiction.
3. *Battered women* were involved in the criminal justice system due to retaliation against male abusers. Few had prior criminal involvement;
4. *Drug-connected women* also had a limited criminal record but sold or manufactured drugs as a result of relationships with criminal intimates or family members;
5. The *Other* category includes *economically motivated* women whose crimes involved monetary gain related to greed or contending with poverty. These women, as a group, lacked abuse histories or substance dependence.

More recently, Salisbury and Van Voorhis (2009) examined 313 female probationers to test how three unique pathways to female offending predict recidivism. Information was gathered through gender-specific risk/need assessments and surveys. They found empirical support for the following three pathway models: 1) *childhood victimization path model*--women are victimized as children, contributing to past and current mental illness and substance abuse; 2) *relational path model*--dysfunctional adult relationships lead to victimization, which is in turn linked to reduced self-efficacy, current mental illness and substance abuse; and 3) *social and human capital path model*--deficits in education, family support and functional relationships contribute to decreased self-efficacy as well as financial and employment problems. Each pathway was linked to future incarceration in the probationers studied. This study represents one of few that quantitatively examine the pathways model to female criminality and the authors argue that findings largely support the qualitative work conducted in this area.

Brennan, Breitenbach, and Dieterich (2010) use a gender specific risk/need assessment developed by Van Voorhis and colleagues (described below) to also examine the notion of gendered pathways to crime. They examined the characteristics of 718 incarcerated females nearing release and found support for unique pathways to female offending. They identified eight distinct gendered pathways to crime (which can be summarized into four general pathways to crime): *1) Normal Women*--lower risk, higher need, drug involved females; 2) *Marginalized "Socialized" Offenders*--addicted, vocational/educational deficits, poverty, homelessness, social isolation; 3) *Serious, Chronic, Violent Women Offenders*--high risk, high need, chronic offending, mental health issues, antisocial personality, aggression, victimization history; and 4) *Lifelong Victimization*--high stress, childhood victimization extending to adulthood, addiction, retaliative violence. They conclude that these pathways encompass both gender neutral

predictors of crime (i.e. antisocial attitude/personality, aggression, substance abuse) as well as factors typically categorized as gender specific factors (i.e., victimization history, economic marginalization, mental illness). A distinct body of research on the predictors of offending among female offenders exists, and will be examined next.

Predictors of Offending and the Use of Risk Assessment with Female Offenders

In accordance with the notion that females have unique pathways to criminal behavior, feminists argue that interventions designed for male offenders should not be capriciously applied to females. Consequently, many feminists reject the applicability of risk/need or classification instruments to female offenders (Funk, 1999, Bloom, 2000; Reisig, Holtfreter and Morash 2006, Holtfretter and Cupp, 2007). Of primary concern is the notation that traditional risk/need tools rely on male-centered theories and fail to account for a number of factors that uniquely lead to female offending.

Funk (1999) studied the predictors of risk in a sample of male and female juvenile probationers. She found that risk factors were not the same for girls and boys. In support of the pathways theory, she noted a relationship between abuse, gender and delinquency. Specifically, she found that the relationship between abuse and offending was much stronger for females (r=.41) than males (r=.03), a finding later supported by Hubbard and Pratt (2002). Similarly, Belknap and Holsinger (2006) found support for the pathways model in their study of self-reported risk factors of incarcerated girls and boys. They found that girls had significantly higher rates of victimization, which they were more likely to associate with their offending. Likewise, girls were more likely to report mental health symptoms, including self-injurious behaviors. Belknap and Holsinger (2006) emphasized the importance of broadening the definition of risk factors by incorporating gender-specific needs.

In a study using national-level longitudinal survey data, Daigle, Cullen and Wright (2007) also examined predictors of male and female delinquency. They identified factors from traditional criminological theories, as well as life-course and feminist theories. Findings showed that while some predictors were similar between males and females, others were unique to each gender. With regard to the feminist variables, they found that depression was uniquely predictive of female offending, yet they also found that victimization predicted delinquency for both males and females. Furthermore, they found that forced sex predicted delinquency, but in the opposite direction. Overall however, they concluded that because there are variables that are uniquely predictive of male and female offending, general theories of crime may not be able to fully address the predictors of delinquency.

Holtfretter and colleagues have conducted a series of studies on predictability of a popular risk/need assessment, namely the Level of Service Inventory-Revised (LSI-R), with women. Holtfretter, Reisig and Morash (2004) found that economic marginalization (i.e. poverty) was predictive of recidivism, and that the LSI-R was no longer predictive for women once poverty was controlled. They concluded that the LSI-R does not adequately account for the economic marginality common among female offenders.

Reisig, Holtfretter and Morash (2006) assessed the predictability of the LSI-R across female pathways to crime. Guided by Daly's pathways model, they subdivided women into groups that appeared to follow "gendered pathways" to crime and those who were more economically motivated to engage in criminal behavior. Given that women in this later group had an absence of significant victimization or economic marginalization, they were designated as more closely resembling typical male criminality (Reisig et al., 2006). Findings suggested that the LSI-R was in fact predictive for women they classified as "economically motivated" to

engage in crime (r =.29). However, the LSI-R was not correlated with recidivism for women who followed gendered pathways to crime (r = -.14). They therefore concluded while use of the LSI-R may be appropriate for certain types of female offenders, it misclassifies women who are economically marginalized or have gendered contexts to their offending patterns.

Finally, Holtfretter and Cup (2007) conducted a recent literature review of 11 studies on the predictability of the LSI-R for women. While the findings made it impossible to conclude the LSI-R had no predictive value for women, they suggested that this tool is better able to predict more extreme forms of recidivism, which are relatively rare for women. They cautioned that correctional administrators who rely too heavily on this tool with women may be misguided.

Gender-Specific Programming

Over the past decade, several gender-specific curricula have emerged that are designed for the treatment of females. Some have been designed specifically for offenders, while others are designed to meet needs common among female offenders. *Girls Circle* was developed in 1996 as a support group aimed at helping girls to make healthy life decisions. This curriculum was not developed for offenders in particular; in fact it tends to be used for at-risk populations in multiple settings such as schools, churches, counseling centers, activity centers, as well as the juvenile justice system. The curriculum model is based upon relational-cultural theory, resiliency practices and skills training, and program themes include friendship, being a girl, body image, individuality, peer relationships, identity, diversity and paths to the future (Irvine and Roa, 2010). This curriculum is rated as "promising" by the Office of Juvenile Justice and Delinquency Prevention (OJJDP). Studies have been conducted on this program in 2005, 2007 and 2010 that used quasi-experimental, single sample pre-posttest designs examining measures such as self-worth, self-reliance, body image, and communication, as well as program

satisfaction. The most recent study conducted by Irvine and Roa (2010) found a significant increase in body image, communication with adults, and self-efficacy. The program also had high satisfaction rates. One of the criticisms, from the generalist perspective, is that this program fails to target criminogenic needs linked to recidivism among girls. Studies conducted have not used a comparison group to determine whether the program is effective in reducing recidivism to address its application to offenders. Gender-specific supporters, however, would likely argue that recidivism is too narrow a definition of correctional program efficacy.

Helping Women Recover is another curriculum developed by Covington in 1999 and revised in 2008, to treat females with substance addiction. This program was developed using theories of addiction, trauma and women's psychological development. It contains four modules: self, relationship, sexuality and spirituality (Covington, 2002). In 2003, Covington developed *Beyond Trauma,* a curriculum designed to address female trauma which also integrates the trauma-substance abuse connection. Like Helping Women Recover, this curriculum uses cognitive-behavioral, expressive arts and relational theory to address trauma needs. *Voices: A program of Self-Discovery and Empowerment for Girls* was also developed, by Covington (2004), using the same theoretical underpinnings, as a gender-responsive alternative for youthful offenders.

A study was conducted in 2008 examining the use of Covington's adult curricula to treat women with co-occurring substance abuse and mental health disorders (Covington, Burke, Keaton and Norcott, 2008). Women who successfully completed these programs were assessed as several time points using scales that measured trauma symptomology, substance use and depression. They used a single-sample pre-posttest design with 192 women and found significant improvement in the anxiety, depression, sleep disturbance and dissociation subscales.

They found that 99 percent of successful completers remained drug, alcohol and crime free during the course of the program. Just 29 females completed a 6 month follow-up and 72 percent of those reported continued abstinence (Covington et al, 2008). A limitation of this study is the lack of a comparison group and long-term follow-up to more effectively test the programs' effectiveness in treating substance abuse and criminality.

Seeking Safety is another curriculum designed to address substance use and trauma, but more specifically, individuals dually diagnosed with a Substance Use Disorder and Post-Traumatic Stress Disorder. Unlike the other programs, this curriculum was not developed exclusively for women, and the primary treatment model upon which it was developed is cognitive-behavioral therapy. Nonetheless, given the target needs addressed, this curriculum is often used among criminal justice services which target females. The emphasis in this curriculum is the development of "safe coping skills" (Najavits, 2002). Several studies have been conducted on the efficacy of this curriculum. An outpatient study with 17 females found significant improvement in drug and alcohol use, suicide risk, as well as other mental health, problem solving and general functioning indicators (Najavits, Weiss, Shaw and Muenz, 1998).

Another small study on women in a prison setting found that over half of the women no longer met the criteria for PTSD by the end of the 3-month treatment. While no use was found via urinalysis tests during incarceration, by 3 months after release, 35 percent of the sample were using, and 33 percent were returned to prison. There was no comparison sample in this study (Zlotnick, Najavits and Rohsenow, 2002). However, in a randomized controlled trial comparing *Seeking Safety (SS)* to individual relapse prevention treatment (RPT) and a treatment as usual group (TAU), both SS and RPT found significant reductions in substance use, PTSD and other psychiatric symptoms, while there was no significant change for the TAU group (Hien, Cohen, Litt, Miele, and Capstick, 2002). Improvements were maintained in both the SS and RPT samples at

six months, but not at nine months. Researchers concluded that brief CBT-based interventions can be effective with a difficult to treat population.

Finally, *Moving On* is a gender-responsive cognitive behavioral program developed specifically for female offenders; it has both an adult and juvenile version. This program incorporates motivational interviewing, cognitive behavioral theory and relational theory to target needs specific to female offenders. Gehring, Van Voorhis, and Bell (2010) examined 190 probationers to test the effectiveness of *Moving On, a*nd found evidence to support the effectiveness of this intervention over a matched comparison sample. Successful program completers had significantly lower rates of recidivism (arrest, conviction and incarceration) over the matched comparison cases, with differences between the groups ranging from 10 to 13 percent. This study lends support to both the generalist notion that cognitive behavioral interventions apply to women and well as men, but also to the feminist perspective that females benefit from an intervention designed to target needs specific to women. Findings from both Seeking Safety and Moving On transition nicely to the next section which discusses the researchers that blend notions from both the generalist and gender-specific perspective.

The Generality-Specificity Merger

While there are researchers who clearly fall into the generality or specificity camp, others have found a way to merge key ideas from both (Blanchette and Brown, 2006; Giordano, Deines and Cernkovich, 2006; Hubbard and Matthews, 2007; and Gehring, Van Voorhis and Bell, 2010). One key crossover area relates to correctional program components. Koons, Burrow, Morash, and Bynum (1997) conducted a nation-wide qualitative study surveying expert and offender perceptions of program elements leading to successful outcomes for female offenders. The study found that many of the same program elements described by Andrews and colleagues

were important for effective treatment of females (e.g. warm and dedicated staff, a range of treatment targets, prosocial modeling, structured and individualized programming, skill acquisition, and brokerage). Hence, while feminists argue for gender-specific programming, effective programs for males and females are still likely to share many of the same core characteristics.

In reviewing the six guidelines for gender-specific treatment outlined by Bloom, Owen and Covington (2003), several overlaps are noted with the general principles of effective intervention developed by Andrews and Gendreau. Consistencies include the need for trained and qualified staff, an environment that is respectful and treats offenders with dignity, the need to address education and vocational limitations of offenders, the importance of addressing key responsivity issues such as mental health and victimization where appropriate, and the emphasis on brokerage with the community. Of course, while Bloom et al. (2003) designed these principles to be specific to women, close review suggests that many are not conceptually dissimilar from the principles identified by Andrews, Gendreau, and other generalists. In fact, Blanchette and Brown (2006) argue that interventions based upon the principles of effective intervention leave room for incorporating offender strengths and that generalists have empirically demonstrated the need for empathic and treatment-oriented staff. Likewise, Blanchette and Brown (2006) challenge core criticisms levied by feminist researchers against the responsivity principle by arguing that cognitive behavioral treatment is appropriate for female offenders and in no way "dehumanizes" women or girls. This assertion is consistent with the viewpoint of some leading feminist scholars (see Worell and Remer, 2003).

Blanchette and Brown (2006) provide five core recommendations based upon the extant literature on female offenders. First, they argue against strict application of gender-neutral *or*

gender-specific theories of female criminality; they instead promote "gender-informed" theories that operationalize theoretical constructs to account for gender differences. For example, social control theory, which focuses on the role of marital attachment and employment should incorporate constructs that account for female bonding, such as attachment to children (Blanchette and Brown, 2006). This may involve the development of new theories or the adaptation of those existing. Second, they call for gender-informed assessment models that build upon the work of Andrews, Gendreau and colleagues, such that the principles of effective classification are applied to women. They however argue that female specific issues, such as *relative* risk of reoffending be incorporated into intervention strategies. Third, Blanchette and Brown (2006) argue for the use of actuarial assessment tools that are built from the ground up on the population for which they will be applied, namely, women and girls. They argue that strength and protective factors be incorporated into the scoring criteria, so as to support a strength-based approach to treating female offenders. Forth, they have reformulated the general responsivity principle for females to incorporate "gender-specific" best practice principles. They caution, however, that because there is little empirical data that has informed gender-specific best practices, this new approach should be empirically tested. Finally, they argue for a "reconciliatory rather than an adversarial approach to both research and operational practice concerning girls and women" (Blanchette and Brown, 2006, p. 146).

Giordano, Deines and Cernkovich (2006) also argue that the either/or dichotomy between the generality-specificity debate need not exist. Like Blanchette and Brown, they view these perspectives as not fundamentally incompatible and call for theory integration. Through a qualitative study, Giordano et al. (2006) concluded that not only is female crime driven by economic marginality and disadvantaged neighborhoods, but also through attitudes consistent

with antisocial behavior. Similarly, Salisbury (2008) found in examining four path analyses to female offending that risk/need factors from both the gender-responsive and traditional (i.e. social learning) perspectives contributed to the predictability of female pathways to offending. With regard to risk assessment, some scholars who study women are less adversarial about use of actuarial tools to classify women offenders. Van Voorhis (2005) acknowledges the data that support the predictive validity of risk assessments such as the LSI-R with women; she argues, however, to avoid over-classification[14], risk assessment tools must be validated on female populations. Van Voorhis and Presser (2001) found in a national survey of state prisons, 36 states had not validated their classification system on female inmates. Additionally, tools such as the LSI-R that include both static and dynamic risk factors fail to address the unique needs of women. Such gender-specific factors would serve to identify valuable gender-specific needs as treatment targets and may increase predictability of these instruments.

Accordingly, with funding from the National Institute of Corrections, Van Voorhis and colleagues developed a series of gender responsive instruments designed to increase the predictability of mainstream risk assessments as well as identify needs specific to women in corrections (Van Voorhis, Wright, Salisbury and Bauman, 2010). Their research borrowed from both the generalist perspective that emphasizes the need to assess risk and identify dynamic needs, as well as the feminist perspective that considers women's' unique pathways to criminal behavior.

Van Voorhis et al. (2010) designed both a "trailer" which can be added to an existing risk/need tool, such as the LSI-R, as well as a stand-alone tool for the assessment and classification of female offenders. They utilized probation, prison and pre-release samples from

[14] Over-classification of female offenders occurs when instruments designed for men classify women so that they have a higher level of supervision or custody level than is necessary (Hardyman and Van Voorhis, 2004).

four different states to develop the gender responsive tools. They found that both gender- neutral (e.g. criminal history, education, employment and antisocial associates) and gender responsive factors (e.g. depression/anxiety, psychotic symptoms, parental stress, and victimization) were predictive[15].

Interestingly, Van Voorhis et al. (2010) found that the "big four" factors touted as most predictive of recidivism—criminal history, antisocial attitudes, antisocial associates and antisocial personality (Andrews and Bonta, 2010), did not predict as robustly for women. Instead, education, employment and financial problems (typically labeled as moderate predictors) were especially potent gender-neutral predictors for women. Furthermore, when cumulative scales were developed, Van Voorhis et al. (2010) found that both gender-neutral and gender responsive scales were predictive of recidivism, with r values that ranged from .16 to .36 for the gender neutral scale and values ranging from .27 to .34 for the gender responsive scale. However, when gender neutral and responsive factors were combined, r values increased in 7 of the 8 models to a high of .38. Hence, both gender-neutral and gender responsive factors were shown to be important in predicting of female offending. Like Funk (1999), Van Voorhis et al. (2010) concluded that predictability can be improved when gender-responsive factors are incorporated into risk/need tools.

Summary

The philosophies for managing and treating female offenders vary widely. Generalists argue that well supported theories of criminal behavior, i.e. social learning and behavioral theories are gender-neutral and apply equally to males and females. While differences between male and female offenders are recognized by these scholars, they do not support the notion that

[15] Significant variables and strength of the relationship varied by the sample studied (probation, prison or pre-release) as well as the geographic location.

the principles of effective classification and treatment are not applicable to female offenders. On the other hand, feminist scholars argue that women's pathways to criminal behavior differ substantially from that of men, and this pathway is strongly influenced by the disenfranchisement of women in society. They argue that male-driven theories and interventions should not be applied to women. Finally, there are scholars who subscribe to core components of both the generality and specificity philosophies, recognizing the empirical support for application of evidence-based principles to females, as well as the need to further study feminist contentions and create gender-responsive interventions that address the unique needs of women.

Chapter 2 provided an overview of the literature on effective program characteristics. The "what works" component examined the overall support for rehabilitation in corrections. The "principles of effective intervention" section reviewed the literature on those principles associated with effective rehabilitation of offenders. Finally, the "what works with females" segment examined the perspectives of both the "generality" and "specificity" camps. Together, this review provides background for the current study which will examine effective program characteristics, and what of these are similar between women and men, and what, if any, differ between the genders.

CHAPTER 3
METHODOLOGY

Introduction

This section of the dissertation will focus on what methods were used for data collection and data analysis. Included in this section are the following: 1) a review of the research questions; 2) a description of the type of facilities under review in both Ohio and Pennsylvania; 3) a description of study participants as well as the method used for matching treatment and comparison cases in both Ohio and Pennsylvania; 4) the data collection procedures for individual and program level data; 5) a description of key measures used in the study; and 6) study design and analysis techniques.

Two outcome studies conducted by the University of Cincinnati, Center for Criminal Justice Research are the source of the data used for this dissertation. These studies examine a sample of offenders participating in a residential community correctional facility in Ohio or Pennsylvania between 2006 and 2007[16]. One was conducted for the Ohio Department of Rehabilitation and Correction, and the other for the Pennsylvania Department of Corrections. While the process for collecting the program-level data was similar for both studies, there were differences in the type of offender data available and way in which individual-level data were collected. As such, the methods of collecting offender data as well as sample descriptions will be reported separately for Ohio and Pennsylvania, but the process for program level data collection will apply to both states.

[16] 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 (Latessa, Lovins and Smith, 2010).

Research Questions

As outlined in Chapter 1, this dissertation is designed to explore differences in effective treatment characteristics by gender. In other words, does the effectiveness of a specific program approach, such as the program's therapeutic model, vary based upon the gender of the population being served? In answering this question, program indicators of treatment effectiveness will be examined separately for male and female participants. Hence, program characteristics that are generally effective at reducing recidivism will be identified (as evident by the fact that it predicted outcome for men and women) as well as those that might be unique to one gender. Thus, the research questions being asked in this dissertation include the following:

1. What program characteristics are important for *both* male and female offenders?
2. What program characteristics are *more* important for male offenders?
3. What program characteristics are *only* important for male offenders?
4. What program characteristics are *more* important for female offenders?
5. What program characteristics are *only* important for female offenders?
6. What program characteristics are important for *neither* adult male nor female offenders?

Ohio Sample

The Ohio data incorporate two treatment samples: offenders sentenced to one of the state's Community Based Correctional Facilities (CBCFs), and offenders placed in an Ohio halfway house (HWH). This study also uses two comparison groups: 1) parolee/Post Release Control (PRC)[17] offenders released from an Ohio institution but not exposed to either HWH or

[17] Post Release Control (PRC) is a period of community supervision conducted by the Adult Parole Authority that follows an offender's release from prison. Offenders must have been sentenced to prison for a crime occurring on or after July1, 1996 to be placed on PRC. Offenders sentenced to prison for crimes occurring prior to this date would be released on parole. July 1, 1996 marks the date when Senate Bill 2 or the "truth-in-sentencing" law was enacted. The Ohio Revised Code mandates PRC for some offenders, whereas it is discretionary for others. Furthermore,

CBCF intervention; and 2) offenders placed on state Intensive Supervision Probation (ISP), which was used as a comparison group for probationers in both treatment samples (Latessa et al., 2010). As part of the outcome study, a site visit was conducted with each facility under examination. The treatment sample for each Ohio program consisted of offenders admitted to the program six months prior to and six months following the date of the site visit. Hence, the sample size is different for each program based upon the number of offenders admitted to the facility within this year timeframe. What follows is a description of the CBCF and HWH sample along with a description of the comparison samples and process used for matching treatment and comparison cases. Finally, a description of the offender-level variables and the procedures used for collecting these data in Ohio is provided.

CBCF Sample Description

Ohio's Community Based Correctional Facilities (CBCFs) are residential programs designed to serve adult felony probationers as a last alternative to prison. Ohio's CBCFs were first opened in the late 1970s as a response to prison overcrowding. These facilities allowed for local sanctioning of lower level felony offenders. The operation of a CBCF involves a partnership between state and local governments. These facilities are funded primarily through the Ohio Department of Rehabilitation and Correction (ODRC), but are overseen by a local facility governing board.

CBCFs provide comprehensive programming aimed at meeting multiple offender needs, such as substance abuse, criminal attitude, family issues, anger management, education and employment needs, and emotional wellness. The programs also emphasize effective reentry and restitution to the local community. The CBCF programs range in size to accommodate roughly

---

offenders that commit a 1st degree felony or sex offense are not eligible for PRC. Community supervision under PRC generally functions similar to that of parole supervision.

50 to 200 offenders. The CBCF offenders in the study sample participated in one of 20 Ohio CBCF programs in operation in 2006 (Latessa et al., 2010)[18]. Half of these facilities served females while the others exclusively treated males.

HWH Sample Description

Ohio halfway houses are community-based residential programs designed to serve adult offenders released from state prisons, referred by the Courts of Common Pleas, or sanctioned due to a violation of community supervision. Hence, Ohio's halfway houses serve a wide array of offenders, typically consisting of parolees, offenders on post-release control, individuals released from an institution on transitional control status, and probationers. The Ohio halfway houses provide an array of services to assist offenders in the reentry process. Common services include employment readiness and job placement, educational programming, and drug/alcohol treatment. Some also provide specialized treatment, such as sex offender treatment or programming for offenders with mental health issues.

Unlike CBCFs, which are minimum-security locked facilities, HWHs tend to be staff-secure facilities. The per diem cost to house an offender in a HWH is less than a CBCF, due in part to facility security and services offered by these programs. The HWH sample of offenders in the original study participated in one of 44 Ohio HWH programs in operation in 2006 (Latessa et al., 2010). Sixteen of the 44 facilities served female offenders.

Ohio Comparison Group and Matching Process

Comparison cases in Ohio consisted of either offenders placed on parole/Post Release Control (PRC) or offenders placed on state Intensive Supervision Probation (ISP). Since CBCFs

[18] Of the 20 CBCFs in the original study, 8 programs treated both males and females at the same site, and were therefore divided by sex for the purposes of examining gender differences, for a total of 28 CBCF programs.

are designed to serve probationers, CBCF treatment cases were matched to ISP offenders not receiving either CBCF or HWH intervention. Since HWH's serve a variety of offender types, different comparison samples were used for the HWH cases: HWH parolees/PRC offenders were matched to parolees/PRC offenders not exposed to HWH placement within the study timeframe; HWH probationers were matched to offenders placed on ISP. ODRC provided the list of offenders participating in CBCF and HWH programs within the sampling timeframe as well as the list of prospective parolee and ISP comparison cases. Duplicate offenders were identified in both the treatment and comparison groups. Whatever intervention the offender was admitted to *first* marked their designated group. For example, if an offender received both CBCF and ISP intervention within the sampling timeframe, s/he was kept in whichever group had the first admission date.

Ohio treatment and comparison cases were matched on the following factors: *gender* (male/female), *race* (White/non-White), *sex offender status* (sex offender/non-sex offender), *county* (large, medium and small)[19] and *risk* (low, moderate, and high)[20] (Latessa et al., 2010). With regard to the matching process, the values for the variables for the treatment case were stored and then all matched comparison cases were selected with one randomly pulled and marked as the comparison case. The matching process resulted in a one-for-one match between

[19] It would have been preferable to match treatment and comparison cases by county of supervision, so that the treatment case was supervised at the same county as that of the comparison case, controlling for differences in how county courts process and supervise offenders. However, when too many treatment cases had to be dropped due to no county match, the decision was made to collapse the county variable by size so that treatment cases supervised in large, medium and small counties were matched to comparison cases from similar sized counties. Of note, *county of supervision* was coded according to location of the facility, rather than by what county the offender would be supervised in upon discharge from the facility. Hence, offenders placed in a facility located in a small county were then matched to a comparison case from the same sized county. The alternative would have been to match offenders by *county of conviction*. When data were retrospectively compared to ascertain the difference had cases been matched by this variable (collapsed by size), few differences emerged suggesting that offenders convicted in a small county were likely placed in a facility in the same sized county. Furthermore, matching by *county of conviction* would not have necessarily controlled for community supervision differences for those cases not supervised in the same county where convicted.

[20] A detailed description of the risk measure used for matching is described below in the *Ohio Measures* section of this chapter.

treatment and comparison cases, using the identified matching variables. Those cases that could not be matched were dropped from the sample.

Table 3.1 provides a breakdown of the Ohio programs. This table lists the facilities that participated in this study, with their respective sample size. The total sample includes the treatment and matched comparison cases for all program participants. A total of 19,270 offenders were included in the sample. Table 3.1 also lists the sample of offenders that successfully completed the programs. This sample size is 12,250 cases[21]. The average successful termination, which helps to identify what portion of the participants enrolled in the treatment program were included in the final sample of successful completers was 65 percent, with a range among programs of 13 to 97 percent.

For the purpose of examining gender differences with regard to effective treatment characteristics, co-ed programs that served both males and females were divided into two separate samples so that separate treatment effects could be calculated. Hence, in the original study, there were 20 CBCFs and 44 HWHs; once the co-ed programs were divided so that all programs were comprised of one gender, the resulting program sample size was 28 CBCFs and 51 HWH programs, totaling 79 Ohio programs. Twenty-five of these programs (or 32%) served females.

[21] Programs ranged in number of participants, so they were weighted by sample size. A common method of doing so is weighting each effect size estimate by (N-3) where N equals the number of cases in each program (Latessa, , Lovins, Smith, and Makarios, 2010). Hence, the resulting sample size for the all participant sample in Ohio was 19033; the resulting sample size for successful completers in Ohio was 12,013. .

Table 3.1: Ohio Program Demographics/Descriptives

Program	Treatment Sample	Matched Comparison Sample	Total Sample	Successful Completer Total Sample	Successful Termination Rate[1]	Serves Males	Serves Females	CBCF	HWH
Alternatives (MALE)	401	401	802	456	57%	X			X
Alternatives (FEMALE)	23	23	46	36	78%		X		X
Alvis House Alum Creek	242	242	484	228	46%	X			X
Alvis House Breslin	37	37	74	38	47%		X		X
Alvis House Cope (MALE)*	46	46	92	18	25%	X			X
Alvis House Dunning	67	67	134	80	58%		X		X
Alvis House Ohiolink (MALE)	31	31	62	32	52%	X			X
Alvis H--OhioLink (FEMALE)	16	16	32	24	75%		X		X
Alvis House Price Hall	87	87	174	100	55%	X			X
Alvis House Veterans	69	69	138	90	63%	X			X
ARCA	79	79	158	90	57%		X		X
Booth House/Salvation Army	69	69	138	68	46%	X			X
CATS Female RTP	61	61	122	106	89%		X		X
CATS Male RTP	124	124	248	144	54%	X			X
CATS Therapeutic Community	72	72	144	120	80%	X			X
CCA RTC I (MALE)	25	25	50	34	68%	X			X
CCA RTC I (FEMALE)	48	48	96	74	77%		X		X
CCA RTC II	145	145	290	226	77%	X			X
Cincinnati VOA Drug/Alcohol	173	173	346	78	22%	X			X
Cincinnati VOA SAMI	38	38	76	12	13%	X			X
Cincinnati VOA Sex Offender Tx	76	76	152	58	37%	X			X
Community Transition Center	161	161	322	226	70%	X			X
CompDrug	266	266	532	232	42%	X			X
Courage House	20	20	40	24	64%		X		X
Crossroads	135	135	270	162	60%	X			X
CTCC Canton	192	192	384	196	50%	X			X
Dayton VOA	218	218	436	120	27%	X			X
Diversified	140	140	280	136	48%	X			X
EOCC Female	39	39	78	76	97%		X	X	
EOCC Male	100	100	200	174	88%	X		X	
Franklin (MALE)	329	329	658	498	74%	X		X	
Franklin (FEMALE)	80	80	160	122	76%		X	X	
Fresh Start (MALE)	166	166	332	216	65%	X			X
Fresh Start (FEMALE)	15	15	30	12	40%		X		X

Table 3.1 Continued: Ohio Program Demographics/Descriptives

Program	Treatment Sample	Matched Comparison Sample	Total Sample	Successful Completer Total Sample	Successful Termination Rate[2]	Serves Males	Serves Females	CBCF	HWH
Harbor Light—Corrections (MALE)	370	370	740	342	46%	X			X
Harbor Light—Corrections (FEMALE)	28	28	56	36	64%		X		X
Harbor Light--Drug/Alcohol	74	74	148	128	89%	X			X
Licking-Muskingum	105	105	210	154	71%	X		X	
Lorain-Medina (MALE)	99	99	198	166	84%	X		X	
Lorain-Medina (FEMALE)	38	38	76	60	79%		X	X	
Lucas (MALE)	155	155	310	236	76%	X		X	
Lucas (FEMALE)	42	42	84	74	88%		X	X	
Mahoning	160	160	320	276	87%	X		X	
Mansfield VOA	102	102	204	70	34%	X			X
MonDay (MALE)	198	198	396	338	85%	X		X	
MonDay (FEMALE)	110	110	220	178	81%		X	X	
NEOCAP (MALE)	233	233	466	298	64%	X		X	
NEOCAP (FEMALE)	60	60	120	108	90%		X	X	
Northwest CCC	105	105	210	154	75%	X		X	
Nova House (FEMALE)*	9	9	18	14	78%		X		X
Oriana CCTC (MALE)	251	251	502	264	53%	X			X
Oriana CCTC (FEMALE)	23	23	46	34	74%		X		X
Oriana Cliff Skeen	121	121	242	160	67%		X	X	
Oriana Crossweah	107	107	214	170	80%	X		X	
Oriana RCC	103	103	206	146	69%		X		X
Oriana RIP	272	272	544	264	48%	X			X
Oriana SHARP	40	40	80	46	58%	X			X
Oriana Summit	226	226	452	298	62%	X		X	
Oriana TMRC	297	297	594	326	56%	X			X
Pathfinder (MALE)	116	116	232	108	47%	X			X
Pathfinder (FEMALE)	51	51	102	62	61%		X		X
River City (MALE)	233	233	466	370	79%	X		X	
River City (FEMALE)	89	89	178	156	88%		X	X	
SEPTA	112	112	224	164	69%	X		X	
SOS	130	130	260	140	56%	X			X
Spencer House	11	11	22	18	85%	X			X
STAR	102	102	204	152	77%	X		X	
STARK (MALE)	178	178	356	308	87%	X		X	

Table 3.1 Continued: Ohio Program Demographics/Descriptives

Program	Treatment Sample	Matched Comparison Sample	Total Sample	Successful Completer Total Sample	Successful Termination Rate[1]	Serves Males	Serves Females	CBCF	HWH
STARK (FEMALE)	46	46	92	84	91%		X	X	
Talbert House Beekman	135	135	270	136	48%	X			X
Talbert House CCC	208	208	416	374	90%	X		X	
Talbert House Cornerstone	76	76	152	112	74%	X			X
Talbert House Pathways	86	86	172	112	71%		X		X
Talbert House Springrove	234	234	468	260	55%	X			X
Talbert House Turtle Creek	166	166	332	238	70%	X			X
Toledo VOA	255	255	510	266	52%	X			X
West Central	178	178	356	282	78%	X		X	
WORTH (MALE)	129	129	258	192	74%	X		X	
WORTH (FEMALE)	42	42	84	70	83%		X	X	
ALL FACILITIES	**9635**	**9635**	**19270**	**12250**	**65%**	**54**	**25**	**28**	**51**

[1]The successful termination rate was derived from original ODRC data that includes matched and unmatched treatment cases
*Represents programs where both genders were present in the original sample, but one gender was dropped as there were too few cases to calculate the predicted probability of re-offense.

Ohio Measures

Procedures for Ohio Offender-Level Data Collection

Independent Variables

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), which is ODRC's main inmate database. These data included demographic characteristics, the current offense, offense history, county of supervision, identified needs, services delivered, termination type, and employment (Latessa et al., 2010). All offender background data were provided by ODRC.

Since individual-level data for the various samples included in the study were derived from different ODRC data sources, common data had to be identified across all sources. Demographic data available for analyses include age, race, gender, and marital status. *Age* was coded as actual age in years; *race* was coded as White or non-White; and *marital status* was coded as married or single/not married.

Criminal history and current offense information includes prior incarcerations, prior convictions, offense type, offense level, sex offense and county of supervision. The variable *prior incarcerations* was coded in three ways: 1) number of prior incarcerations; 2) as a dichotomous variable with zero representing the absence of priors, and one representing the presence of priors, and 3) as a categorical variable with zero representing no priors, one representing one prior, and two representing more than one prior incarceration. *Prior supervision* and *sex offense* was simply coded as a dichotomous variable, with zero representing no and one representing yes. *Current offense type* was coded using the following categories:

1=violent crime/person; 2=sex; 3=drug; 4=property; 5=traffic/DUI; 6=other. *Current offense level* was coded as 1=Felony 1; 2=Felony 2; 3=Felony 3; 4=Felony 4; and 5=Felony 5 or Misdemeanor level offense. Finally, *county of supervision* was coded by collapsing Ohio's 88 counties into the following categories: 1=Large: population above 600,000; 2= Medium: population 250,000 to 600,000; and 3=Small: population below 250,000.

Few offender need variables were consistently available across datasets. All need variables were coded as a dichotomous variable with 1 indicating the need is present and 0 indicating the need is absent. Available need data included *substance abuse problem* (drug or alcohol), *current employment problem*, and *current emotional problem*. The manner in which the need data were measured varied from one data source to the next. In the parole database, need assessment information in the substance abuse, personal/emotional, and employment domain was used to ascertain whether the need was present or absent. On a four point Likert scale, the upper two "moderate to significant need" scales were coded as yes. To the contrary, the CCIS database codes drug and alcohol history and referral (which was collapsed for a substance abuse need variable), whether the offender was employed at arrest or referred to employment intervention, and whether counseling was needed. These CCIS variables were used to code substance abuse, employment and emotional need[22].

Table 3.2 provides descriptive statistics for the *all participants* of Ohio programs based upon group membership (treatment or comparison group) and gender (male/female). Recall that cases were matched on gender and race, so virtually no differences exist between the treatment and comparison groups for these variables. Not surprisingly, the majority of participants were male (86.2%); with regard to race, most participants were White (56.6%). Treatment participants were significantly more likely to be single (89.2%) relative to the comparison

[22] Some common offender background variables were not consistently available across datasets, such as education.

Table 3.2: Descriptive Statistics for *All Participants* in OH by Group Membership and Gender

Variable	Treatment Group	Comparison Group	Females	Males
	% (N)	% (N)	% (N)	% (N)
Gender				
Male	86.2 (8302)	86.2 (8302)	--	--
Female	13.8 (1333)	13.8 (1333)	--	--
Race[c]				
White	56.6 (5456)	56.6 (5456)	69.5 (1854)	54.6 (9058)
Non-white	43.4 (4179)	43.4 (4179)	30.5 (812)	45.4 (7546)
Marital Status[ac]				
Married	10.8 (1034)	14.2 (1322)	14.5 (378)	12.1 (1978)
Single/not married	89.2 (8577)	85.8 (7973)	85.5 (2227)	87.9 (14323)
Age Category[ac]				
16 to 23	28.5 (2747)	25.6 (2470)	19.3 (515)	28.3 (4702)
24 to 30	23.8 (2295)	26.1 (2516)	24.5 (653)	25.0 (4158)
31 to 39	24.8 (2386)	24.8 (2393)	31.1 (829)	23.8 (3950)
40+	22.9 (2207)	23.4 (2256)	25.1 (669)	22.8 (3794)
Mean Age[c]	33.1	33.3	33.9	33.1
SD	10.0	10.4	9.3	10.3

[a]Significant difference at the .001 level between groups

[c]Significant difference at the .001 level between genders

sample where 85.8 percent of offenders were not married. There was no significant difference between the treatment and comparison sample in terms of average age (33.1 and 33.3 respectively). With regard to differences by gender, females were significantly more likely to be White (69.5% versus 54.6%) and married (14.5% versus 12.5%). Most female offenders were older, falling into the 31 to 39 age category (31.1%), whereas the bulk of males fell into the 16 to 23 year old category (28.3%).

Table 3.3 provides descriptive statistics for *successful completers* of the Ohio programs. Since cases were matched on gender and race, no differences exist between the treatment and comparison groups for these variables. Within this sample, most completers were White males. Treatment participants were still more likely to be single (87.8% versus 84.5% in the comparison group). There was no significant difference between the treatment and comparison sample in terms of average age (33.6 and 33.3 respectively). Regarding gender differences, like with the full sample, females were significantly more likely to be White (69.8% versus 58.3%) and to be married (14.8% versus 13.1%). Females also tended to again be older than the males (34 versus 33 years old on average).

Table 3.4 presents the criminal history and need factors for Ohio participants in the full sample. Data are again broken down by group membership and gender. Offenders were matched on *county of supervision*, and on a risk measure that incorporates several of the variables presented in this table. With regard to prior criminal history, 30.2 percent of treatment participants had 2 or more prior incarcerations, which is significantly higher than the comparison sample (24.4%). However, a significantly larger proportion of the comparison sample had a previous conviction (57.2% versus 41.1%). With regard to the current offense level, the treatment sample had a higher proportion of felony 1 and 2 offenses. Most offenders' instant

Table 3.3: Descriptive Statistics for *Successful Completers* in OH by Group Membership and Gender

Variable	Treatment Group	Comparison Group	Females	Males
	% (N)	% (N)	% (N)	% (N)
Gender				
Male	83.9 (5137)	83.9 (5137)	--	--
Female	16.1 (988)	16.1 (988)	--	--
Race[b]				
White	60.2 (3687)	60.2 (3687)	69.8 (1380)	58.3 (5994)
Non-white	39.8 (2438)	39.8 (2438)	30.2 (596)	41.7 (4280)
Marital Status[ab]				
Married	12.2 (746)	14.6 (868)	14.8 (287)	13.1 (1327)
Single/not married	87.8 (5364)	85.4 (5071)	85.2 (1646)	86.9 (8789)
Age Category[c]				
16 to 23	24.9 (1523)	25.3 (1551)	17.7 (349)	26.5 (2725)
24 to 30	24.8 (1518)	25.9 (1584)	24.1 (476)	25.6 (2626)
31 to 39	25.3 (1550)	24.0 (1470)	30.9 (611)	23.4(2409)
40+	25.0 (1534)	24.8 (1520)	27.3 (540)	24.5 (2514)
Mean Age[c]	33.6	33.3	34.3	33.3
SD	10.1	10.6	9.3	10.5

[a]Significant difference at the .001 level between groups

[b]Significant difference at the .05 level between genders

[c]Significant difference at the .001 level between genders

offense was drug related, followed by violent/person offenses. Few offenders were convicted most recently on a sex offense (3.8%). Offenders were most likely to come from large counties (42.8%). Finally, more treatment cases were identified as having a substance abuse problem (90.6% versus 79.9%) and an employment problem (53.1% versus 42.4%); yet fewer treatment cases were classified with an emotional problem (31.7 versus 35.3).

Table 3.4: Descriptive Statistics for Criminal History and Need Factors for *All Participants* in OH

Variable	Treatment Group % (N)	Comparison Group % (N)	Females % (N)	Males % (N)
Prior Incarcerations[ac]				
0 to 1	69.8 (6721)	75.8 (7302)	85.3 (2274)	70.8 (11749)
2 or more	30.2 (2914)	24.2 (2333)	14.7 (392)	29.2 (4855)
Previous Conviction[ac]				
No	58.9 (5677)	42.8 (4119)	67.9 (1809)	48.1 (7987)
Yes	41.1 (3957)	57.2 (5516)	32.1 (857)	51.9 (8616)
Offense Level [ac]				
Felony 1	8.9 (859)	7.6 (728)	3.7 (98)	9.0 (1489)
Felony 2	14.4 (1386)	13.8 (1326)	9.1 (242)	14.9 (2470)
Felony 3	21.7 (2091)	22.2 (2136)	16.8 (447)	22.8 (3780)
Felony 4	22.6 (2177)	22.2 (2138)	20.4 (545)	22.7 (3770)
Felony 5/M	32.4 (3122)	34.3 (3307)	50.0 (1334)	30.7 (5095)
Offense Category[ac]				
Violent/person	29.2 (2811)	28.6 (2759)	21.0 (560)	30.2 (5010)
Sex	3.8 (364)	3.8 (364)	0.8 (20)	4.3 (708)
Drugs	32.1 (3096)	28.2 (2719)	39.3 (1048)	28.7 (4767)
Property	21.4 (2061)	25.9 (2498)	30.0 (799)	22.6 (3760)
Traffic/DUI	2.8 (269)	1.3 (128)	1.6 (43)	2.1 (354)
Other	10.7 (1034)	12.1 (1167)	7.4 (196)	12.1 (2005)
County of Supervision[c]				
Small	20.1 (1934)	20.1 (1934)	15.9 (424)	20.7 (3444)
Medium	37.1 (3578)	37.1 (3578)	38.8 (1034)	36.9 (6122)
Large	42.8 (4123)	42.8 (4123)	45.3 (1208)	42.4 (7038)
Substance Abuse Prob.[ac]				
No	9.4 (903)	20.1 (1932)	11.8 (314)	15.2 (2521)
Yes	90.6 (8732)	79.9 (7703)	88.2 (2352)	84.8 (14083)
Employment Problem[ac]				
No	46.9 (4516)	57.6 (5548)	47.2 (1259)	53.0 (8805)
Yes	53.1 (5119)	42.4 (4087)	52.8 (1407)	47.0 (7799)
Emotional Problem[ac]				
No	68.3 (6578)	64.7 (6232)	49.4 (1317)	69.2 (11493)
Yes	31.7 (3057)	35.3 (3403)	50.6 (1349)	30.8 (5111)

[a]Significant difference at the .001 level between groups

[b]Significant difference at the .05 level between genders

[c]Significant difference at the .001 level between genders

Table 3.4 also examines gender differences related to criminal history in the full sample. Male offenders had a significantly higher rate of both prior incarcerations (29.2 versus 14.7 with two or more prior incarcerations) and prior convictions than females (51.9 versus 32.1 respectively). Half of the females in the sample had a felony 5 or Misdemeanor as the instant offense versus just 30.7 percent of males. Females were also more likely to be convicted of drug and property offenses, and less likely to have a violent offense. Females were significantly more likely to be supervised in large versus small counties. Finally, with regard to needs, females were significantly more likely to be identified with substance abuse (88.25 versus 84.8%), employment (52.8% versus 47%) and emotional problems (50.6% versus 30.8%), all the need areas that were examined.

Table 3.5 examines the same variables for the *successful completer* sample. When comparing the treatment and comparison samples, like in the full sample, treatment participants had significantly more prior incarcerations, but were less likely to have a previous conviction. Offense level and category also looked similar to the full sample; just over one third of offenders had a Felony 5 or Misdemeanor level offense, with fewer than 20 percent of offenders with Felony 1 or 2 level offenses. Most offenders' instant offense was drug related, followed by violent/person offenses. Regarding county of supervision, offenders were evenly split with 38.6 percent supervised in medium and 38.9 percent in large counties. Of note, a smaller percentage of offenders in the successful completer sample came from large counties relative to the full sample. With regard to need areas, more treatment cases were identified as having a substance abuse problem (91.2% versus 80.4%) and an employment problem (42.2% versus 37.1%); like the full sample, fewer treatment cases were classified with an emotional problem (30.6% versus 33.3%). In examining criminal history and need factors by gender among successful completers,

Table 3.5: Descriptive Statistics for Criminal History and Need Factors for *Successful Completers* in OH

Variable	Treatment Group % (N)	Comparison Group % (N)	Females % (N)	Males % (N)
Prior Incarcerations[ac]				
0 to 1	75.1 (4600)	80.6 (4936)	86.6 (1712)	76.2 (7824)
2 or more	24.9 (1525)	19.4 (1189)	13.4 (264)	23.8 (2450)
Previous Conviction[ac]				
No	61.9 (3792)	46.9 (2870)	69.0 (1364)	51.6 (5298)
Yes	38.1 (2333)	53.1 (3255)	31.0 (612)	48.4 (4768)
Offense Level [c]				
Felony 1	7.5 (462)	7.1 (434)	3.7 (74)	8.0 (822)
Felony 2	12.4 (762)	12.9 (789)	9.4 (185)	13.3 (1366)
Felony 3	21.1 (1295)	21.1 (1295)	16.7 (330)	22.0 (2260)
Felony 4	24.4 (1497)	22.6 (1384)	20.8 (411)	24.0 (2470)
Felony 5/M	34.4 (2109)	36.3 (2223)	49.4 (976)	32.7 (3356)
Offense Category[ac]				
Violent/person	26.5 (1621)	29.1 (1782)	21.3 (421)	29.0 (2982)
Sex	3.2 (193)	3.2 (193)	.5 (10)	3.7 (376)
Drugs	35.1 (2147)	30.5 (1869)	39.1 (773)	31.6 (3243)
Property	20.6 (1263)	24.0 (1473)	30.4 (601)	20.8 (3135)
Traffic/DUI	3.6 (222)	1.5 (93)	1.8 (36)	2.7 (279)
Other	11.1 (679)	11.7 (715)	6.8 (135)	12.3 (1259)
County of Supervision[c]				
Small	22.6 (1383)	22.6 (1383)	17.2 (340)	23.6 (2426)
Medium	38.5 (2358)	38.5 (2358)	40.0 (790)	38.2 (3926)
Large	38.9 (2384)	38.9 (2384)	42.8 (846)	38.2 (3922)
Substance Abuse Prob.[ac]				
No	8.8 (538)	19.6 (1203)	11.9 (236)	14.6 (1505)
Yes	91.2 (5587)	80.4 (4922)	88.1 (1740)	85.4 (8769)
Employment Problem[ac]				
No	57.7 (3537)	62.9 (3852)	51.8 (1023)	62.0 (6366)
Yes	42.3 (2588)	37.1 (2273)	48.2 (953)	38.0 (3908)
Emotional Problem[ac]				
No	69.4 (4250)	66.7 (4086)	50.9 (1006)	71.3 (7330)
Yes	30.6 (1875)	33.3 (2039)	49.1 (970)	28.7 (2944)

[a]Significant difference at the .001 level between groups

[b]Significant difference at the .05 level between genders

[c]Significant difference at the .001 level between genders

results are also similar to that of the full population. Females had significantly lower rates of criminal justice involvement and committed less serious and less violent crime. However, women in the successful completer sample still were identified with significantly higher needs in the areas of substance abuse, employment and emotional problems.

Risk Measure

Since a common method for assessing risk was not used in the state of Ohio at the time data were collected for the study, a risk measure had to be developed for purposes of matching treatment and comparison cases by risk[23]. Theoretically relevant variables were selected and cross-tabulations and Chi-square analyses were conducted to determine which variables were significantly related to outcome. Table 3.6 lists the variables that create the risk scale; included are having two or more prior incarcerations, having a previous conviction, having a lower level felony (3-5) or misdemeanor offense, engaging in a property offense as the instant offense, having a substance abuse problem, having an employment problem[24], and being age 40 or below[25]. Variables were weighted so that stronger factors were given more weight; the difference in the percentage of any new conviction based upon the presence or absence of a risk factor served as the weight. Separate weights were created for males and females in the sample (see Table 3.6). The factors were then added together to create an overall risk score for each gender. The male risk score ranges in value from 0 to 64.5, while the female risk score ranges from 0 to 43.1.

[23] Data used to develop the risk measure included successful and unsuccessful program completers from the treatment sample.
[24] See the description of substance abuse and employment problem above under *independent variables.*
[25] Of note, in developing the risk scale, variables were examined separately for males and females to identify those most strongly correlated with recidivism by gender. However, based on the available offender-level variables, the same factors were predictive of recidivism for males and females. Hence, while the weights and cut-offs differ by gender, the factors that predict risk are the same in the risk measure used for this study.

Table 3.6: Risk Assessment Factors and Weights

Factor	Male Weight	Female Weight
Prior Incarcerations		
0 to 1	0	0
2 or More	12.4	11.7
Previous Conviction		
No	0	0
Yes	7.2	4.4
Offense Level		
Felony 1 or 2	0	0
Felony 3 - 5/M	7.6	7.4
Offense Category		
Non-Property Offense	0	0
Property Offense	10.3	3.1
Substance Abuse Problem		
No	0	0
Yes	6.2	6
Employment Problem		
No	0	0
Yes	10.9	4.8
Age Category		
Above 40	0	0
40 or Below	9.9	5.7
Possible Score Range	0-64.5	0-43.1

Once the risk scale was calculated, a visual inspection between the risk score and re-conviction was conducted to develop appropriate cutoff scores for risk levels. This resulted in three groups for both the male and female risk scales: *low*, *moderate*, and *high*. The recidivism rates for these three groups for both the males and females are reported in Table 3.7. As expected, recidivism rates increase as offenders' classification levels increase. Also, consistent with the literature on female offenders, rates of recidivism for females were substantially lower than that of the males (VanVoorhis and Brown, 1996). The correlation for the male risk score and any conviction for this study sample was .23; the correlation for the female risk score and any conviction was .20.

Table 3.8 shows the Ohio risk classification by gender and group membership. Since risk category was uses as a matching variable, no differences exist for risk category between the treatment and comparison samples. For both males and females, the majority of offenders fell in the moderate risk category; overall, 70 percent of the sample was moderate risk. The average risk score in both the male and female sample was significantly higher in the treatment group.[26] It should also be noted that the average risk score for females was more than 10 points lower than for males.

Outcome Data

Recidivism data for both the experimental and comparison groups were collected by University of Cincinnati researchers via the Ohio Law Enforcement Gateway (OHLEG) system. ODRC had access approved for a select group of UC researchers by the Ohio Attorney General's office in order to access offender files. Collection of the recidivism data began in April 2009 and

[26] While the large sample size contributed to a significant difference in average risk score, there was limited substantive difference with the treatment group being less than 1 point higher than the comparison group for both the male and female sample.

Table 3.7: Recidivism Rates by Risk Categories and Gender*

Risk Category	Recidivism Rate
Male[a]	
Low (0-16)	18.7
Moderate (17-41)	36.6
High (42+)	53.4
Pearson's r[b]	0.23
Female[a]	
Low (0-18)	11.0
Moderate (19-31)	20.4
High (32+)	35.1
Pearson's r[b]	0.20

*Recidivism is measured by any new incarceration

[a]Significant difference at the .001 level

[b]Significant difference at the .01 level

Table 3.8: Risk Classification for Ohio Participants by Group Membership

Variable	Treatment Group % (N)	Comparison Group % (N)
Male Risk Categories		
Low (0-16)	7.1 (589)	7.1 (589)
Moderate (17-41)	70.7 (5871)	70.7 (5871)
High (42+)	22.2 (1841)	22.2 (1841)
Female Risk Categories		
Low (0-18)	25.9 (345)	25.9 (345)
Moderate (19-31)	65.6 (874)	65.6 (874)
High (32+)	8.6 (114)	8.6 (114)
Overall Risk Categories		
Low	9.7 (934)	9.7 (934)
Moderate	70.0 (6746)	70.0 (6746)
High	20.3 (1955)	20.3 (1955)
Average risk scores		
Males[a]		
Mean (N)	33.9 (8302)	33.1 (8302)
S.D.	11.1	11.9
Females[b]		
Mean (N)	23.0 (1333)	22.2 (1333)
S.D.	6.7	7.0

[a]Significant difference at the .001 level

[b]Significant difference at the .01 level

ended in September 2009. Recidivism data collection occurred in two phases: 1) locating and printing offense records for the identified treatment and matched comparison cases from OHLEG; and 2) entering data from the offense records into a database. All researchers were trained on both accessing records from OHLEG and coding the recidivism data.

Before accessing criminal records from OHLEG, data coders were provided with a list of offender names, social security numbers, dates of birth, gender, and follow-up dates for the treatment and comparison cases. Follow-up dates for recidivism collection was individualized for each offender, depending upon his or her termination date from a program (for treatment cases), or admission date to parole or ISP (for comparison cases). A two-year follow-up timeframe was used. The lists used to collect the OHLEG data were categorized by program and sample. Coders were instructed to match cases from OHLEG on at least two of the three key identifiers (name, date of birth and social security number). Offender criminal records were printed and data were coded from these records.

Data collected from the OHLEG records included: 1) misdemeanor conviction, 2) date of first misdemeanor conviction, 3) type of misdemeanor conviction (most serious)/citation number, 4) felony conviction, 5) date of first felony conviction, 6) type of felony conviction (most serious)/citation number, 7) probation/parole violation, 8) date of probation/parole violation, 9) probation/parole violation citation number; 10) sex offense conviction, and 11) any arrest[27]. These data were coded directly into a secure database, with a separate database created

[27] More detailed information was not collected on new arrests as concerns were expressed from collateral sources that arrest data coded within OHLEG had limited reliability.

for each of the programs and each of the programs' matched comparison group[28]. ODRC also provided information on which offenders returned to an Ohio prison within the two year follow-up timeframe. However, ODRC was unable to reliably discern between returns for a technical violation or new offense for the ISP population.

While data on several potential recidivism variables were collected, the primary outcome variable that will be used to identify the program characteristics related to recidivism reduction is *any new incarceration.* This variable was selected as it was the only common recidivism measure between the OH and PA data. While it would have been ideal to be able to tease out those offenders returned to ODRC for a technical violation, these data were not available for the Ohio sample. Hence, effect sizes for treatment programs may be lower as offenders could conceivably be returned to ODRC (or PADOC) for failing to complete treatment. Nonetheless, this variable offers consistency between states and measures which offenders stay involved in the criminal justice system, regardless of reason for return or sentence to prison.

Pennsylvania Sample

In Pennsylvania, offenders from the treatment sample also participated in one of two types of community-based residential programs: Community Correctional Centers (CCC) and Community Contract Facilities (CCFs). Much like Ohio HWHs, CCCs are state-operated facilities designed primarily to assist offenders in transitioning back to the community. Like Ohio's CBCFs, Community Contract Facilities are designed to provide more intensive programming; however, CCFs primarily serve offenders released from prison versus CBCFs which tend to house probationers. Hence, unlike the Ohio sample, in Pennsylvania the

[28] Data were organized by program so that quality assurance could easily be performed. Researchers were required to identify which program they selected for both pulling OHLEG cases and coding data. Five to ten percent of cases from each program and matched comparison group were audited to ensure correct coding of the cases.

comparison group consisted exclusively of parolees not exposed to CCC or CCF placement within the study's designated timeframe. What follows is a description of the CCC and CCF sample, as well as the process used for matching treatment and comparison cases. Likewise, a description of the offender-level variables and procedures used for collecting these data in Pennsylvania is provided.

CCC Sample Description

Pennsylvania operates 14 Community Correctional Centers located in urban communities throughout the Commonwealth. Two of these facilities exclusively serve females while the rest serve male offenders. The primary goal for offenders participating in a CCC is to develop linkages with the community, including family and community support as well as employment. Offenders participating in CCCs are required to pay rent, cook their own meals, and meet other financial obligations such as restitution and child support; the goal of this is to mimic the responsibilities offenders will have once fully reintegrated to the community. In terms of programming, these facilities primarily provide employment counseling. They refer out for specialized treatment needs such as mental health or drug/alcohol treatment. Twelve of Pennsylvania's 14 CCCs participated in the study. Two of the twelve served female offenders.

CCF Sample Description

The Pennsylvania Department of Corrections also contracts with 38 Community Contract Facilities (CCFs) to provide a residential option with more specialized treatment services than that offered at a CCC. Hence, typical CCF referrals include offenders that have a higher substance abuse need or offenders from rural areas that are closer to their home communities. Both CCCs and CCFs are ultimately designed to aid offenders in reentering the community following incarceration at a state facility. A total of 42 programs were identified across the 38

contract facilities in the original study[29]. Of the 42 programs, 29 percent (12 programs) served female offenders.

Pennsylvania Comparison Sample and Matching Process

The identified comparison cases in Pennsylvania consisted of parolees not exposed to either CCC or CCF intervention. The parolees were released from a PADOC institution within the same timeframe that the treatment cases were released. Each member of the treatment group, whether from the CCC or CCF sample, was matched to a comparison case on the following measures: *sex* (male/female), *race* (White/non-White); *committing county* (one of Pennsylvania's 67counties); *Level of Service Inventory-Revised category* (Low, Medium, High)[30], and *sex offense* (yes/no). Hence, matching variables were similar to Ohio's matching process; however, in Pennsylvania offenders could be matched county for county versus by county size. Likewise, since the state uses a uniform measure of risk on all offenders (the LSI-R), offenders could be matched using this classification tool (versus in Ohio where a risk scale had to be created). Research staff from the Pennsylvania Department of Corrections conducted the matching. Like in Ohio, the end product was a one-for-one match between the treatment and comparison cases. Table 3.9 provides a breakdown of the Pennsylvania programs. This table lists the facilities that participated in this study, with their respective sample size. The total sample includes the treatment and matched comparison cases for all program participants. A total of 7,738 offenders were included in the sample. Table 3.9 also lists the sample of offenders that successfully completed the programs. The average successful termination rate was 86 percent,

[29] For purposes of the study, some of the larger CCF facilities were subdivided into unique programs, such as substance abuse or dual diagnosis programs. Some of the smaller programs were dropped from the current analyses, once programs were subdivided by gender.
[30] A detailed description of this tool is described in the *Pennsylvania Measures* section of this chapter.

Table 3.9: Pennsylvania Program Demographics/Descriptives

Program	Treatment Sample	Matched Comparison Sample	Total Sample	Successful Completers Total Sample	Successful Termination Rate	Male Sample	Female Sample	CCC	CCF
ADAPPT--DNA	34	34	68	28	41%	X			X
ADAPPT--Group Home (MALE)	200	200	400	356	89%	X			X
ADAPPT—Group Home (FEMALE)	29	29	58	52	89%		X		X
Alle-Kiski Pavilion	148	148	296	258	87%	X			X
Allentown CCC	75	75	150	124	83%	X		X	
Atkins House	12	12	24	18	75%		X		X
Capitol Pavilion/Conew (MALE)	136	136	272	234	86%	X			X
Capitol Pavillion/Conew (FEMALE)	19	19	38	36	95%		X		X
Conewago Place (MALE)	100	100	200	194	97%	X			X
Conewago Place (FEMALE)	11	11	22	22	100%		X		X
Cone. Wern.—Alcohol (MALE)*	27	27	54	42	78%	X			X
Cone. Wern.--Group Home (MALE)*	106	106	212	182	86%	X			X
Cone. Wern.—PennCapp (MALE)*	75	75	150	146	97%	X			X
DRC—Alcohol (FEMALE)*	4	4	8	8	100%		X		X
DRC--Group home (MALE)	78	78	156	98	63%	X			X
DRC—Group home (FEMALE)	8	8	16	12	75%		X		X
DRC—Dual Diagnosis (MALE)*	24	24	48	32	67%	X			X
Eagleville D&A	67	67	134	110	82%	X			X
Erie CCC	99	99	198	190	96%	X		X	
Gateway—Braddock	95	95	190	158	83%	X			X
Gateway—Erie	69	69	138	122	88%	X			X
Gaudenzia—Comm. Ground (MALE)*	12	12	24	24	100%	X			X
Gaudenzia--Concept 90 (MALE)	6	6	12	12	100%	X			X
Gaudenzia—Concept 90 (FEMALE)	7	7	14	12	86%		X		X
Gaudenzia—Erie (MALE)	37	37	74	62	84%	X			X
Gaudenzia—Erie (FEMALE)	28	28	56	46	82%		X		X
Gaudenzia Philly House	33	33	66	50	76%	X			X
Gaudenzia Siena Residential	67	67	134	114	85%	X			X
Gaudenzia Siena HWH	121	121	242	170	70%	X			X
Gaudenzia West Chester (MALE)*	25	25	50	50	100%	X			X
Hannah House	33	33	66	52	79%		X		X
Harrisburg CCC	129	129	258	226	88%	X		X	
Johnstown CCC	81	81	162	148	91%	X		X	
Joseph Coleman--Harmony	162	162	324	248	77%	X			X

Table 3.9 Continued: Pennsylvania Program Demographics/Descriptions

Program	Treatment Sample	Matched Comparison Sample	Total Sample	Successful Completers Total Sample	Successful Termination Rate	Male Sample	Female Sample	CCC	CCF
Joseph Coleman--Tranquility	71	71	142	138	97%	X			X
Keenan House (MALE)	69	69	138	106	77%	X			X
Keenan House (FEMALE)	12	12	24	20	83%		X		X
Kintock--Erie Avenue	247	247	494	372	75%	X			X
Liberty Management	109	109	218	160	73%	X			X
Luzerne	72	72	144	96	67%	X			X
Minsec Broad Street	86	86	172	130	76%	X			X
Minsec Chester	134	134	268	238	89%	X			X
Minsec of Scranton	128	128	256	224	88%	X			X
Minsec York Street	60	60	120	94	78%	X			X
Penn Pavilion	115	115	230	196	85%	X			X
Philadelphia CCC #2	22	22	44	40	91%	X		X	
Philadelphia CCC #3	17	17	34	30	88%		X	X	
Philadelphia CCC #4	28	28	54	50	93%	X		X	
Philadelphia CCC #5	33	33	66	56	85%	X		X	
Pittsburgh CCC	18	18	36	36	100%		X	X	
Renewal, Inc. (MALE)*	245	245	490	423	86%	X			X
Scranton Catholic (MALE)	41	41	82	76	93%	X			X
Scranton Catholic (FEMALE)	6	6	12	12	100%		X		X
Scranton CCC	48	48	96	88	92%	X		X	
Self Help Movement	44	44	88	76	86%	X			X
Sharon CCC	45	45	90	78	87%	X		X	
Tranistional Living Ctr.	20	20	40	40	100%		X		X
York CCC	33	33	66	64	97%	X		X	
Youthbuild/Crispus Attucks	9	9	18	18	100%	X			X
Total	**3869**	**3869**	**7738**	**6497**	**86%**	**45**	**14**	**12**	**47**

*Represents programs where both genders were present in the original sample, but one gender was dropped as there were too few cases to calculate the predicted probability of re-offense.

with a range of 41 to 100 percent. This is substantially higher than Ohio's average completion rate of 65 percent. The sample size for successful completers in PA is 6,497 cases[31].

Like with Ohio, in order to examine effective program characteristics by gender, co-ed programs were divided into two separate samples (male/female) so that separate treatment effects could be calculated. Hence, in the original study, there were 12 CCCs and 42 CCFs; once the co-ed programs were divided the resulting program sample size was 12 CCC and 47 CCF programs. In total, there were 59 Pennsylvania programs identified for the current study. Fourteen of these programs (or 24%) served females.

Pennsylvania Measures

Procedures for Pennsylvania Offender-Level Data Collection

Independent Variables

Offender data were provided by the Pennsylvania Department of Corrections (PADOC). Demographic/criminal history data included: name, date of birth, social security number, sex, race, age at release, marital status, education (highest level completed), reading level, employment status, offense type, and level of offense seriousness. Offender-level data were also provided on the community or institutional placement, including location of current CCC or CCF, community supervision type, supervision level (ranging from low to intensive), time spent in prison, adjustment to institution, referral services, and status of discharge from program and parole. Other variables provided by the PADOC include behavioral indicators related to alcohol/drug use and assaultive behavior. *Age at release* was coded as actual age in years, and then categorized into age ranges; *race* was coded as White or non-White; and *marital status* was

[31] Like with the Ohio sites, PA programs ranged in number of participants, so they were also weighted by sample size. The same method of weighting each effect size estimate by (N-3) was used, where N equals the number of cases in each program. Hence, the resulting sample size for all participants in the PA sample was 7,561. The successful completer weighted sample size in PA was 6,320.

coded as married or not married. Criminal history included *current offense seriousness,* coded as low, medium or high, as well as *sex offender status* and *employed at arrest,* both coded as yes or no. Pennsylvania also identified problem indication variables such as *indication of an alcohol, drug or assaultive behavior problem;* these were all coded as dichotomous variables.

Table 3.10 provides descriptive statistics for *all participants* of Pennsylvania programs based upon group membership and gender. With regard to differences in group membership, since cases were matched on gender and race, no differences exist between the treatment and comparison groups for these variables. The majority of participants were male, with just 5.8 percent of the sample being women. With regard to race, unlike in Ohio, most participants were Non-White (57.6%). However, like in Ohio, treatment participants were more likely to be single (86.6%), which was significantly more than the comparison sample (83.5%). Offenders in the comparison sample were younger than the treatment sample, with the average age for comparison cases at 33.4 versus 35.7 for the treatment group. Likewise, when examined by category, the bulk of the treatment cases were 40 or older while most comparison cases fell in the 24 to 30 year old range. Considering the age-crime curve[32], this slight difference favors the treatment sample in terms of impact on recidivism.

Table 3.10 also examines demographic differences between the male and female samples for all participants. For females, the majority of participants were White (55.8%) while just 41.6 percent of the male sample was White. Females were significantly more likely to be married (16.7%) versus males (14.9%). Finally, females were slightly older with an average age of 36 in comparison to 34.5 for males.

[32] The age-crime curve suggests that the peak age of engagement in criminal behavior is late adolescence/young adulthood (Blumstein, 1995).

Table 3.10: Descriptive Statistics for *All Participants* in PA by Group Membership and Gender

Variable	Treatment Group	Comparison Group	Males	Females
	% (N)	% (N)	% (N)	% (N)
Sex				
Male	94.2 (3645)	94.2 (3645)	--	--
Female	5.8 (224)	5.8 (224)	--	--
Race[c]				
White	42.4 (1642)	42.4 (1642)	41.6 (3034)	55.8 (250)
Non-white	57.6 (2227)	57.6 (2227)	58.4 (4256)	44.2 (198)
Marital Status[a]				
Married	13.4 (518)	16.5 (640)	14.9 (1083)	16.7 (75)
Not Married	86.6 (3351)	83.5 (3229)	85.1 (6207)	83.3 (373)
Age[ac]				
16 to 23	7.1 (274)	14.9 (578)	11.2 (818)	7.6 (34)
24 to 30	28.8 (1115)	33.0 (1278)	31.3 (2285)	24.1 (108)
31 to 39	30.4 (1177)	25.9 (1002)	27.9 (2032)	32.8 (147)
40+	33.7 (1303)	26.1 (1011)	29.6 (2155)	35.5 (159)
Mean Age[ac]	35.7	33.4	34.5	36.0
S.D.	9.5	10.1	9.9	9.2

[a]Significant difference at the .001 level between groups;

[b]Significant difference at the .05 level between groups

[c]Significant difference at the .001 level between genders

[d]Significant difference at the .05 level between genders

Table 3.11 provides descriptive statistics on the same variables for *successful completers* of Pennsylvania programs. Recall that Pennsylvania as a whole had a relatively high successful completion rate across programs (86%), which suggests there should be nominal differences between the full sample and successful completers. Nonetheless, with regard to differences in group membership, 6.1 percent of the successful completer sample was females, which is slightly higher than the full sample, and a higher proportion were White (44% versus 42.4%). Comparison cases in the successful completer sample were significantly more likely to be married (16.3% versus 13.7%). Like in the full sample, offenders in the comparison sample were significantly younger than the treatment sample, with the average age for comparison cases at 33 versus 36 for the treatment group. Table 3.11 also examines gender differences in the successful completer sample. White was still the dominant race for females versus non-White for males. There was still no difference in the proportion of married men versus women, relative to the full sample. Finally, females in the successful completer sample were still significantly older than male participants (35.9 versus 34.6 years old), which was very similar to the full sample findings.

Table 3.12 presents the criminal history and need factors for the full sample in Pennsylvania. Offenders were matched on *sex offender status*, which means no group membership differences for this variable; just 1 percent of offenders in the study had an instant offense classified as a sex offense. With regard to current offense seriousness, treatment cases were more likely than comparisons to have a "high" classification level (17.2% versus 9.6%), and less likely to be classified as "low" (9.1% versus 15.8%); nearly three quarters of both treatment and comparison cases were considered medium level seriousness. Significantly more offenders from the comparison sample had indication of an alcohol problem (63.9% versus

Table 3.11: Descriptive Statistics for *Successful Completers* in PA by Group Membership and Gender

Variable	Treatment Group	Comparison Group	Males	Females
	% (N)	% (N)	% (N)	% (N)
Sex				
Male	93.9 (3050)	93.9 (3050)	--	--
Female	6.1 (198)	6.1 (198)	--	--
Race[c]				
White	44.0 (1428)	44.0 (1428)	41.3 (2628)	57.6 (228)
Non-white	56.0 (1820)	56.0 (1820)	56.9 (3473)	42.4 (168)
Marital Status[b]				
Married	13.7 (445)	16.3 (528)	14.8 (905)	17.2 (68)
Not Married	86.3 (2803)	83.7 (2721)	85.2 (5196)	82.8 (328)
Age[ad]				
16 to 23	6.8 (222)	14.8 (480)	11.0 (669)	8.6 (33)
24 to 30	28.4 (921)	33.6 (1092)	31.4 (1914)	25.0 (99)
31 to 39	30.1 (978)	26.0 (846)	27.9 (1701)	31.1 (123)
40+	34.7 (1127)	25.6 (831)	29.8 (1817)	35.6 (141)
Mean Age[ad]	36.0	33.3	34.6	35.9
S.D.	9.6	10.0	9.9	9.3

[a]Significant difference at the .001 level between groups;

[b]Significant difference at the .05 level between groups

[c]Significant difference at the .001 level between genders

[d]Significant difference at the .05 level between genders

Table 3.12: Descriptive Statistics for Criminal History and Need Factors for *All Participants* in PA by Group Membership and Gender

Variable	Treatment Group	Comparison Group	Males	Females
	% (N)	% (N)	% (N)	% (N)
Current Offense Seriousness[ac]				
Low	9.1 (354)	15.8 (613)	12.3 (896)	15.8 (71)
Medium	73.7 (2851)	74.5 (2883)	74.0 (5395)	75.7 (339)
High	17.2 (664)	9.6 (373)	13.7 (999)	8.5 (38)
Sex Offender[d]				
No	99.0 (3831)	99.0 (3831)	99.0 (7214)	100 (448)
Yes	1.0 (38)	1.0 (38)	1.0 (76)	0.0 (0)
Indication of Alcohol Problem[bc]				
No	39.6 (1533)	36.1 (1397)	37.4 (2727)	45.3 (203)
Yes	60.4 (2336)	63.9 (2472)	62.6 (4563)	54.7 (245)
Indication of Drug Problem				
No	21.6 (835)	20.9 (807)	21.1 (1538)	23.2 (104)
Yes	78.4 (3034)	79.1 (3062)	78.9 (5752)	76.8 (334)
Indication of Assaultive Behavior[c]				
No	34.0 (1317)	33.7 (1302)	32.3 (2353)	59.4 (266)
Yes	66.0 (2552)	66.3 (2567)	67.7 (4937)	40.6 (182)
Employed at Arrest[a]				
No	78.1 (3022)	72.9 (2819)	75.5 (5502)	75.7 (339)
Yes	21.9 (847)	27.1 (1050)	24.5 (1788)	24.3 (109)

[a]Significant difference at the .001 level between groups;

[b]Significant difference at the .05 level between groups

[c]Significant difference at the .001 level between genders

[d]Significant difference at the .05 level between genders

60.4%); however, no significant differences existed between the groups for indication of a drug or assaultive behavior problem. Finally, comparison cases were significantly more likely to be employed at arrest (27.1% versus 21.9%).

Table 3.12 also examines criminal history and need factors for all participants by gender. There were significant differences in current offense seriousness with females being more likely than males to be classified as low (15.8% versus 12.3% respectively) and less likely to be high (8.5 versus 13.7). There were no female sex offenders in the sample. Unlike in Ohio, women in Pennsylvania were significantly less likely to have indication of an alcohol problem than men (54.7% versus 62.6%). They were also less likely than men to have indication of assaultive behavior (40.6% versus 67.7%). There were no gender differences related to drug problem or employed at arrest.

Table 3.13 presents the criminal history and need factors for Pennsylvania *successful completers* by group membership and gender. Due to the high successful completion rate, findings for the successful completer sample related to both group membership and gender were very similar to that of the full sample. The treatment sample had a higher classification level related to offense seriousness. However, this group had a lower indication of alcohol problems and was more likely to be employed at arrest. With regard to gender, females also had less serious offenses, and less indication of alcohol or assaultive behavior problems.

Risk Measure

Unlike Ohio, Pennsylvania conducts a validated risk assessment on all offenders within the PADOC system using the Level of Service Inventory-Revised (LSI-R). The Level of Service Inventory-Revised (Andrews and Bonta, 1995) is a dynamic risk tool that both classifies offenders according to risk for recidivism and identifies offenders' criminogenic needs. This tool

Table 3.13: Descriptive Statistics for Criminal History and Need Factors for *Successful Completers* in PA by Group Membership and Gender

Variable	Treatment Group	Comparison Group	Males	Females
	% (N)	% (N)	% (N)	% (N)
Current Offense Seriousness[a c]				
Low	9.5 (308)	16.4 (533)	12.8 (779)	15.7 (62)
Medium	73.0 (2371)	73.9 (2400)	73.2 (4467)	76.8 (304)
High	17.5 (569)	9.7 (316)	14.0 (855)	7.6 (30)
Sex Offender[d]				
No	99.0 (3215)	99.0 (3215)	98.9 (6035)	100 (396)
Yes	1.0 (33)	1.0 (33)	1.1 (66)	0.0 (0)
Indication of Alcohol Problem[b d]				
No	39.8 (1292)	36.2 (1177)	37.5 (2290)	45.2 (179)
Yes	60.2 (1956)	63.8 (2072)	62.5 (3811)	54.8 (217)
Indication of Drug Problem				
No	22.9 (744)	21.3 (693)	22.0 (1342)	24.0 (95)
Yes	77.1 (2504)	78.7 (2556)	78.0 (4759)	76.0 (301)
Indication of Assaultive Behavior[c]				
No	35.3 (1148)	34.3 (1114)	33.2 (2023)	60.4 (239)
Yes	64.7 (2100)	65.7 (2135)	66.8 (4078)	39.6 (157)
Employed at Arrest[a]				
No	78.3 (2542)	72.2 (2345)	75.2 (4587)	75.8 (300)
Yes	21.7 (706)	27.8 (904)	24.8 (1514)	24.2 (96)

[a]Significant difference at the .001 level between groups;

[b]Significant difference at the .05 level between groups

[c]Significant difference at the .001 level between genders

[d]Significant difference at the .05 level between genders

was developed using theoretically relevant variables, and it has been validated on a variety of offender populations (see Vose et al., 2010). This assessment includes 54 items, which fall into the following 10 domains: Criminal History; Education/Employment; Financial; Family/Marital; Accommodations; Leisure/Recreation; Companions; Alcohol/Drug Problems; Emotional/ Personal; and Attitudes/Orientation. Items in each domain are scored as either *yes* or *no* on a scale of 0 to 3[33]. PADOC has conducted validation studies on the LSI-R, resulting in cutoff scores that are normed on a Pennsylvania population of adult inmates; the most recent validation study resulted in the following classification cutoffs: Low 0-20, Medium 21-28, High 29-54. These classification categories were used to match offenders by risk level.

Finally, Table 3.14 shows the LSI-R risk classification by gender and group membership. Since risk category was uses as a matching variable, no differences exist for risk category between the treatment and comparison samples. For both males and females, the majority of offenders fell in the moderate risk category; males, however, had a lower proportion of low risk offenders while the categories for females were more evenly distributed. Overall, 42.2 percent of the sample was moderate risk, followed by 33.7 percent being high risk, and 24.1 percent classified as low risk. There were no significant differences in the average risk scores between the treatment and comparison groups for either the male or female sample. Unlike the Ohio sample, the female's mean risk assessment score was nearly identical to that of the males.

Outcome Data

In Ohio, outcome data were collected by University of Cincinnati staff; the process for collecting such data is detailed above. In Pennsylvania, outcome data were provided by

[33] Zero to 1 indicates a problem area with need for improvement and 2 or 3 indicating the item poses either no problem is a minor problem. The instrument offers descriptors of how each item is to be scored, based on this continuum.

Table 3.14: LSI-R Classification for PA Programs by Group Membership

Variable	Treatment Group	Comparison Group
	% (N)	% (N)
Male LSI-R Categories		
Low (0-20)	23.7 (863)	23.7 (863)
Moderate (21-28)	42.7 (1557)	42.7 (1557)
High (29-54)	33.6 (1225)	33.6 (1225)
Female LSI-R Categories		
Low (0-20)	31.7 (71)	31.7 (71)
Moderate (21-28)	33.9 (76)	33.9 (76)
High (29-54)	34.4 (77)	34.4 (77)
Overall LSI-R Categories		
Low (0-20)	24.1 (934)	24.1 (934)
Moderate (21-28)	42.2 (1633)	42.2 (1633)
High (29-54)	33.7 (1302)	33.7 (1302)
Average LSI-R scores		
Males		
Mean (N)	25.5 (3645)	25.3 (3645)
S.D.	7.5	6.9
Females		
Mean (N)	25.4 (224)	24.7 (224)
S.D.	8.6	7.9
Total		
Mean (N)	25.5 (3869)	25.3 (3869)
S.D.	7.5	7.0

PADOC. Primary recidivism measures in Pennsylvania differ from that of Ohio. Primary Pennsylvania measures include number of arrests, technical violation, re-incarceration and a combined variable measuring any failure. In Ohio, reason for incarceration could not reliably be determined; this variable includes re-incarceration for either a new crime or technical violation. Pennsylvania was able to disentangle incarceration for a new offense versus a technical violation[34]. However, in order to have a consistent outcome variable between the two states, *any incarceratio*n (for new crime or technical violation) was selected for use as the outcome variable.

Program Level Data Collection

Data Collection Process

For both Ohio and Pennsylvania, a list of all residential community corrections sites to be included in the study was provided by the State. In all, the University of Cincinnati research teams visited 115 programs across the state of Ohio and Pennsylvania (64 Ohio and 51 Pennsylvania programs)[35]. Site visits began in early August 2006, and were concluded by December of 2006. Site visits to the facilities generally occurred weekly within this time frame, and were typically conducted by 3 to 5 researchers. All researchers selected to conduct site visits for this project were trained on the data collection materials, as well as effective practices in corrections. Data were typically collected at a program within one eight-hour day, although the time used to collect data varied depending on the size of the program and availability of

[34] PADOC used only parolees as comparison cases. In Ohio, the difficulty in disaggregating incarceration due to a new offense versus technical violation was primarily related to probation cases, whose outcomes the state had more difficulty tracking.

[35] Programs within facilities or agencies were identified based upon whether there were separate treatment sites, and whether sites offered distinct program models or served separate populations. If a program served both males and females, these were only identified as separate programs if programming between the genders differed significantly. For these programs, separate program data was collected. However, for the purposes of the current study, *all* co-ed programs were subdivided by gender despite similar programming, for the purpose of examining gender differences.

treatment groups for observation. Follow-up phone calls were also used if additional program information was needed following the site visit.

Nine tools were used to collect data at each site: 1) program director interview guide, 2) staff interview guide, 3) participant interview guide, 4) group observation form, 5) file review checklist, 6) curriculum review form, 7) staff member survey, 8) staff attitudinal survey, and 9) program summary form[36]. The interview guides were designed so that researchers could conduct a semi-structured interview aimed at determining how the program operated, particularly with regard to treatment delivery. The group observation form was used to code group facilitation practices for each session observed. Information for each curriculum used by the program was coded on the curriculum review form. Furthermore, 10 open and 10 closed client files were reviewed at each site, primarily to corroborate information from the interviews about assessments conducted, needs being targeted, risk level of offenders in the program and discharge documentation. The surveys were provided to staff prior to the site visit to collect data on staff qualifications, experience, training and attitude toward offender rehabilitation. Finally, a program summary form was completed for each site to produce average ratings across researcher observations using all of the available data collected during the site visit. A database with over 1,000 variables was created using the data from the program summary form. This was the database used for the majority of the analyses to follow.

Site visits were scheduled on a day that key programming could be observed and key staff members were available for interviews. Where this was not possible, a researcher was either sent back to a facility for additional group observation or interviews, or follow-up phone interviews were conducted. At each site, the following individuals were interviewed: a program

[36] As a research project overseen by the University of Cincinnati Institutional Review Board (IRB), consents for participation in the research study were also used.

and/or clinical director; treatment providers including therapists, case managers, group facilitators, intake staff, employment specialists, aftercare specialists, mental health specialists or any staff involved in program delivery; a sample of custody staff and supervisors; quality assurance/accreditation managers; and program participants. Likewise, prior to the site visit, the program director was provided with a checklist of materials for review to help prepare for the data collection process. Materials included treatment manuals, assessments, policy and procedures, written information on reinforcers and sanctions, admission and completion criteria, and any research studies conducted on the program. This information was reviewed during the site visit or copies were provided to research staff.

Program Level Measures

Two instruments were used to develop the program-level data collection tools for this project: the Evidence Based Correctional Program Checklist (CPC) and the Core Correctional Practices section of the Correctional Program Assessment Inventory-2000 (CPAI-2000). These instruments measure the degree with which correctional programs follow the principles of effective intervention[37]. Several studies conducted by the University of Cincinnati, including the original 2002 HWH/CBCF study, were used to develop and validate the indicators on the CPC.[38] These studies generated strong relationships between outcome and the overall scores, domain areas, and individual items (Holsinger, 1999; Lowenkamp and Latessa, 2003, Lowenkamp, 2003; Lowenkamp and Latessa, 2005a; Lowenkamp and Latessa, 2005b). The data collection tools for

[37] The CPC was derived frome the Correctional Program Assessment Inventory developed by Gendreau and Andrews in 1989. The CPC does include a number of items not contained in the CPAI and it omits items that were not found positively correlated with recidivism (University of Cincinnati, 2005).

[38] These studies involved nearly 500 correctional programs of varying types. A large part of this research involved the identification of program characteristics that were significantly related to outcome (University of Cincinnati, 2005).

the current study were designed to expand item definitions on these instruments so that in-depth program data could be collected.

The CPC measures two components of programs: *capacity*, or the degree to which the program has the capability of using evidence-based practices and *content*, the current assessment and treatment practices employed by the program (University of Cincinnati, 2005). Program capacity evaluates 1) program leadership and development, 2) staff characteristics and 3) quality assurance. More specifically, the *program leadership and development* domain of the CPC considers the credentials of the program director as well as the program director's involvement in the development of the program, selection of staff and delivery of services. Program funding and sustainability, as well as the process for piloting new interventions are also considered. The *staff characteristics* section identifies the credentials of the treatment staff, in addition to staff training and support as well as attitudes of the staff related to programming. This domain also identifies the use of clinical supervision by the program. *Quality assurance* is the final CPC capacity area. It measures how strategies employed by a program are monitored for fidelity, including observation of service delivery and surveying client satisfaction with the program. Additional quality assurance items include the collection of recidivism data by the program or undergoing process and/or outcome evaluations, and the results of such assessments.

Program *content* on the CPC scrutinizes offender assessment and treatment practices. *Offender assessment* measures the use of actuarial, standardized risk/need assessments as well as the use of other tools that measure criminogenic needs. Likewise, the CPC determines whether a range of key responsivity factors are assessed using validated tool(s). The assessment section also evaluates the program's has eligibility criteria. The items under the *treatment characteristics* domain, examine the following: 1) whether the primary treatment targets of the

program are criminogenic; 2) if the program model is centered around social learning 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 (University of Cincinnati, 2005).

As indicated previously, over 1000 program-level variables were originally coded from the Ohio and Pennsylvania studies. Those variables deemed both theoretically and empirically relevant based on the research questions were selected for use in the current study. Some variables were selected based on the ongoing RNR versus gender-responsive debate outlined in Chapter 2 (e.g., treatment targets, program model, use of specific cognitive-behavioral strategies, use of risk assessment, assessment of criminal attitude, and application of the risk principle to females). Other variables were chosen that might be linked to key gender-specific theoretical approaches such as relational theory (e.g., staff selection and evaluation, role of leadership in the program, family interventions, staff ratio, co-ed design, and program harmony). Still other variables were chosen simply for exploratory purposes (e.g., program budget, years of operation, and support from the community). In particular, items where there is theoretic opposition were chosen so that an empirical examination of long-argued gender differences in corrections could be examined. Items in four of the five CPC domains were studied: 1) Leadership and development; 2) Staff characteristics; 3) Assessment; and 4) Treatment. Quality assurance measures were *not* included in these analyses as there is limited theoretical rationale to suggest

there are gender differences related to the need for monitoring program fidelity. Likewise, there was limited variation on items in this section for both OH and PA.

Tables 3.15 through 3.18 describe what variables from each of the CPC domains were selected. Table 3.15 describes the 10 leadership variables examined in this study. Table 3.16 lists the 13 variables selected to examine differences in staff characteristics. Table 3.17 describes the 18 assessment variables. Finally, table 3.18 reviews the treatment variables, which includes 57 measures of how treatment or programming is employed. These tables identify how the 98 select variables were originally measured in the OH and PA studies and how these variables were re-coded for the current analyses.

Table 3.15: Program Measures Related to Leadership and Development

Variable	Initial Measurement	Recode
Program Budget	Estimate ranging from 100,000 to 6 million	0 to 2.5 million=1 2.6 to 6 million=2
Years of operation	# of years	0 to 5=1 6 or more=2
Staff to resident ratio	Number of full time staff employed/program capacity	One staff to three residents or less=1 More than one staff per three residents=2
Rating of the support of the program by the community at-large:	# on a scale of 1-10	0 to 6=0 7 to 10=1
Program director's involvement in the training of new staff:	No involvement Training is limited to on-the-job training Formal training provided	No (none or on-the-job only)=0 Yes (formal only)=1
Program director involvement in providing direct services to offenders:	Provides no direct service to offenders Provides direct services only as a back-up when regular staff cannot Provides regular direct services to offenders	No (none or back-up only)=0 Yes (regular direct services)=1

Table 3.15 Continued: Program Measures Related to Leadership and Development

Variable	Measurement	Recode
How harmoniously do staff and management work together?	Major problems Moderate problems Minor problems No problems	Moderate to minor problems=2 No problems=0
Are males and females sharing living space (i.e. occupying living space at the same time)?	No Yes N/A	No=0 Yes=1 No co-ed=99
Are males and females sharing educational services?	No Yes N/A	No=0 Yes=1 No co-ed=99
Are males and females sharing visitation?	No Yes N/A	No=0 Yes=1 No co-ed=99

Table 3.16 Program Measures Related to Staff Characteristics

Variable	Measurement	Recode
What 5 characteristics are most important when hiring new treatment staff:	List Characteristics	Believes treatment works for offenders (Y/N) Experience in a correctional environment (Y/N)
Years of experience working in a correctional treatment setting	List # of years	Less than 2 years=0 Two or more years =1
Education level of staff responsible for treatment delivery	Less than high school High School/GED Associate's degree Bachelor's degree Graduate Degree	70 percent of staff have at an Associate's degree or higher in a helping profession (Y/N)
List training topics for formal initial trainings:	List topics	Training on theory/philosophy of program (Y/N)
Who regularly attends clinical or treatment team meetings?	List positions	Clinical meeting scale: 4 or more meetings/month (Y/N) Time spent reviewing cases (Y/N) Involvement of director (Y/N) Involvement of case manager (Y/N) Involvement of security staff (Y/N) 1 or fewer characteristics=0 2 or more characteristics=1
How many times per month are clinical meetings held	# of meetings per month	
Are offender cases reviewed?	No Yes	

Table 3.16 Continued: Program Measures Related to Staff Characteristics

Variable	Measurement	Recode
What factors are staff evaluated on?	List Factors	Soft skills scale: Creativity (Y/N) Openness to Supervision (Y/N) Team Player (Y/N) 0 to 1=0 2 to 3=1 Hard skills scale: Paperwork (Y/N) Attendance (Y/N) Dress (Y/N) Productivity (Y/N) 0 to 2=0 3 to 4=1 Communication/relationship scale: Crisis/de-escalation (Y/N) Boundaries (Y/N) Negative interaction with clients (Y/N) Firm but Fair (Y/N) Empathic (Y/N) 0 to 2=0 3 to 5=1
Who provides clinical supervision?	Licensed psychologist Licensed social worker or counselor Non-licensed but certified clinical staff Non-licensed staff Other	Clinical supervision provided by a Licensed social worker/counselor (Y/N)

Table 3.16 Continued: Program Measures Related to Staff Characteristics

Variable	Initial Measurement	Recode
Staff value treatment efforts Treatment Staff Security Staff	Rate on a scale of 0-10	Treatment staff rating 0-8 rating 9-10 rating Security staff rating 0-7 rating 8-10 rating
Staff Turnover	Very big problem Problem Minor issue No problem	Problem=1 Minor or no problem=0

Table 3.17: Program Measures Related to Assessment

Variable	Initial Measurement	Recode
Program Assessments	Assessment Name	Risk/Need (Y/N)
	Purpose	LSI-R (Y/N)
	Standardized?	Validated need assessment tools (Y/N)
	Summary Score?	Validated substance abuse tool (Y/N)
	Validated?	Criminal attitude (Y/N)
	Normed?	Validated responsivity tools (Y/N)
	% of offenders that receive the assessment	Biopsychosocial tool (Y/N)
		Mental health assessment (Y/N)
		Assess past abuse (Y/N)
		Number of assessments conducted
		0-3=0
		4 or more=1
Who selects offenders for the program?	Program/Clinical Director	Program Director (Y/N)
	Case manager/therapists	Intake Department (Y/N)
	Licensed Professionals	
	Parole/probation officers	
	Referral source	
	Intake department	
	Other	
Are there written exclusionary criteria that are consistently followed?	No	No=0
	Yes	Yes=1
What characteristic might exclude an offender from the program?	List criteria	Excludes offenders with mental illness (Y/N)

Table 3.17 Continued: Program Measures Related to Assessment

Variable	Initial Measurement	Recode
How are staff trained on the use of assessments?	Trained by watching others conduct assessments Attended formal training on validated instruments Unknown	Staff trained by watching other staff conduct assessments (Y/N)
What is the percentage of low risk offenders in the program?	List %	Below 20%=0 20% or more=1
Are reassessments conducted by the program?	No Yes	No=0 Yes=1

Table 3.18: Program Measures Related to Treatment

Variable	Initial Measurement	Recode
What characteristics/offender issues does the program regularly target?	List characteristics	Tier 1 criminogenic needs scale: Criminal attitude (Y/N) Peer associates (Y/N) Anger/hostility (Y/N) Self-control (Y/N) Interpersonal conflict resolution (Y/N) 0 to 2=0 3 or more=1 Tier 2 criminogenic needs scale: Alcohol or drug problems (Y/N) Structured use of leisure time (Y/N) Vocational achievement (Y/N) Educational achievement (Y/N) Relationship with significant others (Y/N) 0 to 3=0 4 or more=1 Gender specific needs scale: Mental health (Y/N) Self-esteem (Y/N) Childhood abuse/neglect (Y/N) Trauma/PTSD (Y/N) Relationship with children (Y/N) Parenting skills (Y/N) Economic/social needs (Y/N)

Table 3.18 Continued: Program Measures Related to Treatment

Variable	Initial Measurement	Recode
What characteristics/offender issues does the program regularly target? *Continued*	List characteristics	0 to 2=0 3 or more=1 Gender responsive need areas targeted: Drug/Alcohol (Y/N) Vocational achievement (Y/N) Educational achievement (Y/N) Family affection/communication (Y/N) Family problem solving (Y/N) Offender relationship with significant other (Y/N) Low self-esteem (Y/N) Mental health issues (Y/N) Economic/social needs (Y/N) Childhood abuse/neglect (Y/N) Trauma/PTSD (Y/N) Offender relationship with children (Y/N) Parenting skills (Y/N)
Name of groups delivered by the program	List group name	Group Interventions offered: Thinking for a Change (Y/N) Other cog-based group (Y/N) Substance abuse (Y/N) Anger management (Y/N) Employment/vocational education (Y/N) Family (Y/N)

Table 3.18 Continued: Program Measures Related to Treatment

Variable	Initial Measurement	Recode
Name of groups delivered by the program *continued*	List group name	Gender-specific (Y/N) Mental health (Y/N) Life-skills (Y/N) Eclectic (Y/N) Number of groups provided 0 to 6=0 7 or more=1
How are offenders assigned to groups?	All residents attend all groups Residents are assigned based on need Residents choose which groups they attend Residents assigned based on counselor instinct/decision Other	Groups assigned based on need (Y/N)
Are males and females kept in separate treatment groups?	No Yes N/A	Program mixes gender in groups (Y/N)
Is a curriculum used for each group?	No Yes	Program uses structured curricula (Y/N)

Table 3.18 Continued: Program Measures Related to Treatment

Variable	Initial Measurement	Recode
Does the program teach the offenders to plan or rehearse alternatives to problem situations (e.g. role play)?	No Yes	Use of role play (Y/N)
Does the program train offenders to practice new behaviors in increasingly difficult situations (e.g. graduated rehearsal)?	No Yes	Use of graduated rehearsal (Y/N)
What treatment model(s) are used by the program?	List model(s)	Use of a cognitive behavioral treatment model (Y/N)
Excluding aftercare, what is the average length of stay of the program?	List average LOS	Average length of stay Less than 4 months=0 4 or more months=1
Do higher risk offenders get more treatment than low or moderate risk offenders?	No Yes	Variation of treatment by risk level (Y/N)

Table 3.18 Continued: Program Measures Related to Treatment

Variable	Initial Measurement	Recode
Are higher risk offenders separated from lower risk offenders in the program?	No Yes	Offenders separated by risk level (Y/N)
Are any of the following considered when assigned offenders to staff caseloads?	No, offenders assigned to caseloads based on availability/open slot Offenders initially assigned based on caseload size, but mat be switched Level of offender motivation Offender personality Offender learning style Cognitive limitations of offender Gender Age Other	Cases are assigned by caseload size (Y/N)
How well are offenders monitored in the community?	No monitoring Ineffective monitoring Some monitoring Adequate monitoring Highly effective monitoring	Community monitoring while on pass Ineffective to some monitoring=0 Adequate to highly effective monitoring=1
Rate the use of rewards/incentives by the program	Rewards are not used by the program Rewards are used sparingly Moderate use of rewards Liberal use of rewards	None or sparingly=0 Moderate to liberal=1

Table 3.18 Continued: Program Measures Related to Treatment

Variable	Initial Measurement	Recode
What type of incentives and rewards are used by the program?	List type	Use of early release as incentive (Y/N)
How are reinforcers administered?	List administration method	Offenders told why being reinforced (Y/N)
Rate the use of punishers/ consequences by the program?	Punishers/consequences are not used Liberal use of punishers Moderate use of punishers Punishers are used sparingly	Sparingly=0 Moderate to Liberal =1
What type of punishers or consequences are used by the program?	List type	Use of Therapeutic Community strategy as punishment (Y/N) Use of isolation as a punishment (Y/N) Removal of pass as a punishment (Y/N)
How are punishers administered?	List administration method	Punishers individualized (Y/N)
What is the approximate ratio of reinforcers to punishers	Reward ratio Punishment ratio	Reinforcement to punishment ratio: Punishers outweigh reinforcers=0 Reinforcers outweigh punishers=1
How are close relations/family/ friends of the offenders involved with the program	List involvement	Given program overview only (Y/N) Family activities (Y/N) Family treatment intervention (Y/N)
What percentage of families is involved in programming?	List percentage	Less than 40%=0 40% or more=1

Table 3.18 Continued: Program Measures Related to Treatment

Variable	Initial Measurement	Recode
Who provides aftercare for offenders once they complete the program?	List provider type	External provider (Y/N) Internal provider (Y/N)
What services are provided as part of aftercare?	List service(s)	Self-help/AA only (Y/N) Relapse Prevention (Y/N) Case Management (Y/N)
Average Length of aftercare	List length	1 to 3 months=0 4 or more months=1

Analyses

The analyses that follow are reflective of each of the four domains listed above. Ultimately, analyses for this study will explore which treatment measures were predictive of recidivism, and how that differed by gender. The OH and PA original samples included programs that served exclusively males (N=64), exclusively females (N=13) and both males and females (N=38). In order to examine effective program characteristics by gender, those co-ed programs that served both males and females were separated, so that separate effect sizes were calculated for males in the program and for females in the program. Hence, a co-ed program became two separate programs (e.g., WORTH male and WORTH female).

Program characteristics for each co-ed program were duplicated for the male and female samples.[39] In cases where separation by gender resulted in very few cases in a program sample (either male or female), those cases were simply dropped from the sample and data was examined for only the dominate gender group. This occurred in 10 cases, where the sample sizes for a particular gender did not allow for calculation of the predicted probability. Separation of co-ed programs by gender resulted in duplicating 23 programs, and increasing the overall number of programs from 115 to 138.

Data were examined for both the full sample and successful treatment completers only. Both samples were examined to help rectify debate in the field related to whether program effect sizes should be based on all offenders exposed to the intervention, or only those that successful completed treatment. Some argue for the examination of all offenders exposed to an intervention since higher risk offenders who are more likely to recidivate are also are more likely to be

[39] At the time of data collection for OH and PA, a decision was made as to how to collect data on the program. The initial decision to score a co-ed program as a single program was based on the fact that the program structure and programming characteristics appeared similar for male and females residents. Hence, it was deemed appropriate to simply duplicate program characteristics when co-ed programs were separated, for purposes of comparison.

unsuccessfully discharged, which skews results to favor the treatment group (Lowenkamp, 2004; Latessa et al., 2010). Others argue that including the full sample is not an accurate measure of program effectiveness as those unsuccessfully terminated fail to receive the full dosage of the intervention. Thus, both an "all participant" sample and the "successful completer" sub-sample were examined[40].

Calculating Effect Sizes

In order to identify which program characteristics were related to reductions in recidivism, the predicted probability of recidivism was calculated for each program. This process involved running descriptive statistics for each program on the following variables: *gender* (male/female); r*ace* (White/Non-white); a*ge* (age at program discharge); *sex offender status* (yes/no); *total risk score*; and s*ample* (treatment/comparison)[41]. Next, multivariate logistic regression models were calculated, using *any new incarceration* as the dependent variable and the above 6 variables as the independent variables. A predicted probability calculator was used to compute the predicted probability of recidivism for each program. This calculator used means from descriptive statistics and slopes from logistic regression models to calculate the predicted probability of any new incarceration. Probabilities were calculated for both the treatment and matched comparison sample, and then subtracted to identify each program's respective effect size. This process was repeated for the successful completer sample so that each program had an effect size based on the full sample and an effect size based on successful completers only.

[40] Only those comparison cases matched to the successful completers were retained in the comparison samples for the studies.

[41] These six variables were available for cases in both the PA and OH study. Recall that risk score for the PA sample is based on the LSI-R, and for OH is based on the risk score designed for the study.

The decision was made to calculate predicted probabilities for all programs so that a uniform measure of effect size was used in OH and PA[42]. Given that the control variables did not differ substantially from those that treatment and comparison cases were matched on, the predicted probabilities used in this study resembled mean difference used in the original OH study and predicted probabilities (using slightly different control variables) in the original PA study.

Weighting

When effect size estimates derived from programs with different sample sizes are examined, it is vital that estimates are weighted so that larger programs are given greater weight (Rosenthal, 1991; Wolf, 1986). This is based on the assumption that larger sample sizes produce more reliable estimates. One way to do this is to weight effect size estimates by the number of cases in each sample. 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 (27,008) instead of the number of actual programs (138). To address this issue, the current research standardizes the weights so that the mean of the weights is 1 and N equals 138. 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 (Latessa et al., 2010).

In the current study, once effect sizes were calculated for each program, these effect sizes were weighted using standardized weights. Given that there were substantially fewer female programs, and female programs tended to be smaller, separate weights were calculated for the male and female programs. Hence, since the research questions examine what program

[42] In the original studies, predicted probabilities were calculated for PA, but mean differences were used as the effect size in the OH study.

characteristics were important for male programs (comparing all programs serving males), and what characteristics were important for female programs (comparing all programs serving females), it was important to weight separately so that male programs were weighted relative to other male programs, and female programs were weighted relative to other female programs. Weights were also calculated separately for the full sample and successful completer only sample, since weights are based on sample sizes, and sample sizes decreased for the successful completer samples.

Bivariate Relationships and Scales

Once predicted probabilities were calculated and cases were weighted by sample size, analyses were conducted beginning with examination of the bivariate relationship between the program variable of interest and the average program effect size. Variables were coded to be either dichotomous or categorical, and the outcome variable (effect size) was metric. Hence, t-tests were used to examine this relationship for dichotomous variables, and ANOVA for categorical variables.

Pearson's correlation coefficient was used to identify the strength of the relationship between the program trait and outcome. This measure was chosen because it is widely understood and easy to interpret. The correlation coefficient *r* ranges between -1 and 1 whereby negative values indicate an increase in the probability of recidivism, whereas positive values denote a positive program outcome. A value at or near zero indicates that the item has little relationship with outcome.

Once all of the bivariate relationships were examined, aggregate scales were developed. Individual program characteristics that were theoretically important *and* found to have a significant relationship with program effectiveness were included in scales used to measure each

of the four domains listed above (Leadership Characteristics, Staff Characteristics, Treatment and Assessment). Included in the scales were items that did not reach statistical significance, which may have been based on the relatively small sample size (N=138 overall with N=99 for male and N=39 for female programs), but that had theoretical significance. Hence, non-significant items that approached a correlation of .10 *and* were theoretically relevant may have also been included as they provided a substantive contribution to the scale. Correlations were calculated for the scales as well, using Pearson's r. Cronbach's alpha was also examined for each scale to determine the scale's reliability, or the degree to which items measure the same construct.

Summary

Chapter 3 provides a description of the methods used to collect and analyze data for this dissertation. This chapter reviewed in detail a description of the Ohio and Pennsylvania samples, including the process used to match treatment and comparison cases to determine program effect sizes. Furthermore, the procedure used for collecting the offender and program level data was described for both states. Likewise, this chapter provided a detailed description of the offender and program level measures, as well as offender and program descriptive data for each state. Finally, the techniques that will be used to analyze the data were described. The next chapter will provide results based upon the methods of analysis described above.

CHAPTER 4
RESULTS

Introduction

Chapter four of this dissertation will provide a presentation of the findings from the current study. As a reminder, this dissertation poses the following research questions:

What program characteristics are important for *both* adult male and female offenders?
What program characteristics are *more* important for adult male offenders?
What program characteristics are *only* important for adult male offenders?
What program characteristics are *more* important for adult female offenders?
What program characteristics are *only* important for adult female offenders?
What program characteristics are important for *neither* adult male nor female offenders?

The "importance" of each program characteristic is measured based on that item's correlation with an improved effect size. Remember that effect size is measured via incarceration for either a new crime or technical violation. Data are presented separately for males and females so that gender differences can be examined. Data are also presented for both the full sample of all program participants and successful completers only. The total number of offenders in the study is 27,008 in the full sample and 18,747 in the successful completer sample. In all, there were 23,894 males and 3,114 females. Recall also from chapter three that the programs that were originally co-ed were subdivided by gender so that separate effect sizes could be calculated for the male and female samples. For the current study, the overall sample is 138 programs, with 99 serving males and 39 serving females. The unit of analysis for this study is the program.

Chapter 4 will begin with a brief description of overall aggregate differences between male and female programs by state (OH and PA), along with a distribution of effect sizes. Next bivariate analyses will examine each program characteristics' relationship with recidivism. Presentation of bivariate analyses will be broken down into the following program characteristic categories: Leadership/Design, Staff, Assessment Tools and Characteristics, Treatment Targets, Group Interventions, and General Treatment Characteristics. Data from the bivariate analyses will be used to construct scales identifying the relationships between Leadership, Staff, Assessment and Treatment categories and program effectiveness[43]. These overall scales will be calculated separately for males and females. Correlations between program characteristic categories and state will also be examined.

Aggregate Findings by Gender and State

Before examining individual program characteristics, data on the overall treatment effects, as well as the differences in treatment effects for males versus females will be explored. The overall treatment effect for programs when the full sample is examined is -12.8 (SD=12.4). This improves when only successful completers are included (ES=-4.9; SD=14.7), but the average remains negative. This indicates that the average predicted probability of recidivism for the matched comparison samples was lower that the treatment samples, suggesting that overall, those offenders participating in the programs were more likely to be returned to prison. This finding is in part due to the outcome variable of choice. Recall that *any incarceration* was selected as the outcome variable since this was the only common variable used to measure recidivism in the Ohio and Pennsylvania studies. As noted previously, *any incarceration*

[43] Recall from Chapter 3 that these categories are based on the Evidence-based Correctional Program Checklist (CPC), an assessment designed to measure program fidelity that is correlated with a program's ability to reduce participant recidivism (Lowenkamp and Latessa, 2005).

includes both new crimes and technical violations. Incorporating technical violations increases the negative effects for treatment programs since failed treatment often leads to probation or parole revocation. This bias is particularly problematic for the "All Participant" sample, which includes unsuccessful discharges.

Although the average treatment effect was negative for programs, there was some variation, albeit limited for the full sample, in effectiveness of programs at reducing the rate of incarceration. Figure 4.1 shows the effect size distribution for the full sample of participants. This graph indicates that 84 percent of programs had a negative effect size[44]. Figure 4.2 examines successful completers only. As expected, fewer programs (59%) showed negative treatment effects when only successful completers were considered.

Figure 4.3 examines average treatment effects by gender. This figure suggests that female programs as a whole tended to produce higher treatment effects, significantly so in the full sample (t= -1.895; p= .06)[45]. However, different results are seen when treatment effects are split by gender and state (see Figure 4.4). In Ohio, males in programs produced significantly higher treatment effects in the successful completer sample than female participants (t= 2.418; p=.02). Yet, in Pennsylvania, females' effect sizes were 10 to 12 percentage points higher than males, indicating a better treatment effect among programs serving women. Hence, the overall higher treatment effects for females in the study are clearly driven by the Pennsylvania data.

Figure 4.5 examines the distribution of effect sizes by gender. In the full sample, 87 percent of male programs had a negative effect size versus 76 percent of female programs.

[44] Outliers in the distribution were examined and all had small sample sizes (fewer than 25 cases); these programs were therefore not omitted from the analyses.

[45] Note that $p \leq .10$ was used as the significance level since the number of female programs was limited (n=39). Also note in the tables that follow that sample sizes do not always equal 99 for male programs and 39 for female programs due to missing/not applicable data as well as weighting. In particular, cases in the successful completer sample were dropped due to low weights, which are based on a small number of successful program completers.

Figure 4.1: Effect Size Distribution for the Full Sample

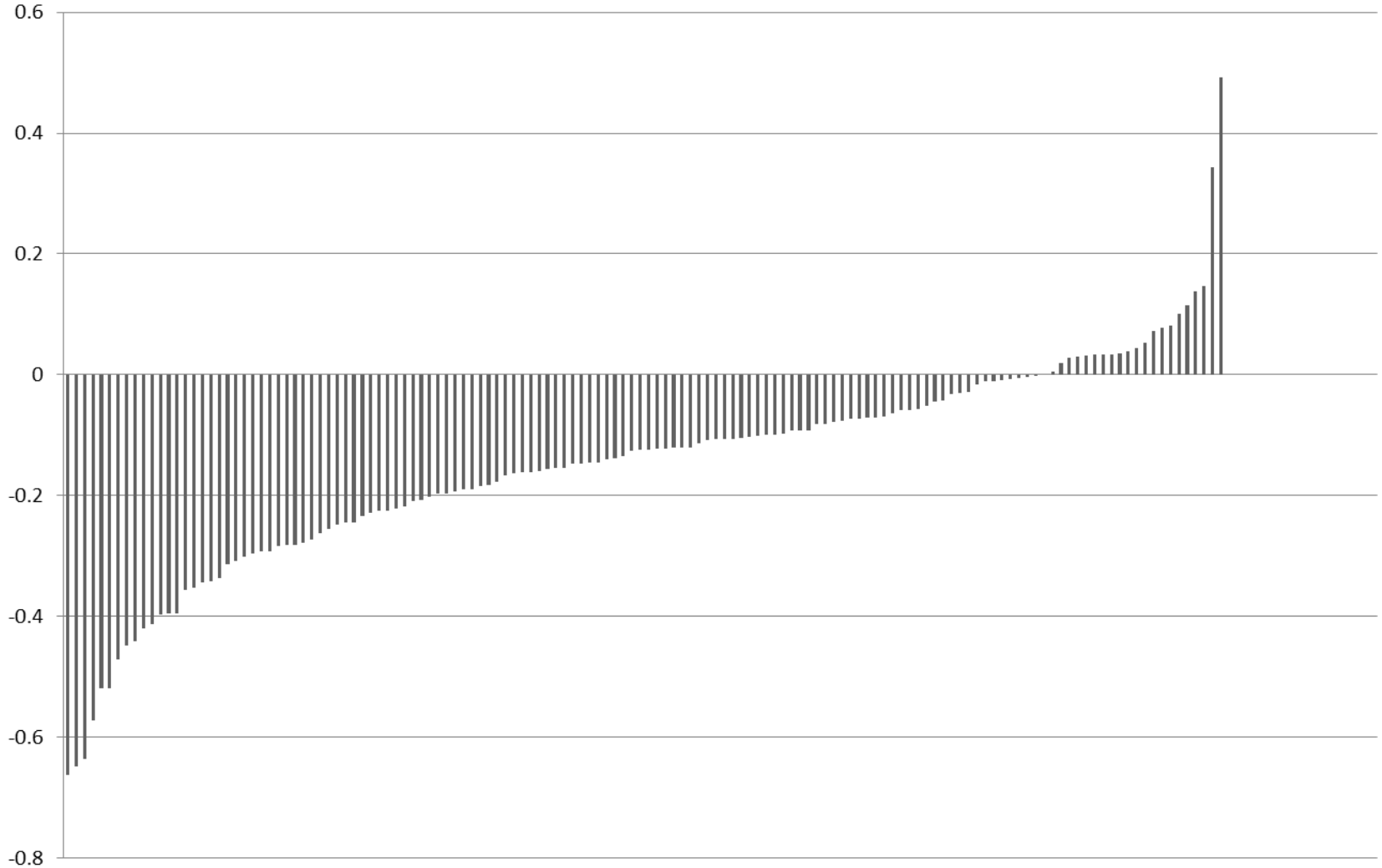


Figure 4.2: Effect Size Distribution for the Successful Completer Sample

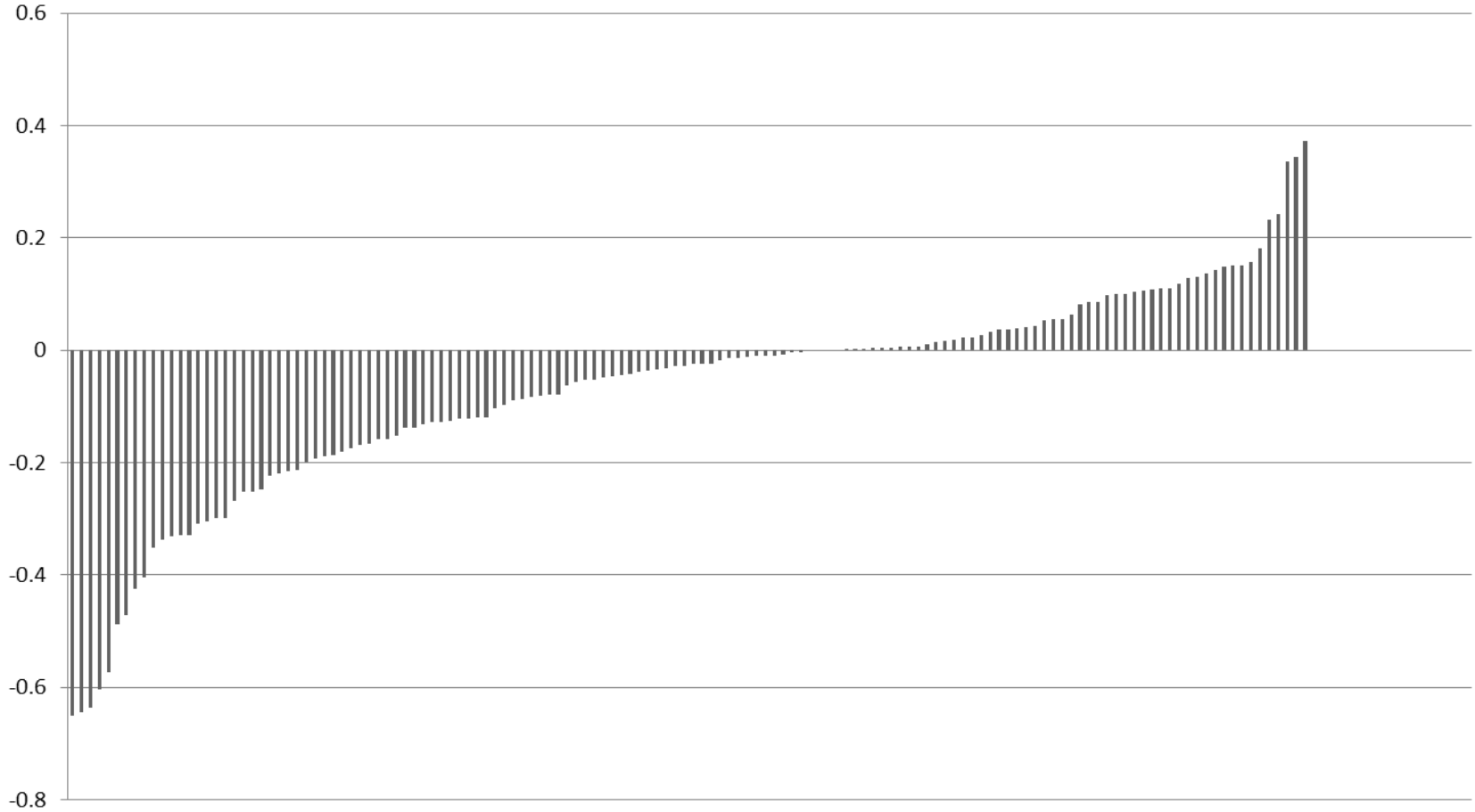
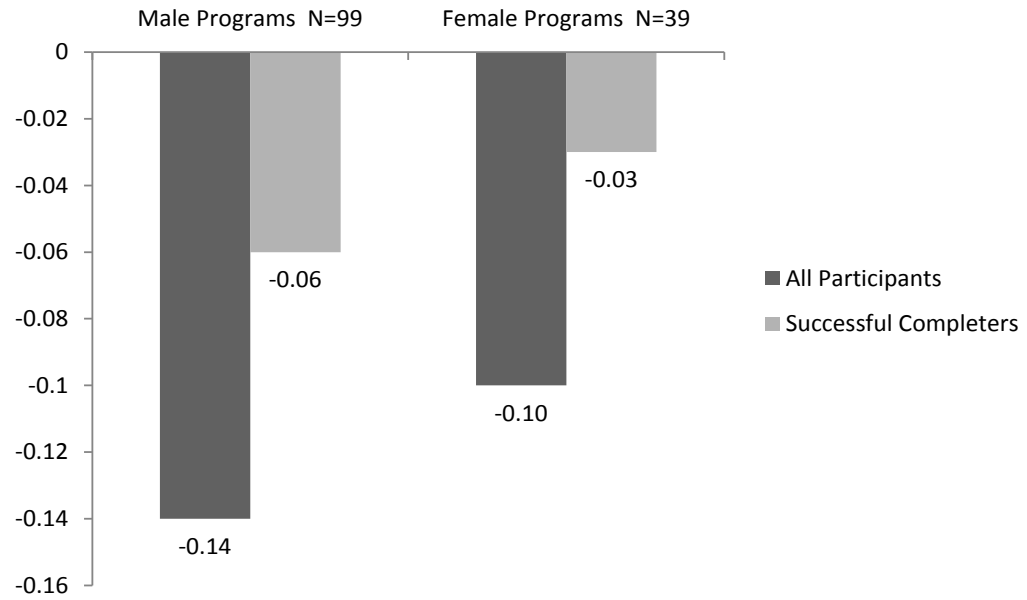
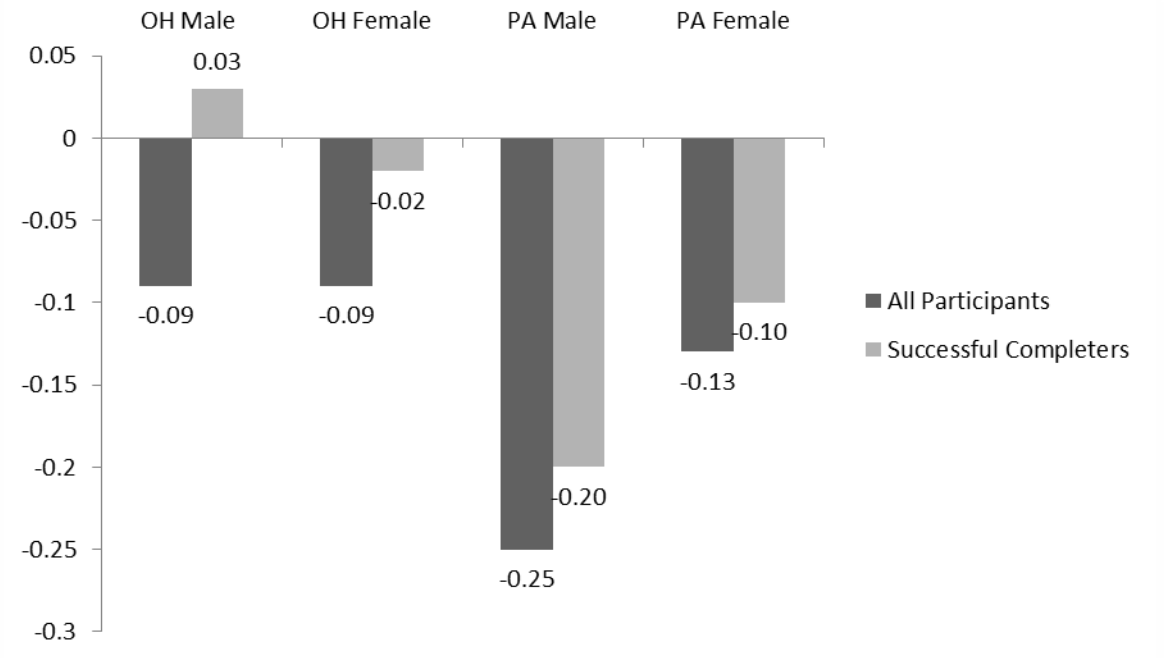


Figure 4.3: Overall Treatment Effect by Program Gender



All Participants: t= -1.895; p=.06
Successful Completers: t= 2.418; p=.02

Figure 4.4: Overall Treatment Effect by Program Gender and State



OH All Participants: t= -0.125; p=.90; OH Successful Completers: t=2.418; p=.02
PA All Participants: t= -1.640; p=.11; PA Successful Completers: t= -1.468; p=.15

Figure 4.6 examines the effect size distribution for successful completers only. Here, 61 percent of male programs had a negative treatment effect versus 55 percent of female programs. Hence, overall, a higher proportion of female programs showed positive treatment effects and there is variability in the data, with some programs producing positive outcomes.

While the average effect size for the programs being examined is negative, the variability in the effect sizes allows for the research questions to be addressed. Of note, the analyses to follow will examine the *difference* in predicted probability of recidivism by program characteristic and gender. This allows for comparison between programs with an average negative treatment effect to determine whether a particular characteristic *improves* the outcome, even if the average effect size is still negative overall. Nonetheless, it is important to note that characteristics shown to "improve" program outcomes are still, oftentimes associated with overall negative treatment effects. Hence, programs that possessed these positive program characteristics show improved results, but may still have performed more poorly than the matched comparison sample. This will be noticed throughout this section of the dissertation.

Figure 4.5: Effect Size Distribution for the Full Sample by Gender

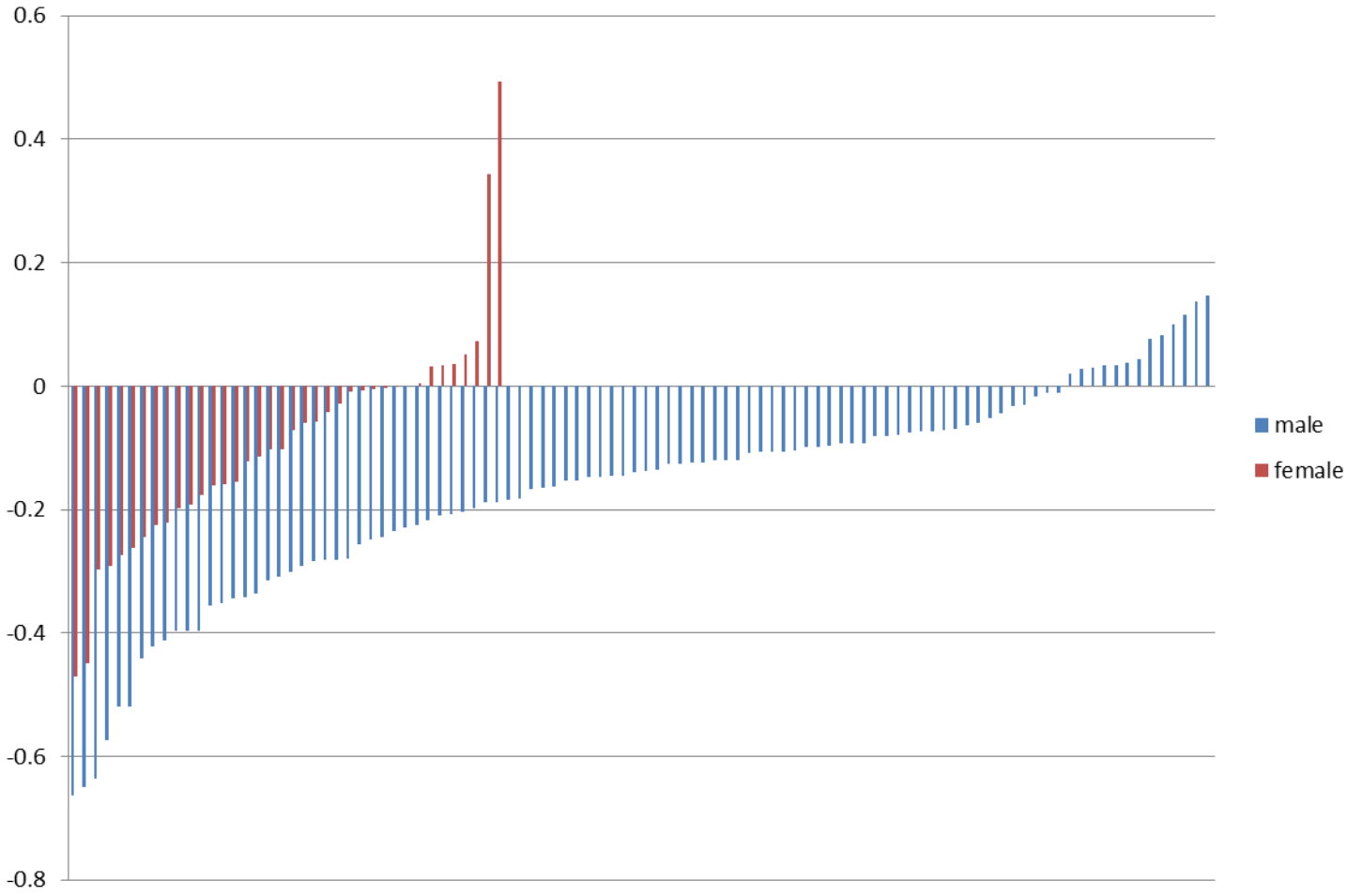
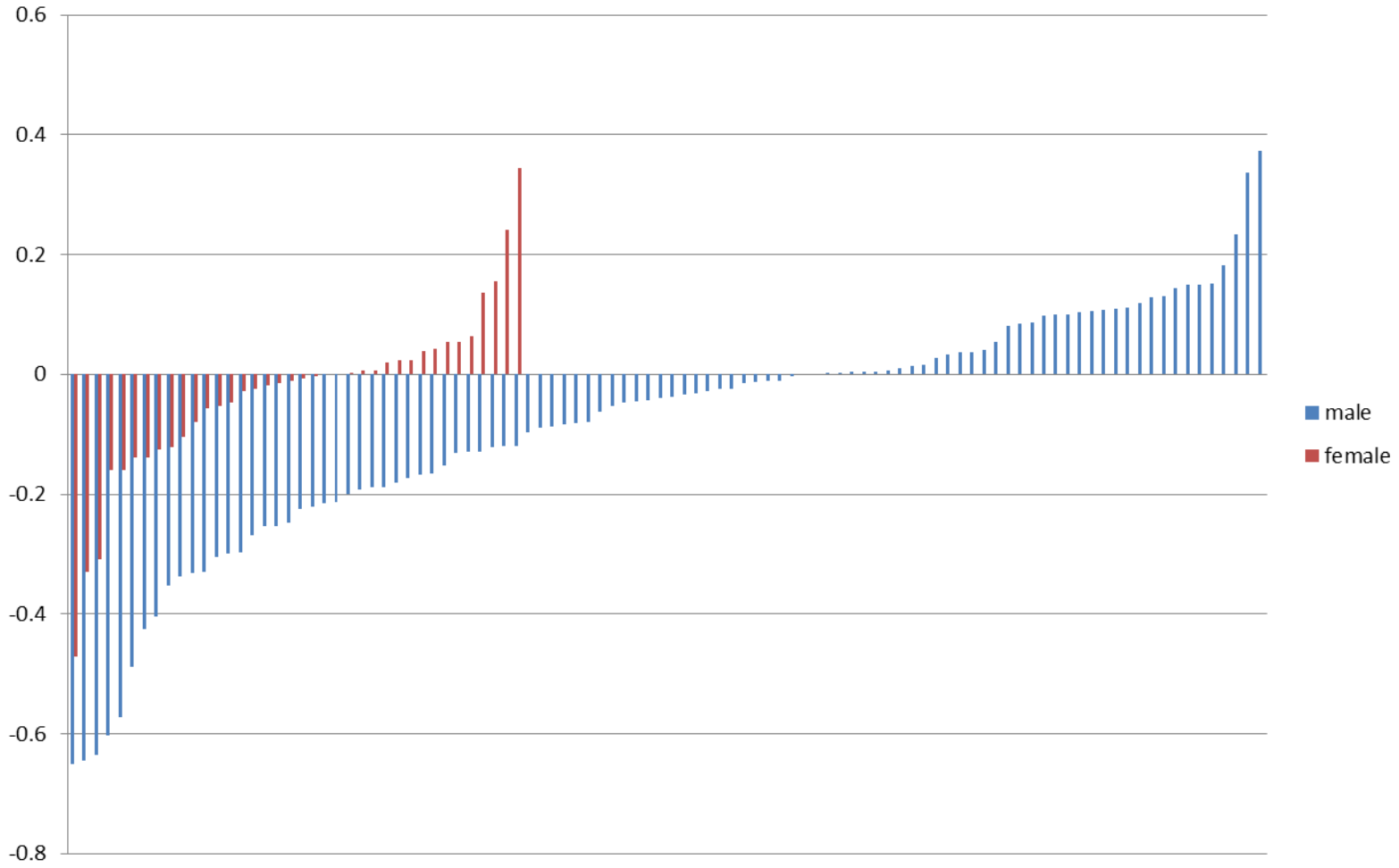


Figure 4.6: Effect Size Distribution for the Successful Completer Sample by Gender



Bivariate Results

This next section of the dissertation will examine each program characteristics' relationship with recidivism. Presentation of bivariate analyses will be broken down into Leadership/Design, Staff, Assessment Tools and Characteristics, Treatment Targets, Group Interventions, and General Treatment Characteristics. Note that some program characteristics had limited variation, suggesting it was a characteristic that few programs had (e.g. aftercare), or a characteristic that nearly all programs had (e.g. use of CBT components). Lack of variation on an item makes the results less stable, rendering conclusions about the findings less sound. Findings for all variables examined can be found in the tables, despite variation issues. However, variables that lacked variation are highlighted in grey in the tables to remind readers to interpret these findings with caution[46].

Leadership and Design Characteristics

The first set of analyses examines how program leadership and design characteristics contribute to effective treatment strategies for males versus females in programs. There are ten items in this category, four related to design and function, three items to leaderships' role, and three devoted to how co-ed programs are structured. *Co-ed education* is highlighted in this table as this item lacks variation; conclusions related to this item are therefore not drawn. Table 4.1 shows that a higher program budget appears significantly related to improved program effects for males in both the full and successful completer sample (r=.29 and .36 respectively). This finding may be tied to the resources needed to implement evidence-based treatment strategies.

[46] For male programs, cells with fewer than 18 cases are highlighted and conclusions about findings are withheld; for female programs, cells falling below 8 cases (approximately 20% of the sample) are highlighted.

Differences in effect size were not significant for programs serving females, although the direction of the relationship was similar to that of males in the full sample.

Years of operation tends to be tied with a program's stability and role as part of the local correctional system. Interestingly, for females in the full sample, years of operation had a negative correlation with treatment effects; programs 16 years or older had a treatment effect of -.18 versus -.07 for those operating fewer than 6 years. Years of operation was not correlated to recidivism for males. Perhaps newer programs that serve females tend to be more mindful of the gender-responsive literature, as opposed to more stable programs that may be more likely to conduct "business as usual'.

The next variable in Table 4.1 examines the staff to resident ratio in programs. This accounts for all program employees, relative to the program's capacity for residents. The gender-specific literature emphasizes women's relational nature, and more heavily staffed programs may have more time to spend with participants. Findings, however, suggested that more heavily staffed programs neither hindered nor benefited programs, regardless of gender. The gender-specific literature also speaks to the importance of linking females to community resources to assist in addressing their array of needs. Hence, the program's rating of support from the community with which it is housed was included. This item had a negative correlation with outcome for males in the full sample, suggesting that a higher rating by program staff for community support was associated with poorer outcome. These findings did not hold, however, when successful completers only were examined.

The next three items examine the level of involvement and effectiveness of the program leader. In the full sample, having a program director involved in the delivery of clinical training

Table 4.1: Program Leadership/Design Characteristics

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Program budget								
0-2.5 mil	-.16 (42)	.29[c]	-.12 (12)	.18	-.10 (43)	.36[b]	-.04 (11)	-.06
2.6-6 mil	-.09 (28)		-.08 (9)		.01 (28)		-.05 (9)	
Years of operation								
0-5	-.14 (64)	-.03	-.07 (26)	-.44[b]	-.06 (66)	.06	-.01 (26)	-.28
16+	-.15 (35)		-.18 (10)		-.04 (33)		-.08 (9)	
Staff to resident ratio								
1 to 3 or less	-.13 (28)	-.09	-.10 (18)	-.05	-.07 (29)	-.02	-.03 (17)	-.04
More than 1 to 3	-.15 (52)		-.12 (12)		-.07 (52)		-.04 (11)	
Support rating from community-at-large								
0-7	-.12 (61)	-.18[d]	-.10 (22)	.09	-.04 (57)	-.06	-.03 (21)	-.03
8-10	-.17 (37)		-.08 (15)		-.09 (40)		-.03 (15)	
PD delivers clinical training								
No	-.15 (55)	.07	-.14 (17)	.36[c]	-.07 (55)	.06	-.04 (16)	.08
Yes	-.13 (44)		-.06 (19)		-.05 (43)		-.02 (20)	
PD delivers services								
No	-.16 (54)	.16	-10 (22)	.03	-.09 (56)	.24[c]	-.03 (22)	.02
Yes	-.12 (44)		-.09 (15)		-.01 (42)		-.03 (14)	

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level

Table 4.1 Continued: Program Leadership/Design Characteristics

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Harmony between staff and management								
Moderate to minor problems	-.15 (74)	.07	-.11 (26)	.16	-.06 (72)	.06	-.03 (26)	.03
No problems	-.13 (25)		-.07 (11)		-.04 (26)		-.02 (10)	
Co-ed—Shared living space								
No	-.08 (19)	-.41[b]	-.04 (10)	-.38[d]	.00 (20)	-.29[d]	.02 (11)	-.44[c]
Yes	-.20 (21)		-.13 (12)		-.11 (22)		-.07 (13)	
Co-ed—Shared education								
No	-.11 (30)	-.26	-.08 (18)	-.19	-.01 (30)	-.34[c]	-.01 (19)	-.36[d]
Yes	-.20 (8)		-.14 (4)		-.16 (9)		-.11 (5)	
Co-ed—Shared visitation								
No	-.09 (28)	-.46[b]	-.08 (18)	-.27	.02 (28)	-.55[a]	-.01 (18)	-.38[d]
Yes	-.24 (10)		-.16 (5)		-.20 (12)		-.11 (5)	

Highlighted items lack variation; findings related to these items are therefore not discussed

was associated with significantly improved effects for programs serving females (r=.36), while there was a limited relationship for males. However, when examining the effects of the programdirector's involvement in delivering direct services, this attribute appears more important for programs serving males, significantly so in the successful completer sample, with an 8 percentage point improvement in effect size. Harmony between staff and management was also examined, based again on theory that females tend to be more relational, and may be more affected by disharmony in a program. Program harmony was not significantly correlated with program outcome for either male or female programs.

The final items look at shared living space and visitation among those programs serving both men and women (defined here as co-ed programs). There were a total of 62 programs that fell into this category. With regard to shared living, this includes any space within the program that men and women are sharing at the same time. In programs where both males and females used the same cafeteria or recreation area, but times were scheduled so that men and women had no contact, this item was scored as no. The data showed that this item is significantly negatively correlated for both men and women in the full and successful completer samples. Programs that shared living space between genders had between a 9 and 12 percentage point lower treatment effect relative to those that kept the living space separate. Co-ed visitation also showed to be problematic for programs serving both males and females. This suggests that keeping living spaces and visitation separate appears to be an important program characteristic for both males and females. The effect of mixing treatment groups will be addressed in the treatment section.

Program Staff Characteristics

This section generally examines the selection of program staff, how staff is trained, supervised and evaluated, and staff philosophy about treatment. A total of 13 items were

examined in this category. One of the 13 items, *Training on Program Theory/Philosophy* lacked variation in the female sample; conclusions are therefore not drawn about this characteristic. The first two items in Table 4.2 look at the selection of staff. Seeking new staff with the belief that correctional treatment works appeared more important for programs serving females. In the full sample, female programs had a 50 percent improvement in effect size when such staff was sought. While this characteristic would seemingly be important for any program, it appeared neither of value or hindrance for programs serving males. Interestingly, for programs serving males, there were significantly poorer treatment effects when staff members with experience in corrections were sought. Correlations for female programs were not significant, but were in the same direction. Perhaps individuals without an extensive corrections background are more open-minded as to how to work most effectively with an offender population.

The next two items examine current staff credentials. The first looks at the impact of the staff's experience in corrections. Experience was defined as at least 75 percent of program staff having 2 or more years of experience in a correctional treatment setting. Interestingly, in programs serving males, having more experienced staff was negatively correlated with outcome (r=-.41 and -.51 respectively). Experience was not significantly related to outcome for females. The opposite trend is true for having educated staff. For programs where at least 70 percent of staff delivering treatment had an associate's degree or higher, outcomes went down for programs serving women. There was a 7 percentage point decrease in effect size for female programs with more educated staff. Education was not significantly related to outcome for males. One would assume that employing more experienced and educated staff would increase treatment effects; perhaps other factors are linked to education and experience, such as burn-out or lack of openness to new treatment strategies. Nonetheless, the variation by gender is perplexing.

Table 4.2: Program Staff Characteristics

Variable		All Participants				Successful Completers			
		Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Selection of new Staff									
Seeks belief that treatment works									
	No	-.14 (56)	.04	-.14 (15)	.32[c]	-.06 (55)	.05	-.06 (14)	.22
	Yes	-.14 (43)		-.07 (22)		-.05 (43)		-.01 (22)	
Seeks any corrections experience									
	No	-.12 (57)	-.22[c]	-.07 (25)	-.11	-.03 (57)	-.18[d]	-.02 (25)	-.06
	Yes	-.17 (42)		-.14 (12)		-.09 (41)		-.04 (12)	
2 or more years relevant experience									
	No	-.08 (36)	-.41[a]	-.16 (9)	.27	.05 (31)	-.51[a]	-.05 (8)	.10
	Yes	-.19 (54)		-.09 (21)		-.12 (60)		-.03 (22)	
Treatment staff higher education									
	No	-.10 (42)	.16	-.06 (18)	-.32[d]	.02 (37)	.12	.00 (17)	-.32[d]
	Yes	-.08 (27)		-.13 (14)		.04 (25)		-.04 (14)	
Training on program theory/philosophy									
	No	-.15 (73)	.16	-.11 (30)	.34[c]	-.07 (72)	.10	-.04 (29)	.13
	Yes	-.11 (26)		-.02 (7)		-.03 (27)		.00 (7)	
Clinical meeting scale									
1 or fewer characteristics		-.17 (67)	.23[c]	-.11 (22)	.06	-.09 (71)	.26[c]	-.03 (23)	-.04
2 + characteristics		-.10 (23)		-.09 (13)		.01 (21)		-.04 (11)	

Table 4.2 Continued: Program Staff Characteristics

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Clinical supervision								
No	-.19 (47)	.37[a]	-.11 (14)	.12	-.12 (50)	.38[a]	-.02 (13)	-.08
Yes	-.10 (51)		-.08 (23)		.00 (49)		-.03 (23)	
Staff Evaluation								
Soft skills								
0 to 1	-.11 (57)	-.31[b]	-.08 (19)	-.15	-.03 (56)	-.20[d]	-.03 (19)	-.01
2 to 3	-.19 (39)		-.11 (18)		-.09 (39)		-.03 (17)	
Hard skills								
0 to 2	-.15 (59)	.05	-.10 (20)	.02	-.07 (60)	.14	-.04 (20)	.19
3 to 4	-.13 (33)		-.09 (13)		-.02 (30)		.00 (12)	
Communication/relationship skills								
0-2	-.14 (74)	.00	-.10 (28)	.07	-.07 (75)	.21[c]	-.04 (28)	.21
3-5	-.14 (22)		-.08 (9)		.01 (20)		.01 (8)	
Treatment staff value treatment efforts								
0-8 rating	-.13 (40)	-.11	-.12 (18)	.20	-.04 (39)	-.08	-.03 (18)	-.03
9-10 rating	-.15 (57)		-.07 (18)		-.07 (58)		-.03 (18)	
Security staff value treatment efforts								
0-7 rating	-.13 (57)	-.07	-.09 (18)	-.07	-.03 (53)	-.15	-.02 (18)	-.08
8-10 rating	-.15 (38)		-.10 (17)		-.08 (40)		-.04 (17)	
Staff turnover								
Problem	-.14 (49)	.02	-.11 (17)	.13	-.04 (44)	.-09	-.01 (16)	-.14
Minor or No Problem	-.14 (50)		-.08 (20)		-.07 (54)		-.04 (20)	

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level

Highlighted items lack variation; findings related to these items are therefore not discussed

Programs that offered clinical meetings with certain positive attributes fared better for male programs in both the full and successful completer sample. The positive attributes consisted of: 1) four or more meetings held per month; 2) time spent reviewing case progress; 3) involvement of program director; 4) involvement of case manager; and 5) involvement of security staff. Here, more attributes were associated with improved treatment effects for males. Findings also support the need for clinical supervision by a licensed social worker or counselor. Programs serving males had significantly higher effect sizes when such clinical supervision was in place (a 9 to 12 percentage point decrease in recidivism). The effect of both clinical meetings and clinical supervision on outcome was non-significant for the female programs.

Staff evaluation of skills is typically thought of as an important component of treatment programs. The next three items in Table 4.2 describes staff evaluation in the following clusters:

1) Soft skills—creativity, open to supervision, team player;
2) Hard skills—paperwork, attendance/tardiness, dress, productivity;
3) Communication/relationship skills—crisis de-escalation, boundaries, avoiding negative interactions with clients, firm but fair, empathic.

With regard to soft skills, more attributes were associated with negative treatment effects for programs, significantly so for the programs serving males (r=-.31 and -.20 respectively). Evaluation of hard skills had no significant impact on treatment effects for either male or female programs. Evaluation of communication/relationship skills had little impact for the full sample, but positive effects were seen when these attributes were evaluated in the successful completer sample, significantly so for the programs serving males. This underscores the importance of evaluating staff's ability to interact appropriately with residents.

The next two staffing items include a 1 to 10 rating of how both treatment and security staff value the treatment efforts by the program. A higher rating for neither treatment nor

security staff was significantly associated with outcome. The final staff item examines staff turnover. This item was included as females could arguably be more likely to bond with staff and be affected by staff instability, according to relational theory. Turnover, however, had no significant impact on program outcomes for either male or female facilities.

Program Assessment Tools

The next section examines the types of assessments used by programs, along with other more general program assessment and admission practices. Table 4.3 describes the various assessment types and offender need areas being formally evaluated by programs. Nine items were included in this area; however, three of the nine items (those highlighted) lacked variability, rendering these findings less reliable. Positive correlations were associated with use of a risk need assessment, but were only significant for programs serving males. In examining use of the Level of Service Inventory-Revised (LSI-R) specifically as a risk/need tool[47], increased program effects were seen with male programs in the successful completer sample (r=.21). Hence, use of a risk need tool at female programs was not shown to have a negative impact, as some feminists argue, but unlike male programs, it was also not significantly correlated with positive outcome.

Table 4.3 also examines the importance of assessing particular offender need areas. Use of any validated need assessments, beyond a general risk need tool, showed limited variation in programs serving females, as did using a criminal attitude tool specifically (few male or female programs used such a tool). Use of a validated substance abuse tool did have variation for both male and female programs; for male programs in the successful completer sample, use of a substance abuse assessment was significantly related to improved treatment effects. Application

[47] The LSI-R was the only instrument specifically examined in these analyses, as it was the most commonly used risk assessment (N=68). Use of this tool specifically has also been debated by those in the gender specific and RNR camp (Holtfretter and Cupp 2007; Lowenkamp et al. 2007).

Table 4.3: Program Assessment Tools

Variable		All Participants				Successful Completers			
		Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Assessments Used by program:									
Risk/Need Assessment									
	No	-.17 (42)	.18[d]	-.15 (10)	.27	-.11 (42)	.31[b]	-.07 (10)	.25
	Yes	-.12 (57)		-.08 (27)		-.01 (56)		-.01 (26)	
Risk assessed using LSI-R									
	No	-.16 (53)	.15	-.11 (14)	.13	-.09 (51)	.21[c]	-.04 (14)	.13
	Yes	-.12 (46)		-.08 (22)		-.02 (47)		-.02 (22)	
Validated need assessment tools									
	No	-.15 (76)	.21[c]	-.11 (30)	.19	-.07 (75)	.12	-.04 (30)	.30[d]
	Yes	-.09 (23)		-.05 (7)		-.02 (23)		.03 (6)	
Criminal attitude tool									
	No	-.15 (83)	.14	-.10 (32)	.16	-.06 (82)	.07	-.04 (32)	.29[d]
	Yes	-.10 (15)		-.05 (5)		-.03(16)		.04 (4)	
Validated substance abuse tool									
	No	-.15 (51)	.07	-.06 (13)	-.24	-.09 (51)	.21[c]	-.02 (13)	-.10
	Yes	-.13 (48)		-.11 (24)		-.02 (47)		-.04 (24)	
Validated responsivity tools									
	No	-.15 (69)	.13	-.11 (21)	.16	-.06 (64)	.04	-.02 (19)	-.12
	Yes	-.12 (30)		-.07 (16)		-.05 (34)		-.04 (17)	
Biopsychosocial tool									
	No	-.12 (56)	-.19[d]	-.09 (23)	-.01	-.03 (57)	-.23[c]	-.02 (23)	-.08
	Yes	-.17 (43)		-.10 (13)		-.10 (41)		-.04 (13)	
Mental health assessment									
	No	-.14 (80)	-.08	-.11 (26)	.19	-.05 (78)	-.12	-.03 (25)	.08
	Yes	-16 (18)		-.06 (11)		-.09 (20)		-.02 (11)	
Assess past abuse									
	No	-.12 (80)	-.24[c]	-.08 (29)	-.23	-.02 (78)	-.32[b]	-.02 (29)	-.26
	Yes	-.20 (15)		-.15 (7)		-.15 (15)		-.08 (7)	

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level; Highlighted items lack variation--findings related to these items are therefore not discussed

of such tools, however, was not associated with decreased incarceration rates for programs serving females. This finding is interesting given that this is one of the criminogenic need areas generally recognized as an important need for female offenders.

Also examined is the use of validated responsivity assessments in programs. Examples of common responsivity areas that might be assessed include mental health, motivation, victimization and learning ability. Programs serving females appeared more likely to use such tools (nearly 50% versus approximately 33% for male programs); however, significant improvement was not seen for either male or female programs when such tools were used. In examining specific types of responsivity assessments, use of a biopsychosocial assessment[48] had no relationship with outcome for female programs, and was negatively related to outcome for programs serving males. In the successful completer male sample, the average predicted probability of recidivism was seven percentage points higher for programs that used such a tool. Use of a mental health assessment was not significantly related to program outcome for male or female programs. Of note though, the correlation was in a negative direction for male programs and positive direction for those serving females. Too few programs assessed past abuse to draw conclusions about these findings.

Program Assessment Characteristics

Aside from the types of tools used by agencies, Table 4.4 displays program assessment characteristics. There were a total of 8 items examined in this area. The first several items examine how offenders are selected for participation in the program. In programs where the

[48] A biopsychosocial assessment can be described as a non-validated semi-structured interview that examines a range of need and responsivity areas. This tool is typically used to diagnose individuals according to the Diagnostic and Statistical Manual IV.

Table 4.4: Program Assessment Characteristics

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Who selects offenders for the program?								
Program Director								
No	-.11 (60)	-.32[b]	-.09 (25)	-.06	.00 (56)	-.44[a]	-.02 (25)	-.20
Yes	-.19 (39)		-.10 (12)		-.14 (42)		-.06 (12)	
Intake Department								
No	-.19 (39)	.31[b]	-.13 (9)	.17	-.12 (42)	.34[a]	-.07 (9)	.21
Yes	-.11 (60)		-.08 (27)		-.01 (56)		-.02 (27)	
Admission Criteria								
Written exclusionary criteria								
No	-.17 (29)	.16	-.15 (8)	.25	-.11 (31)	.21[c]	-.07 (8)	.22
Yes	-.13 (70)		-.08 (28)		-.03 (67)		-.02 (28)	
Excludes offender with mental illness								
No	-.13 (43)	-.03	-.10 (19)	.05	-.05 (41)	-.04	-.02 (17)	-.12
Yes	-.15 (55)		-.09 (18)		-.06 (56)		-.04 (19)	
Proportion of population low risk								
Below 20%	-.11 (80)	-.46[a]	-.11 (12)	.08	-.01 (75)	-.57[a]	-.03 (11)	-.02
20% or more	-.26 (19)		-.09 (25)		-.21 (23)		-.03 (25)	
Number of assessments conducted								
0-3	-.16 (48)	.11	-.11 (14)	.09	-.08 (48)	.14	-.05 (13)	.13
4 or more	-.13 (51)		-.09 (23)		-.04 (50)		-.02 (23)	
Assessment training by observation only								
No	-.09 (30)	-.21[c]	-.07 (10)	-.15	.00 (30)	-.17	.01 (10)	-.27
Yes	-.15 (58)		-.11 (26)		-.06 (55)		-.05 (25)	
Reassessment conducted								
No	-.16 (59)	.22[c]	-.12 (17)	.21	-.07 (55)	.10	-.02 (16)	-.11
Yes	-.10 (37)		-.07 (16)		-.03 (40)		-.04 (18)	

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level

program director makes decisions about what offenders are admitted, there were significant negative treatment effects for programs serving males in both samples (r=-.32 and -.44). On the other hand, for programs that had an intake department that made decisions about admission, correlations were significantly positive for both male programs in both the full and successful completer sample (r=.31 and .34 respectively). The relationship between intake decision and outcome was not significant for female programs, although the direction of the relationship was similar to that of the male sites. This finding may be related to the more collaborative nature of admission decisions by an intake department, and the possibility that assessment results accompany this choice, given a specific department or dedicated staff is making such decisions.

Also related to admission decisions is what exclusionary criteria are used to determine program eligibility. Male programs in the successful completer sample showed significant improvement when written exclusionary criteria were in place (r=.21). The relationship was not significant for programs serving females, although the correlation was in a positive direction. Significant mental health issues tend to be a common exclusionary criteria since many community residential programs cannot accommodate such needs; yet, as previously discussed, it also tends to be an important need area for programs serving females. Hence, this specific exclusionary criterion was examined. Findings indicated that use of mental health as an exclusionary criterion was not strongly associated with outcome for either male or female programs. One who adheres to the gender-specific philosophy might have expected this item to be negatively related to program outcome.

Finally, many of the community residential programs used risk assessment results as part of the exclusionary criteria, primarily to limit the number of low risk offenders in

programming[49]. Programs serving males that enrolled fewer than 20 percent low risk offenders decreased the predicted probability of recidivism by 15 percentage points in the full sample and 20 percentage points among successful completers. There was no significant difference in outcomes related to the proportion of low risk offenders served for female programs. Of note, a higher percentage of female programs served low risk women relative to male programs. There were also fewer females that fell into the high risk category, particularly in the larger Ohio sample. Nonetheless, these results support the risk principle for males, but not females, which is contrary to previous research that found support for the risk principle for females (Dowden and Andrews, 2006; Lovins, et al. 2007).

The next few items in Table 4.4 examine assessment practices. While all correlations were positive, the number of assessment tools used was not significantly related to outcome for neither male nor female programs. Also examined was how evaluators were trained to conduct the risk/need assessment. Programs that had no formal training on the tool, but instead trained staff members by having them observe colleagues conduct the assessment showed an increase in negative treatment effects, particularly for male programs (r=-.21). There was not a significant relationship between risk assessment training and outcome for female programs, although the direction of the relationship was also negative. Finally, the use of reassessment was inspected. Reassessment allows programs to formally examine whether specific need areas are being effectively targeted. Male programs in the full sample that conducted reassessment had significantly improved treatment effects (a 5 percentage point improvement over programs that

[49] Proportion low risk was calculated using the risk assessment developed for the OH study and the LSI-R results available for all PA offenders. Hence, these data are specific to the sample of offenders included in the study versus proportions reported by the program or those coded based on file review. The method used was deemed the most reliable measure of percentage of the population by risk.

did not reassess offenders). There was not a significant relationship between reassessment and outcome for programs serving women.

In all, the use of risk assessment and adherence to the risk principle was clearly demonstrated for programs serving males. While this finding supports the RNR literature, the lack of similar support for female programs does not. On the other hand, use of tools to assess need areas deemed crucial by the gender-specific literature such as substance abuse or mental health was not significantly associated with improved outcome for female programs. Hence, some assessment items appear supportive of RNR whereas others offer more support for a gender specific perspective.

Program Treatment Targets

The next three tables examine program treatment characteristics. This area is subdivided into program treatment targets, specific group interventions, and other more general program characteristics. Table 4.5 examines key target areas identified and treated by the program. There are sixteen variables examined in this area. These were areas that either staff acknowledged as key areas of focus when conducting group, individual or family intervention, or areas identified based on the program's treatment or group schedule. Of the sixteen items, half lacked enough variation to draw meaningful conclusions. These, again, are highlighted in the table and will not be discussed.

The first item on this table examines the number of Tier 1 criminogenic needs targeted. Included here are the following needs: criminal attitudes/orientation/values, antisocial peer associations, anger/hostility, self-control and interpersonal conflict resolution. Contrary to much of the literature on treatment targets associated with decreased recidivism (see Andrews and Bonta 2010), targeting a higher number of these specific characteristics was not significantly

Table 4.5: Program Treatment Targets

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Tier 1 Criminogenic Needs Targeted								
0 to 2	-.15 (37)	.06	-.04 (16)	-.40[c]	-.05 (37)	.00	.01 (15)	-.34[c]
3 or more	-.13 (61)		-.13 (21)		-.05 (60)		-.06 (21)	
Tier 2 Criminogenic Needs Targeted								
0 to 3	-.15 (79)	.08	-.08 (25)	-.22	-.06 (79)	.09	-.02 (24)	-.09
4 or more	-.12 (19)		-.13 (12)		-.03 (20)		-.04 (12)	
Gender Specific Needs Targeted								
0 to 2	-.14 (58)	.01	-.05 (13)	-.29[d]	-.07 (59)	.08	-.01 (13)	-.11
3 or more	-.14 (41)		-.12 (24)		-.04 (39)		-.04 (23)	
Individual Gender Responsive Needs—Gender neutral targets:								
Drug/Alcohol								
No	-.19 (11)	.13	.01 (2)	-.19	-.14 (12)	.20[c]	.03 (2)	-.12
Yes	-.13 (88)		-.10 (35)		-.05 (86)		-.03 (34)	
Vocational achievement								
No	-.24 (6)	.20[c]	-.12 (4)	.08	-.21 (7)	.27[b]	-.09 (4)	.21
Yes	-.13 (93)		-.09 (33)		-.05 (91)		-.02 (32)	
Educational achievement								
No	-.16 (30)	.08	-.12 (7)	.09	-.08 (25)	.08	-.07 (8)	.20
Yes	-.13 (69)		-.09 (30)		-.05 (73)		-.02 (28)	
Family affection/communication								
No	-.12 (74)	-.29[b]	-.11 (29)	.21	-.03 (71)	-.28[b]	-.03 (29)	.03
Yes	-.20 (25)		-.05 (8)		-.13 (27)		-.02 (7)	

Table 4.5 Continued: Program Treatment Targets

Variable		All Participants				Successful Completers			
		Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Family problem solving									
	No	-.12 (77)	-.24[c]	-.11 (31)	.30[d]	-.02 (73)	-.35[a]	-.04 (30)	.22
	Yes	-.20 (21)		-.02 (6)		-.15 (25)		.02 (6)	
Offender relationship with sig other									
	No	-.14 (76)	-.06	-.08 (27)	-.16	-.06 (78)	.04	-.04 (27)	.12
	Yes	-.15 (23)		-.13 (10)		-.04 (20)		-.01 (10)	
Gender specific factors:									
Low self esteem									
	No	-.14 (70)	-.03	-.09 (14)	-.04	-.05 (71)	-.05	-.02 (14)	-.04
	Yes	-.15 (28)		-.10 (23)		-.07 (28)		-.03 (22)	
Mental health issues									
	No	-.16 (50)	.12	-.10 (13)	.00	-.07 (49)	.06	-.04 (12)	.07
	Yes	-.13 (49)		-.10 (24)		-.05 (50)		-.02 (24)	
Economic/social needs									
	No	-.14 (44)	-.02	-.07 (19)	-.22	-.07 (49)	.10	-.03 (18)	.00
	Yes	-.14 (54)		-.12 (18)		-.04 (50)		-.03 (19)	
Parenting Skills									
	No	-.14 (57)	.03	-.04 (9)	-.30[d]	-.06 (55)	.01	.02 (10)	-.33[c]
	Yes	-.14 (41)		-.11 (28)		-.06 (43)		-.05 (26)	
Childhood abuse/neglect									
	No	-.14 (86)	.06	-.08 (23)	-.19	-.06 (85)	.02	-.02 (22)	-.14
	Yes	-.12 (13)		-.12 (13)		-.05 (14)		-.05 (14)	
Trauma/PTSD									
	No	-.15 (85)	.25[c]	-.08 (22)	-.13	-.07 (84)	.19[d]	-.01 (22)	-.21
	Yes	-.06 (13)		-.11 (14)		.02 (14)		-.06 (14)	
Offender relationship with children									
	No	-.14 (84)	-.03	-.06 (24)	-.39[c]	-.06 (84)	.10	-.02 (24)	-.19
	Yes	-.15 (15)		-.16 (12)		-.02 (14)		-.06 (12)	

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level; Highlighted items lack variation--findings related to these items are therefore not discussed

correlated with outcomes for male programs, and was significantly related to recidivism for female programs, but the direction was negative. This suggests that when programs serving females target Tier 1 risk factors, recidivism rates get worse. This finding supports the gender responsive literature, which de-emphasizes focus on Tier 1 risk factors for females, and emphasizes more of the Tier 2 risk factors (Bloom et al. 2003), and fails to support the generalist perspective.

The following factors were coded as Tier 2 criminogenic needs: alcohol or drug problems, structured use of leisure time, vocational achievement, educational achievement, and relationship with significant others. Like with the Tier 1 risk factors, targeting Tier 2 factors had a positive, but non-significant relationship with outcome for programs serving males, and a negative, but again non-significant, relationship with outcome for female programs. Tier 2 need areas are recognized by both the gender specific and generalist literature as important target areas for both male and female offenders; hence, this finding supports neither the generalist nor gender specific perspective. The limited relationship between criminogenic targets and outcome may be related to *how* these needs are targeted, which will be examined in a later section.

Given the topic of this dissertation, in addition to criminogenic needs, a third category, gender specific needs, was also identified. These needs include those advocated as primary targets by the gender responsive literature, but clearly identified as responsivity areas, thus not primary treatment targets associated with recidivism by the RNR literature. Included in the gender specific needs are: mental health, self-esteem, childhood abuse/neglect, trauma/PTSD, relationship with children, parenting skills, and economic/social needs. Like the previous two scales measuring criminogenic targets, this scale had a limited relationship with recidivism for programs serving males. Interestingly, in the full participant sample, female programs that

targeted gender specific needs had a significantly higher failure rate, which remained negative (although non-significant) in the successful completer sample. This finding supports the generalist perspective, and contradicts the gender specific perspective. However, given that outcome in this study is measured by a program's ability to reduce recidivism, it is not surprising that targeting a cluster of factors that historically show a limited relationship with recidivism is not associated with positive treatment effects.

The next set of variables in Table 4.5 examines how individual gender responsive need factors correlate with recidivism. Included are both gender specific needs not typically advocated as an appropriate primary target in male programs (e.g. self-esteem), and gender neutral need areas, advocated as an important need area in both male and female programs (e.g. substance abuse). It is important to note that gender specific needs can certainly occur among male offenders (e.g. PTSD), but correctional programming designed for males does not tend to be structured to systematically address such needs, except through individual or case management sessions. All factors analyzed, whether gender specific or neutral, were selected because they are generally noted by the gender responsive literature as important treatment targets (see Van Voorhis et al. 2010).

Gender neutral factors will be examined first. Note that of the six items analyzed in this section, only one had enough variation for meaningful results. The large majority of programs targeted drug/alcohol issues, vocational and educational achievement, while very few programs targeted family affection/communication and family problem solving. The one gender neutral item with sufficient variation was targeting the offender's relationship with a significant other; this item was not significantly related to recidivism for males or females.

The next set of variables examined in Table 4.5 is the gender specific factors. Here, three of the items (childhood abuse/neglect, trauma/PTSD, and offender relationship with children) were not targeted enough in programs serving males to make a meaningful comparison. However, this was not true for programs targeting self-esteem, mental health, economic/social needs and parenting skills. With regard to self-esteem, while all correlations were slightly negative, this factor generally appeared to have little relationship with program outcome for either male or female programs. Similarly, although in a positive direction, neither male nor female programs that targeted mental health as a need factor experienced a significant improvement in treatment effects. Economic/social needs as a treatment target also showed no significant effect on outcome for either gender. Interestingly, targeting parenting skills had little impact on recidivism for programs targeting males, but female programs showed a significant decline in outcome when this factor was targeted, in both the full and successful completer sample (r=-.30 and -.33 respectively). This finding would stand in contrast to the gender specific literature.

Among the gender specific need areas examined, none had a significant positive relationship with recidivism reduction for programs serving females, which parenting skills actually showing iatrogenic effects for female programs. In general, these findings suggest that while the gender responsive literature suggests that a range of female specific needs should be targeted in programs serving women, targeting such needs do not appear strongly related to positive treatment effects, as measured by reduced recidivism.

Program Group Interventions

While Table 4.5 examined general treatment target areas, Table 4.6 looks more specifically at group interventions offered by the program. This set of items is important, as

correctional programs tend to use group strategies as a primary mode of treatment within the program. This table will examine both the group area being targeted as well as characteristics of the group interventions. There are 16 items in this section. Five of the 16 items had limited variation; few programs offered a gender specific or mental health group, while the large majority of program provided a substance abuse group. Similarly, most programs used a structured curriculum and incorporated some role play into the group interventions. Data on these items are shown in the tables, but discussion of the results will be withheld due to limited variation.

The first item to be examined is use of Thinking for a Change as a treatment curriculum[50]. Use of this curriculum was significantly correlated with positive treatment effects for programs serving males in both samples (r=.25 and .28). Specifically, for successful completers, there was an average nine percentage point reduction in predicted probability of recidivism when this curriculum was used. For female programs, use of this curriculum had no relationship with outcome. The use of other cognitive behavioral groups designed to target general offender needs was also examined (e.g., Moral Recognation Therapy, EQUIP, or a general thinking errors group). Such group interventions had a limited relationship with outcome for either male or female programs.

With regard to anger management, offering this group intervention was significantly related to positive treatment effects for male programs in both samples, with up to an eight percentage point reduction in recidivism. This variable was not significantly correlated with

[50] Thinking for a Change is the only specific curriculum that was examined, as it was one commonly seen across many of the OH and PA programs. There was a significant amount of variation in the type of curricula used for other need areas, such as substance abuse or anger management. Hence, these interventions were examined as type of group versus use of a specific curriculum within that need area. Thinking for a Change is a curriculum whose development was sponsored by the National Institute of Corrections. The curriculum is based on a cognitive-behavioral model and targets cognitive restructuring, problem solving and social skills.

Table 4.6: Program Group Interventions

Variable		All Participants				Successful Completers			
		Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Group Interventions Offered									
Thinking for a Change									
	No	-.16 (61)	.25[c]	-.10 (21)	.03	-.09 (62)	.28[b]	-.03 (20)	.06
	Yes	-.10 (38)		-.09 (15)		.00 (36)		-.02 (16)	
Other Cog -based group									
	No	-.14 (67)	-.03	-.10 (20)	.04	-.05 (65)	-.10	-.03 (20)	.02
	Yes	-.15 (31)		-.09 (16)		-.08 (33)		-.03 (16)	
Substance abuse									
	No	-.18 (29)	.20[c]	-.07 (7)	-.10	-.13 (31)	.32[a]	-.04 (7)	.06
	Yes	-.12 (70)		-.10 (30)		-.02 (67)		-.03 (29)	
Anger management									
	No	-.17 (60)	.30[b]	-.10 (22)	.08	-.09 (59)	.21[c]	-.01 (20)	-.17
	Yes	-.09 (39)		-.08 (15)		-.02 (39)		-.05 (16)	
Employment/vocational education									
	No	-.18 (59)	.17[d]	-.09 (26)	-.03	-.10 (61)	.32[a]	-.03 (24)	-.01
	Yes	-.11 (39)		-.10 (11)		.01 (37)		-.03 (12)	
Family									
	No	-.15 (74)	.08	-.09 (18)	-.07	-.07 (73)	.12	-.04 (18)	.08
	Yes	-.12 (25)		-.10 (18)		-.02 (25)		-.02 (18)	
Life-skills									
	No	-.14 (67)	.01	-.11 (18)	.09	-.06 (68)	.04	-.05 (18)	.20
	Yes	-.14 (32)		-.08 (18)		-.05 (31)		-.01 (18)	
Gender-specific									
	No	-.14 (95)	-.13	-.08 (27)	-.16	-.06 (95)	.04	-.03 (28)	.02
	Yes	-.22 (4)		-.13 (9)		-.02 (3)		-.03 (8)	
Mental health									
	No	-.15 (91)	.14	-.12 (28)	.31[d]	-.06 (91)	.17[d]	-.04 (27)	.17
	Yes	-.08 (7)		-.03 (9)		.04 (7)		.00 (9)	

Table 4.6 Continued: Program Group Interventions

Variable		All Participants				Successful Completers			
		Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Eclectic									
	No	-.12 (75)	-.23[c]	-.10 (28)	.04	-.04 (74)	-.15	-.03 (27)	-.06
	Yes	-.19 (24)		-.09 (8)		-.10 (25)		-.04 (9)	
Groups assigned based on need									
	No	-.17 (45)	.24[c]	-.11 (10)	.05	-.11 (46)	.36[a]	-.04 (10)	.07
	Yes	-.11 (50)		-.09 (27)		.00 (48)		-.02 (26)	
Program mixes genders in groups									
	No	-.13 (62)	-.01	-.08 (22)	-.15	-.06 (62)	.15	-.01 (21)	-.20
	Yes	-.14 (26)		-.11 (13)		.00 (24)		-.05 (13)	
Number of groups provided									
	0-6	-.16 (56)	.21[c]	-.09 (11)	.01	-.09 (55)	.25[c]	-.04 (11)	.08
	7 or more	-.11 (43)		-.10 (26)		-.01 (43)		-.02 (26)	
Program uses structured curricula									
	No	-.26 (9)	.34[a]	-.07 (3)	-.07	-.20 (11)	.37[a]	-.07 (2)	.11
	Yes	-.12 (83)		-.10 (33)		-.03 (80)		-.03 (33)	
Use of role play									
	No	-.19 (20)	.20[c]	-.13 (2)	.07	-.11 (21)	.18[d]	-.09 (2)	.16
	Yes	-.13 (78)		-.09 (35)		-.04 (76)		-.02 (34)	
Use of graduated rehearsal									
	No	-.17 (63)	.31[b]	-.12 (18)	.23	-.09 (62)	.25[c]	-.04 (16)	.12
	Yes	-.09 (35)		-.07 (19)		.00 (36)		-.02 (20)	

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level

Highlighted items lack variation; findings related to these items are therefore not discussed

outcome for female programs. Anger and aggression may be less a need area for females versus males, which may account for these results.

The next item examines programs that offer structured educational and vocational interventions. This might include employment readiness or GED classes. Again, significant positive treatment effects are found for programs serving males in both samples (r=.17 and .32), but there is a minimal relationship between these types of interventions and outcomes for programs serving females (r=-.03 and -.01). Findings related to males support the RNR literature; however, both the RNR and gender specific perspective argues that education and employment are appropriate targets for women too.

Programs serving females were more likely to have a group intervention that targeted family issues (half of female programs versus just 25% of male programs). Family groups had no significant relationship with outcome for either male or female programs. This finding also stands in contrast with both the RNR and gender specific literature. The finding may be related to programs' typical difficulty with getting families to participate in regular family groups, particularly the higher risk families.

The next group intervention examined is life skills. There were no significant relationships between this intervention and outcome for either male or female programs, which is consistent with the RNR literature. The last category of treatment group is named "eclectic" groups. Examples of groups coded as eclectic include process group, group therapy, meditation, Power of Writing, music therapy, self-discovery, and didactic group. These groups include interventions that target a range of different needs or needs that are not easily categorized into one of the other areas. Eclectic groups were significantly related to negative outcomes for male

programs in the full sample (r=-.23), and there was no relationship with treatment effect for female programs; this is also consistent with the literature on effective treatment strategies.

The next set of items in Table 4.6 examines admission to groups and characteristics of the group interventions. The first item in this section is group assignment based on need assessment results. For male programs, assigning offenders to groups based on need assessments was significantly positively related to outcome in both samples (r=.24 and .36). There was a limited relationship between this item and outcome for female programs. Whether programs should offer mixed or single gender groups has always been a heavily debated topic. Generally, both the RNR and gender responsive literature seems to primarily support single-gendered groups, although the RNR literature is probably more neutral on this issue. There has, however, been little empirical examination of this question. Findings in this area suggest that mixing genders in group provided no significant detriment or benefit to programs as it relates to outcome. Yet, while non-significant, correlations were negative for programs serving females in both samples, suggesting that effect size did decrease when groups were mixed (albeit, not significantly so).

The next item is a proxy measure of dosage for programs. Here, the number of groups offered by the program was examined. This item was broken into two categories (0-6 groups, 7 or more groups). For male programs, there were significant positive correlations between number of groups and outcome in both samples (r=.21 and .25). For female programs, the relationship between number of treatment groups and outcome was limited. While this may suggest that dosage is a bigger concern for programs serving males than females, number of groups as a proxy measure of dosage is problematic in that the number and length of sessions are not factored into this measure. Likewise, since females were as a whole, lower risk, a non-significant relationship with an item that is proxy for dosage may make sense.

The final item examines use of graduated practice within programming. Role play provides offenders the opportunity to practice skills they are learning in the program. This strategy is built into many cognitive-behavioral curricula, such as Thinking for a Change, and as noted earlier, most programs incorporated some role play into group interventions. Graduated practice offers offenders additional opportunity to practice the skills learned in the program in increasingly difficult and life-like situations. This is oftentimes seen in advanced practice or booster sessions offered by a program. Programs serving males showed significant positive relationships between use of graduated practice and outcome in both samples, decreasing the average predicted probability of recidivism by up to 9 percentage points. For female programs, the correlations were not significant. This finding supports the RNR literature for male programs. Although there were not significant correlations for female programs, the direction of the relationship was still positive, suggesting use of a key CBT-technique was not detrimental to females, as suggested by some supporting gender-specific interventions.

General Program Treatment Characteristics

The final bivariate table (Table 4.7) examines general program treatment characteristics. This table contains a number of different characteristics that help describe application of the principles of effective intervention. There are 26 items in this section. Eight of the 26 highlighted items have limited variation, most of which are in the reinforcement and sanctioning section. On a scale from no use to liberal use of reinforcement and of sanctions, nearly all programs were marked with a moderate use of behavioral strategies, rendering prediction with these items difficult. Similarly, most programs did not use early release as an incentive or use traditional therapeutic community strategies as a sanction, but did remove passes as a means of punishment. Other items that lacked variation were use of a CBT model (most did), separating

Table 4.7: General Program Treatment Characteristics

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Use of a CBT model								
No	-.22 (16)	.25[c]	-.08 (3)	-.05	-.17 (18)	.31[b]	-.02 (3)	-.06
Yes	-.14 (60)		-.10 (21)		-.05 (58)		-.04 (20)	
Average length of stay								
Less than 4 months	-.19 (35)	.28[b]	-.07 (10)	-.15	-.13 (36)	.36[a]	-.03 (10)	.00
4 or more months	-.12 (64)		-.11 (27)		-.01 (63)		-.03 (26)	
Variation of treatment by risk level								
No	-.15 (73)	.13	-.12 (22)	.23	-.06 (69)	.05	-.03 (21)	.06
Yes	-.11 (25)		-.06 (14)		-.04 (29)		-.02 (15)	
Offenders separated by risk level								
No	-.15 (85)	.16	-.10 (31)	-.01	-.07 (85)	.28[b]	-.04 (31)	.17
Yes	-.09 (14)		-.10 (6)		.05 (13)		.01 (5)	
Cases assigned by caseload size								
No	-.15 (49)	.07	-.07 (17)	-.21	-.07 (50)	.07	-.02 (18)	-.08
Yes	-.13 (49)		-.12 (19)		-.05 (48)		-.04 (18)	
Community monitoring while on pass								
Ineffective to some monitoring	-.08 (25)	-.24[c]	-.10 (13)	.05	-.01 (25)	-.09	.00 (11)	-.18
Adequate to highly effective	-.14 (52)		-.09 (17)		-.04 (49)		-.03 (17)	

Table 4.7 Continued: General Program Treatment Characteristics

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Use of reinforcers and sanctions								
Use of reinforcers								
None or sparingly	-.13 (35)	-.04	-.09 (7)	-.04	-.06 (35)	.00	-.02 (8)	-.02
Moderate to liberal	-.14 (64)		-.10 (29)		-.06 (63)		-.03 (29)	
Use of early release as incentive								
No	-.15 (77)	.23[c]	-.11 (30)	.20	-.07 (77)	.12	-.04 (30)	.22
Yes	-.09 (21)		-.05 (7)		-.02 (21)		.02 (6)	
Offenders told why being reinforced								
No	-.09 (27)	-.26[c]	-.14 (12)	.30[d]	.02 (26)	-.29[b]	-.07 (12)	.27
Yes	-.16 (71)		-.07 (24)		-.08 (70)		-.01 (24)	
Use of punishers								
Sparingly	-.09 (11)	-.14	.00 (1)	-.10	-.04 (10)	-.04	-.01 (1)	-.03
Moderate to liberal	-.15 (87)		-.10 (36)		-.06 (89)		-.03 (36)	
Use of TC strategy as punishment								
No	-.13 (86)	-.27[b]	-.09 (31)	-.06	-.04 (84)	-.28[b]	-.02 (29)	-.15
Yes	-.23 (13)		-.11 (6)		-.17 (14)		-.06 (7)	
Removal of pass as a punishment								
No	-.22 (10)	.21[c]	-.18 (2)	.21	-.14 (10)	.18[d]	-.14 (3)	.31[d]
Yes	-.13 (89)		-.09 (34)		-.05 (88)		-.02 (34)	
Use of isolation as a punishment								
No	-.17 (67)	.32[a]	-.10 (21)	.08	-.08 (69)	.26[b]	-.04 (21)	.18
Yes	-.08 (32)		-.08 (15)		.01 (29)		-.01 (15)	
Punishers individualized								
No	-.15 (72)	.13	-.11 (28)	.28[d]	-.08 (73)	.22[c]	-.04 (27)	.25
Yes	-.11 (27)		-.04 (9)		.00 (26)		.01 (9)	
Reinforcement to punishment ratio								
Punishers outweigh	-.14 (77)	.11	-.10 (22)	-.01	-.05 (73)	-.03	-.03 (22)	.01
Reinforcers outweigh	-.11 (19)		-.10 (15)		-.06 (21)		-.03 (14)	

Table 4.7 Continued: General Program Treatment Characteristics

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Family Interventions								
Program overview only								
No	-.13 (72)	-.19[d]	-.08 (26)	-.24	-.04 (69)	-.19[d]	-.01 (25)	-.34[c]
Yes	-.18 (27)		-.14 (11)		-.10 (29)		-.08 (11)	
Family activities								
No	-.14 (80)	-.09	-.09 (29)	-.05	-.05 (78)	-.08	-.01 (29)	-.38[c]
Yes	-.16 (19)		-.11 (8)		-.08 (20)		-.10 (8)	
Family treatment intervention								
No	-.13 (75)	-.17	-.09 (24)	-.10	-.03 (72)	-.29[b]	-.01 (24)	-.24
Yes	-.18 (24)		-.11 (13)		-.13 (27)		-.06 (12)	
% of families involved in program								
Less than 40%	-.15 (29)	.09	-.10 (14)	.20	-.08 (30)	.02	-.07 (13)	.28
More than 40%	-.13 (12)		-.06 (8)		-.08 (14)		.00 (9)	
Aftercare								
External provider								
No	-.12 (56)	-.16	-.06 (20)	-.35[c]	-.02 (53)	-.28[b]	.01 (20)	-.44[b]
Yes	-.16 (43)		-.14 (17)		-.11 (45)		-.08 (16)	
Internal provider								
No	-.15 (77)	.09	-.10 (24)	.12	-.06 (76)	.04	-.03 (23)	-.03
Yes	-.12 (22)		-.08 (13)		-.05 (22)		-.03 (13)	
Self-help/AA only								
No	-.12 (58)	-.19[d]	-.12 (23)	.26	-.02 (55)	-.27[b]	-.04 (22)	.13
Yes	-.17 (37)		-.06 (14)		-.11 (40)		-.01 (14)	
Relapse prevention								
No	-.13 (65)	-.05	-.12 (17)	.24	-.05 (64)	.00	-.04 (16)	.09
Yes	-.15 (30)		-.07 (19)		-.05 (31)		-.02 (20)	

Table 4.7 Continued: General Program Treatment Characteristics

Variable	All Participants				Successful Completers			
	Males Mean ES (N)	r	Females Mean ES (N)	r	Males Mean ES (N)	r	Females Mean ES (N)	r
Case management								
No	-.13 (82)	-.08	-.10 (26)	.03	-.05 (80)	-.07	-.03 (24)	.09
Yes	-.16 (13)		-.09 (11)		-.08 (15)		-.02 (12)	
Aftercare length								
1 to 3 months	-.14 (24)	-.19	-.04 (9)	-.50[c]	-.04 (24)	-.31[d]	.02 (10)	-.54[c]
4 or more months	-.19 (11)		-.14 (8)		-.14 (12)		-.09 (8)	

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level

Highlighted items lack variation; findings related to these items are therefore not discussed

offenders by risk level (most did not), and use of case management during aftercare (few male programs did).

The first item to be examined is the program's average length of stay (LOS). Like number of groups, this item can also be considered a proxy measure for program dosage. Interestingly, for programs serving males, a longer LOS (4 or more months) produced significantly stronger treatment effects in both samples (r=.28 and .36), whereas the relationship between this item and outcome for female programs was not significant. This item provides some support for the risk principle, since males tended to be higher risk than females, and showed better results with a longer LOS.

The next item in Table 4.7 more specifically examines application of the risk principle. With regard to variation of treatment by risk level, offering more treatment hours to higher risk offenders was not significantly related to outcome for either males or females. However, all correlations were in a positive direction. This item offers limited support for the RNR literature.

The next item in this table looks at how offender cases are assigned to staff members. When cases are assigned by caseload size only (i.e., the next intake is assigned to the staff member with the least number of cases), there was not significant impact on outcome, although correlations were negative for programs serving females, suggesting failure to match counselor and client may be more harmful to female participants.

Community monitoring while on pass is the next item examined in Table 4.7. Here, staff and residents were asked to rate how effectively offender behavior was monitored while they were on community passes. Interestingly, male programs with limited monitoring had higher treatment effects than programs rated as having adequate to effective monitoring (in the full sample r=-.24). Yet, given that the outcome variable used in this study is any incarceration, the

negative effect is logical as effective community monitoring likely means that more offenders are caught engaging in antisocial or rule violating behaviors in the community, which may lead to more unsuccessful terminations and subsequent incarceration.

The next several items examine use of reinforcement and sanctioning strategies within a program. As indicated previously, several of these items (those highlighted) lacked sufficient variation to draw meaningful conclusions. Some items, however, can be discussed. Interestingly, in programs serving males, when offenders were told why they were being reinforced, treatment effects were significantly worse; however, when staff in female programs specified the purpose of a reinforcer, effects improved significantly (r=.30 in the full sample). These varying effects are difficult to explain, but may be a proxy measure for relationship building, which feminists may argue is particularly important in female programs.

The next set of variables relates to use of punishment or sanctions to correct negative behavior. There were significantly positive effects for male programs that used isolation as a sanction (r=.32 and .26); this item had a limited effect for females in programs. This finding is interesting, given that females tend to be more relational, and that effective sanctions should be aversive. The final two contingency management items examine how punishers are delivered, and the ratio of reinforcers to punishers. The first item looks at whether punishers are individualized. In other words, do programs use group punishers or are sanctions applied only to the person that engaged in the rule violation. Individualizing punishers so that they are applied only to those that engaged in the infraction was associated with positive treatment effects for males and females. Significant positive effects were noted for females in the full sample (r=.28) and males in the successful completer sample (r=.22). The final item in this subsection, the ratio or reinforcers to punishers, shows little relationship with outcome for either males or females.

The next set of items in Table 4.7 examines the use of family interventions in programs. These items are particularly important when examining gender differences due to the emphasis the gender responsive literature places on family involvement[51]. Sites that provided an overview of the program to families, but no actual intervention showed significant negative treatment effects for both the male and female populations. When family activities were offered (e.g. family day, picnics, and holiday parties), treatment effects were significantly worse for female programs in the successful completer sample (a 1% versus 10% increase in recidivism when such activities were offered). This item had no relationship with outcome for male programs. Why family activities would have a negative impact on recidivism for females is unclear. Interestingly, the application of a therapeutic intervention with family involvement did not improve effects. In fact, male programs in the successful completer sample had significantly worse outcomes when a family treatment intervention was applied (r=-.29). Negative effects may have been attributable to the type of family strategies employed.

The final family item examines the percentage of families involved in the program. Although there was a positive correlation for female programs that made the effort to involve at least forty percent of families in programming, the difference was not significant. There appeared to be little relationship between percentage of families involved and outcome for programs serving males. The bulk of family items support neither the gender responsive nor the RNR literature which suggests that family involvement is important. Results may have looked different if the sample included juvenile offenders.

The final set of variables in Table 4.7 address use of aftercare among the programs. Like with family interventions, results were somewhat mixed. When offenders were referred to

[51] While this need area is emphasized in the gender responsive literature, note that it is also a key crimingenic need according to the RNR followers.

external providers for aftercare, treatment effects for both male and female programs were significantly worse. When use of internal providers was examined, there was a limited relationship between this item and outcome, despite gender. When offenders were referred to Alcoholics Anonymous or Narcotics Anonymous only for aftercare, then treatment effects were significantly negative in each of the male samples. This item was not significantly correlated with outcome for female programs, but interestingly, correlations fell in a positive direction. Hence, while use of AA or NA may be useful for females, this is clearly not the case with males.

Incorporation of relapse prevention strategies during aftercare showed non-significant results for both male and female programs. Finally, the length of aftercare did appear to make a difference. In programs where aftercare was in excess of four months, treatment effects were significantly negative for males and females, showing at least a 10 percentage point increase in recidivism with lengthier aftercare. Like the monitoring item under treatment characteristics, this effect may be the result of more intensive supervision and treatment expectations for a longer period of time (resulting in more opportunity for probation or parole violations).

Given the number of treatment characteristics examined (96 variables in all) a summarization of the findings seemed prudent. Table 4.8 provides a duplication of variables described in Tables 4.1 through 4.7, summarizing variables in each of the four program areas: Leadership/Design, Staff, Assessment, and Treatment[52]. This table defines whether the relationship between each variable and incarceration is significant or not, and in what direction (i.e. a positive or negative relationship). It allows for a more succinct examination of the bivariate findings[53].

[52] Items with limited variation (those highlighted) are not included in the summarization tables.
[53] One general point of observation from the summary table is that far more programs serving males reached significance than programs serving females. While this may be attributable to a weak relationship between the variables of interest and outcome, it is also important to note that significance is more difficult to reach with smaller

Table 4.8: Summary Table of Bivariate Findings

Factor	Males Full	Females Full	Males Successful	Females Successful
Leadership/Design				
Program budget	Sig +	NS	Sig +	NS
Years of Operation	NS	Sig -	NS	NS
Staff/resident ratio	NS	NS	NS	NS
Support from community-at-large	Sig -	NS	NS	NS
PD clinical training	NS	Sig +	NS	NS
PD delivers services	NS	NS	Sig +	NS
Staff/management harmony	NS	NS	NS	NS
Co-ed living	Sig -	Sig -	Sig -	Sig -
Co-ed visitation	Sig -	NS	Sig -	Sig -
Staff Characteristics				
Belief treatment works	NS	Sig +	NS	NS
Any corrections experience	Sig -	NS	Sig -	NS
2 or more years of experience	Sig -	NS	Sig -	NS
Education	NS	Sig -	NS	Sig -
Clinical meeting	Sig +	NS	Sig +	NS
Clinical supervision	Sig +	NS	Sig +	NS
Soft skills evaluated	Sig -	NS	Sig -	NS
Hard skills evaluated	NS	NS	NS	NS
Commun/relationship skills evaluated	NS	NS	Sig +	NS
Treatment staff value treatment	NS	NS	NS	NS
Security staff value treatment	NS	NS	NS	NS
Staff turnover	NS	NS	NS	NS
Assessment				
Risk/need assessment	Sig +	NS	Sig +	NS
LSI-R	NS	NS	Sig +	NS
Substance abuse tool	NS	NS	Sig +	NS
Validated responsivity tools	NS	NS	NS	NS
Biopsychosocial tool	Sig -	NS	Sig -	NS
Mental health tool	NS	NS	NS	NS
PD selects offenders	Sig -	NS	Sig -	NS
Intake department selects offenders	Sig +	NS	Sig +	NS
Written exclusionary criteria	NS	NS	Sig +	NS
Excludes mental illness	NS	NS	NS	NS
Low risk above 20%	Sig -	NS	Sig -	NS
Number of assessments	NS	NS	NS	NS
Assessment training-observation only	Sig -	NS	NS	NS
Reassessment	Sig +	NS	NS	NS

sample sizes. This is because there is less confidence in inferring that the difference in effect size is beyond what might happen by chance with a smaller number of cases (i.e. programs). Hence, the trend in non-significant items for programs serving females is likely in part attributable to the small sample size. In particular, few items in the assessment domain reached significance for programs serving females.

Table 4.8 Continued: Summary Table of Bivariate Findings

Factor	Males Full	Females Full	Males Successful	Females Successful
Treatment Characteristics				
Targets				
Tier 1 criminogenic needs	NS	Sig -	NS	Sig -
Tier 2 criminogenic needs	NS	NS	NS	NS
Gender specific needs	NS	Sig -	NS	NS
Relationship with significant other	NS	NS	NS	NS
Low self esteem	NS	NS	NS	NS
Mental health	NS	NS	NS	NS
Economic/social needs	NS	NS	NS	NS
Parenting skills	NS	Sig -	NS	Sig -
Group interventions				
Thinking for a Change	Sig +	NS	Sig +	NS
Other cog group	NS	NS	NS	NS
Anger management	Sig +	NS	Sig +	NS
Employment/vocational education	Sig +	NS	Sig +	NS
Family	NS	NS	NS	NS
Life-skills	NS	NS	NS	NS
Eclectic	Sig -	NS	NS	NS
Assigned based on need	Sig +	NS	Sig +	NS
Mixed gender in groups	NS	NS	NS	NS
Number of groups provided	Sig +	NS	Sig +	NS
Use of graduated rehearsal	Sig +	NS	Sig +	NS
General treatment characteristics				
Average LOS	Sig +	NS	Sig +	NS
Variation by risk	NS	NS	NS	NS
Cases assigned by caseload size	NS	NS	NS	NS
Community monitoring	Sig -	NS	NS	NS
Told why reinforced	Sig -	Sig +	Sig -	NS
Isolation as punisher	Sig +	NS	Sig +	NS
Punishers individualized	NS	Sig +	Sig +	NS
Reinforcers outweigh punishers	NS	NS	NS	NS
Program overview for family	Sig -	NS	Sig -	Sig -
Family activities	NS	NS	NS	Sig -
Family treatment intervention	NS	NS	Sig -	NS
% family involvement	NS	NS	NS	NS
Aftercare by external provider	NS	Sig -	Sig -	Sig -
Aftercare by internal provider	NS	NS	NS	NS
Self-help/AA only as aftercare	Sig -	NS	Sig -	NS
Relapse prevention as aftercare	NS	NS	NS	NS
Aftercare longer than 4 months	NS	Sig -	Sig -	Sig -

NS=not significant (p>.10)
Sig += significant (p≤.10) in a positive direction
Sig - = significant (p≤.10) in a negative direction

Program Scales

Now that individual program characteristic data have been examined, aggregate scales will be reported based on the same four primary programming categories: Leadership/Design, Staff Characteristics, Assessment, and Treatment. Items were chosen for each of the scales based upon that item's relationship with outcome as well as the theoretical relevance of the item[54]. Significant relationship with outcome was not required, particularly since this was difficult to reach among the female samples due to low sample size. Instead, theoretical relevance as well as both the direction of the relationship and proximity to significance was considered.

All dichotomous items were coded as 0 or 1, and categorical variables were coded as 0, 1 or 2. Items were not further weighted based on the strength of their relationship with recidivism. Instead, the value of the item in the scale is based on the above coding. Separate scales were created in each domain for males and females based upon what items appeared theoretically and empirically relevant. The full scales for each programming domain were then broken down into categories, which demonstrate the cumulative predictability of each cluster of related characteristics to incarceration. Correlations between each of the four scales plus the overall scale and recidivism will be reported. Finally, the correlations for each scale will be examined by state.

[54] Some items that were both predictive and theoretically relevant could not be included due to missing data. In developing the scales, missing data in one item meant that no other data on items for that case could be included in the scale. Therefore, items with more than 10% missing data were excluded.

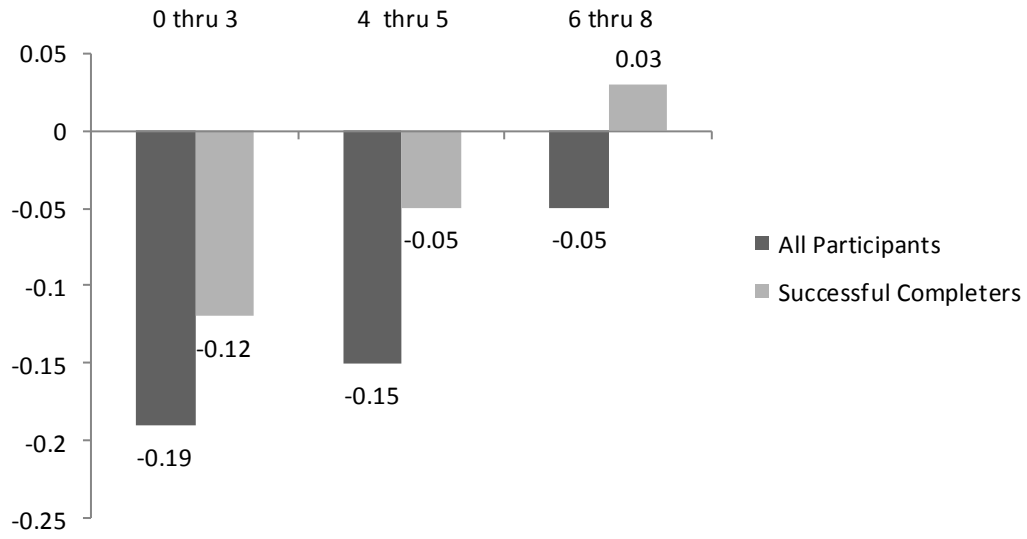
Figures 4.7 and 4.8 present the Leadership and Design scales. Figure 4.3 is the Leadership and Design scale developed for male programs. Program items incorporated in the male scale include[55]:

1) Program director delivers clinical training;
2) Program director delivers services;
3) Co-ed programs have no shared living space;
4) Co-ed programs have no shared education;*
5) Co-ed programs have no shared visitation.[56]

This scale had a range of 0 to 8. Figure 4.3 shows that in the all participants sample, the average effect size for programs scoring 0 through 3 was -.19 percent (N=30), -.15 for programs scoring 4 through 5 (N=51), and -.05 for programs scoring highest (N=18). These differences were significant. For the successful completer sample, there was also a significant difference between category effect sizes. Male programs scoring between 0 and 3 had an effect size of -.12 (N=32), programs scoring mid-range had a treatment effect of -.05 (N=48), and programs scoring 6 through 8 showed a 3 percent improvement in recidivism (N=19). Overall, the culmination of items in the Leadership and Design scale demonstrate significant improvement in treatment effects for male programs in both the all participant and successful completer samples. However, Cronbach's alphas for the 5 leadership and design items for males were .53 for the full sample and .58 for the successful completer sample, which suggests poor internal consistency for this scale.

[55] A full description of these items are available in the previous results section under Bivariate Results—Leadership and Design Characteristics.
*This item lacked variation, but was included in the scale as a theoretically relevant variable.
[56] The Co-ed items were reverse coded since they were significantly negatively correlated with recidivism. Because these items were only coded for programs that served both males and females, a 3 point scale was developed for each of the three items: 0=co-ed with shared living (education or visitation), 1=not co-ed, 2=co-ed with NO shared living (education or visitation). While program budget was significantly correlated, it was not included in the scale due to lack of theoretical relevance as well as missing data.

Figure 4.7: Leadership and Design Scale—Males



All Participants: F=7.250; p=.001; 5 items; $\alpha = .53$
Successful Completers: F=5.781; p=.004; 5 items; $\alpha = .58$

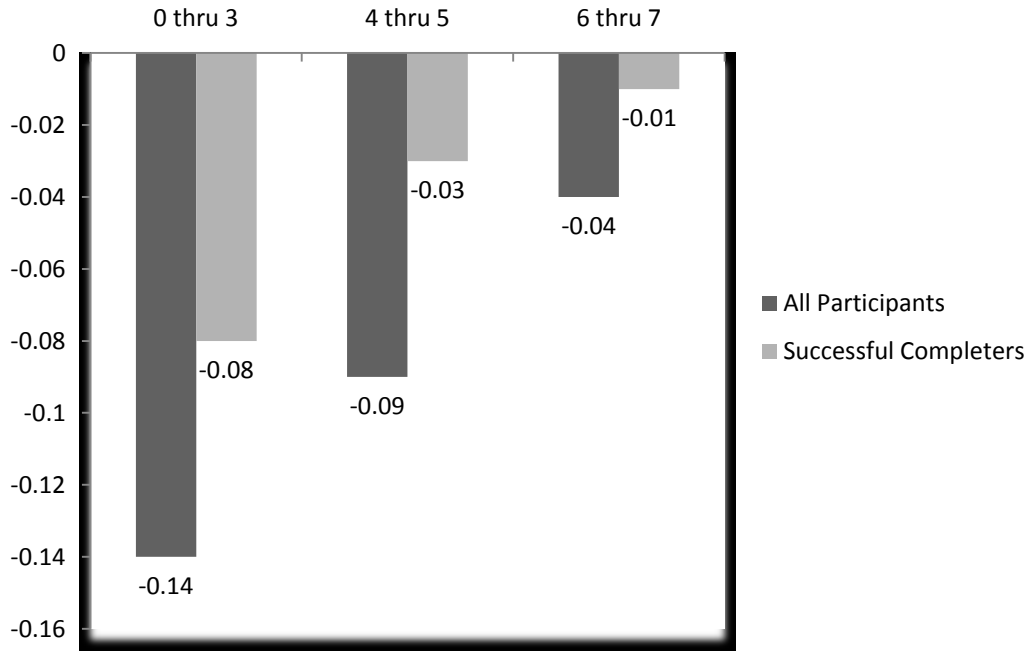
Figure 4.8 examines Leadership and Design characteristics for female programs. Items included in the female scale are:

1) Program director delivers clinical training;
2) Co-ed programs have no shared living space;
3) Co-ed programs have no shared education;*
4) Co-ed programs have no shared visitation.[57]

In comparison to the male leadership scale, item 2 on the male scale was omitted (program director delivers services); all other items are the same. The possible range of scores on this scale was 0 to 7, with actual scores also falling between 0 and 7. Three categories of scores were developed for this scale. Figure 4.4 shows that for the full sample, female programs scoring 3 or fewer points had an effect size of -.14 (N=14), programs scoring 4-5 points had an average effect size of -.09, and programs scoring above 6 to 7 had a 4 percent increase in the predicted probability of recidivism (N=9). This difference between categories was not significant (p=.133). The successful completer sample also showed some improvement with increased score; programs scoring 0 to 3 had an average effect size of -.08 (n=7), programs scoring between 4 and 5 had a -.03 predicted probability of recidivism (n=18), and programs scoring 6 or more had nearly the same likelihood of recidivism as the comparison sample (ES=-.01; n=12). Differences, again, were not significant (p=.345). Cronbach's alphas for the 5 leadership and design items for females were .65 for the full sample and .66 for the successful completer sample, which suggests limited internal consistency for this scale.

*This item lacked variation, but was included in the scale as a theoretically relevant variable.
[57] The same co-ed scales were used for the female programs as males.

Figure 4.8: Leadership/Design Scale--Females



All Participants: F=2.149; p=.133; 5 items; $\alpha = .66$
Successful Completers: F=1.098; p=.345; 5 items; $\alpha = .66$

Figures 4.9 and 4.10 describe Staff Characteristics. The scale for male programs consists of the following items:[58]

1) Training on the theory and philosophy of the program;*
2) Clinical meetings with positive attributes;
3) Evaluation of communication/relationship skills; and
4) Clinical supervision by a licensed social worker or counselor;

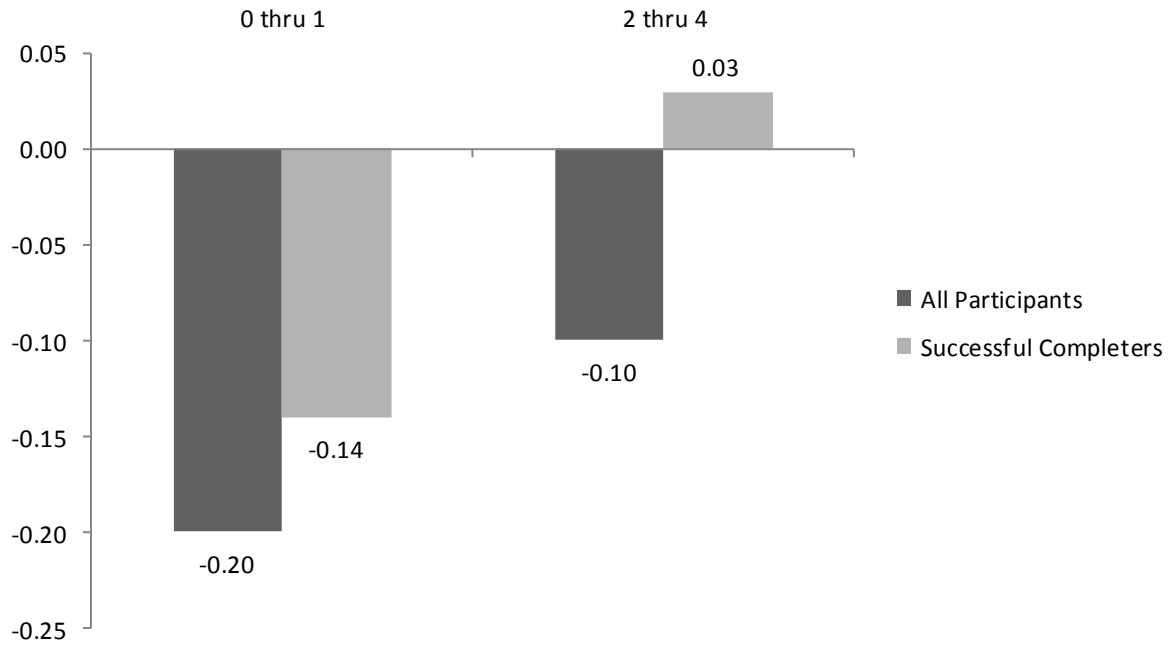
This scale had a possible range of 0 to 4. In the all participants sample, the average treatment effect for programs scoring 0 through 1 was -.20 percent (N=44). This dropped by half when the score in this section exceeded 1 point. For the successful completer sample, Figure 4.5 shows that the average effect size was -.14 for programs scoring 0 to 1 (N=48), with an decrease in recidivism by three percent for programs scoring 2 to 4 points. Differences in recidivism rates were significant for both the full and successful completer samples. Cronbach's alphas for the 5 staff items for males were .36 for the full sample and .40 for the successful completer sample, which suggests poor internal consistency for this scale.

Figure 4.10 depicts the Staff Characteristics scale for the female programs. The scale includes the following items:

1) Seeks staff that believes treatment works;
2) Training on the theory and philosophy of the program;*
3) Evaluation of communication/relationship skills;
4) Clinical supervision by a licensed social worker or counselor; and
5) Treatment staff value treatment efforts.

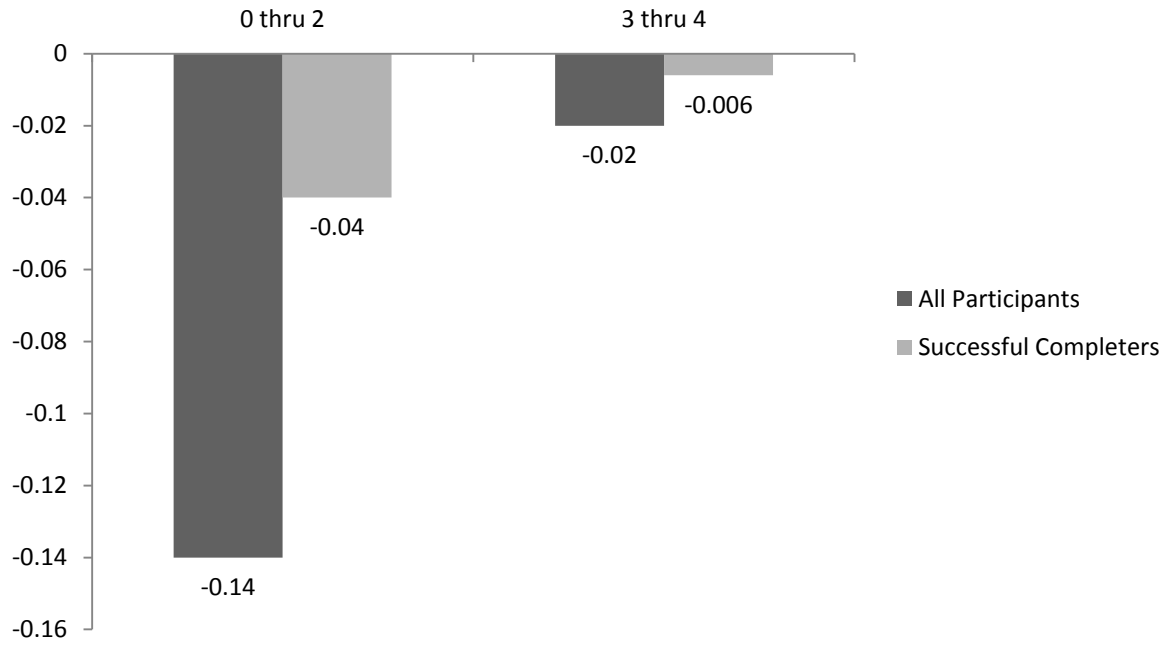
[58] A full description of these characteristics can be found in the Bivariate Analyses—Staff Characteristics section.
* These items lacked variation, but were included in the scale as theoretically relevant variables.

Figure 4.9: Staff Characteristics Scale--Males



All Participants: F=17.633; p=.000; 4 items; $\alpha = .36$
Successful Completers: F=35.196; p=.000; 5 items; $\alpha = .40$

Figure 4.10: Staff Characteristics Scale--Females



All Participants: F=10.907; p=.002; 5 items; $\alpha = .20$
Successful Completers: F=1.106; p=.300; 5 items; $\alpha = .23$

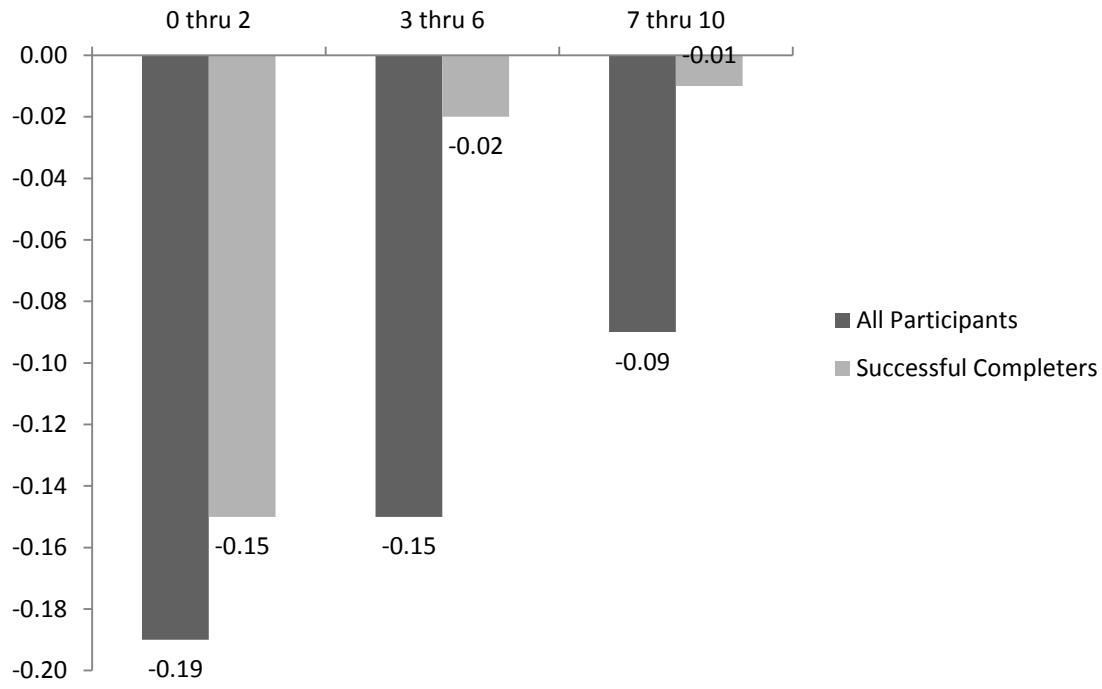
Compared to the same scale for males, the female scale adds two items, and omits item 2 on the male scale; all other items are the same. The added items are related to staff believing in and valuing treatment. This scale has a possible range of 0 thru 5, with the actual range being 0 through 4. In the full sample, when a program scored between 0 and 2 on this scale, the average treatment effect was -.14 (N=23). When a program scored 3 through 4, the average treatment effect increased to -.02 (N=14). In the successful completer sample, the treatment effect for programs scoring in the lower range was -.04 percent (N=22); this effect size is null for programs scoring above 2 (N=14). Differences for the full sample were significant, but not the successful completer sample. Cronbach's alphas for the 5 staff items for females were .20 for the full sample and .23 for the successful completer sample, which suggests very poor internal consistency for this scale.

The next scale to be examined is Assessment. Figure 4.11 depicts the assessment scale for male programs. Items in this scale include[59]:

1) Use of a risk/need assessment;
2) Risk assessed using the LSI-R;
3) Use of a validated substance abuse tool;
4) Use of other validated need assessment tools*;
5) Assessment of criminal attitude*;
6) Offenders selected for the program by an intake department;
7) Written exclusionary criteria in place;
8) Low proportion of the population low risk; and
9) Reassessment conducted.

[59] A full description of these characteristics can be found in the Bivariate Analyses—Assessment section.
* These items lacked variation, but were included in the scale as theoretically relevant variables.

Figure 4.11: Assessment Scale--Males



All Participants: F=5.193; p=.007; 9 items; α =.75
Successful Completers: F=6.699; p=.002; 9 items; α =.78

The range of scores in this scale is 0 to 9. Figure 4.11 shows that in the full sample, programs scoring between 0 and 2 in the assessment category had an average recidivism rate of -.19 (N=20). Programs scoring 3 through 6 increased recidivism by an average of 15 percent (N=54), and programs scoring the highest dropped this rate to -.07 (N=23). In the successful completer sample, recidivism rates ranged from -.15 for the lowest scores (N=23), -.02 for scores between 3 and 6 (N=47), and a rate of -.01 for programs scoring in the highest category (N=25). Differences in recidivism rates between categories are significant in both the full and successful completer categories. Cronbach's alphas for the 10 assessment items were .75 for the full sample and .78 for the successful completer sample, which suggests adequate internal consistency for this scale.

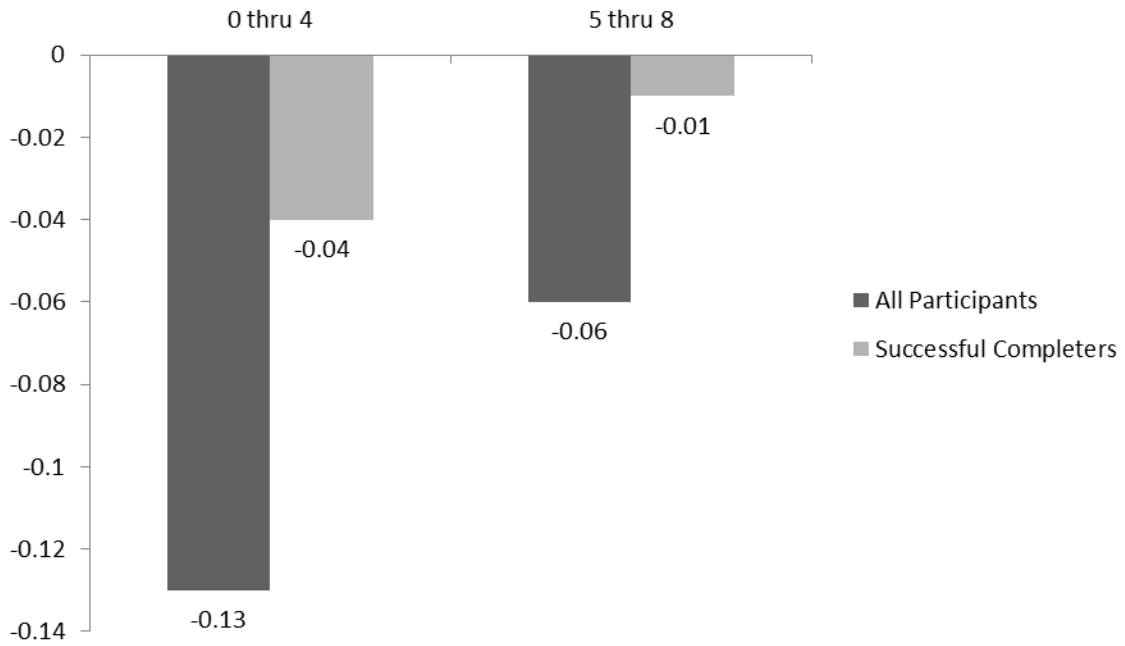
Figure 4.12 examines Assessment items for female programs. Items included in the female scale were as follows:

1) Use of a risk/need assessment;
2) Risk assessed using the LSI-R;
3) Use of other validated need assessment tools;[60]*
4) Assessment of criminal attitude;*
5) Assessment of mental health;
6) Offenders selected for the program by an intake department;
7) Written exclusionary criteria in place; and
8) Reassessment conducted.

All items are the same as the male assessment scale except use of a substance abuse tool, number of assessments conducted, and proportion of low risk items were omitted, and

[60] * These items lacked variation, but were included in the scale as theoretically relevant variables.

Figure 4.12: Assessment Scale--Females



All Participants: F=2.756; p=.107; 8 items; α =.77
Successful Completers: F=.525; p=.474; 8 items; α =.78

assessment of mental health was added. The female assessment scale has a range between 0 and 8. Figure 8 shows that programs falling between 0 and 4 had an average effect size of -.13 (N=18) in the all participants sample, with a drop to -.06 when the assessment score rose to 5 or more (N=16). In the successful completer sample, a 0 to 4 produced an average effect size of -.04 (N=17), which increased to -.01 with a score of 5 to 8 (N=17). These differences in effect sizes were not significant. Cronbach's alphas for the 8 assessment items for females were .77 for the full sample and .78 for the successful completer sample, which suggests adequate internal consistency for this scale.

The next scale to be examined is the largest—Treatment Characteristics. This scale combines items from Table 4.5*: Treatment Targets*, Table 4.6: *Group Interventions*, and Table 4.7: *General Program Characteristics*. Figure 4.13 depicts the Treatment Characteristics scale for the male programs. Items included in this scale are[61]:

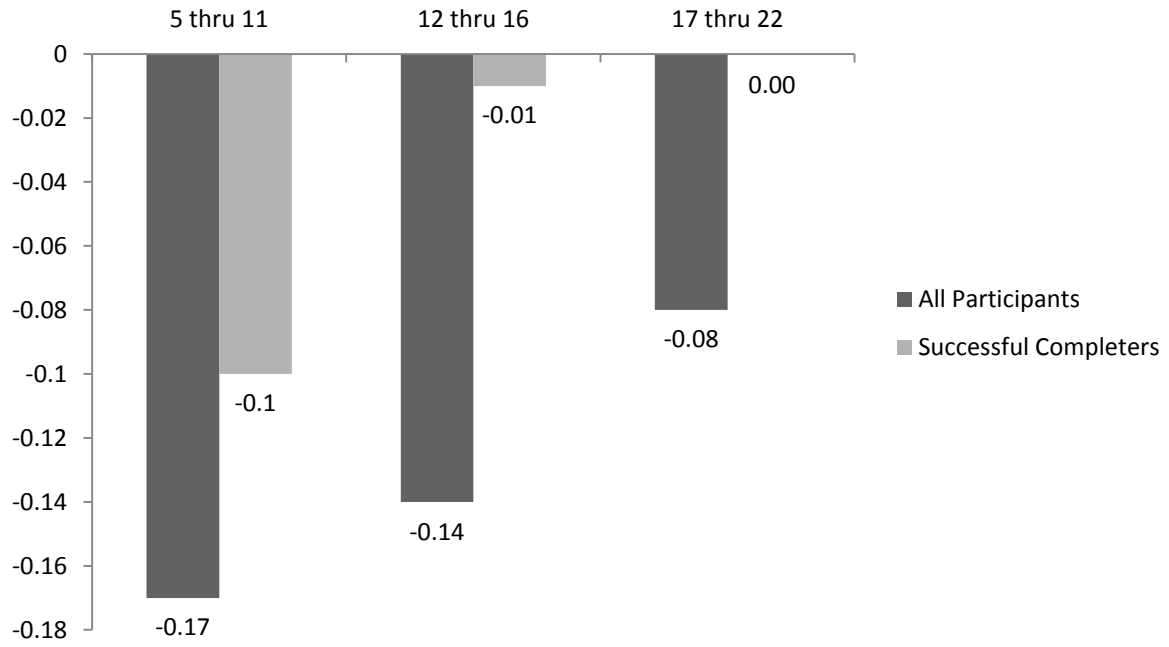
1) Tier 1 criminogenic needs targeted;
2) Tier 2 criminogenic needs targeted;
3) Drug/alcohol problems targeted;*
4) Vocational achievement targeted;*
5) Educational achievement targeted;*
6) Trauma/PTSD targeted;*
7) Thinking for a Change group provided;
8) Substance abuse group provided;*
9) Anger management group provided;
10) Employment or vocational education group provided;
11) Family group provided;

[61] A full description of these characteristics can be found in the Bivariate Analyses—Treatment Characteristics section.
* These items lacked variation, but were included in the scale as theoretically relevant variables.

12) Mental health group provided;*
13) Groups assigned based on need;
14) Program uses a structured curriculum;*
15) Use of role play;*
16) Use of graduated practice;
17) Use of a CBT model*
18) Higher number of groups provided;
19) Longer average length of stay;
20) Variation of treatment by risk level;
21) Offenders separated by risk level*;
22) Early release used as an incentive*;
23) Use of isolation as a punishment;
24) Removal of a pass as a punishment*;
25) Punishers individualized;
26) Aftercare provided internally; and
27) Aftercare does not consist of AA/NA only.

The range on this scale is 0 to 27, with actual scores falling between 5 and 22 points. As seen in Figure 4.13, for the full sample, programs scoring between 5 and 11 had an average recidivism rate of -.17 (N=29). Programs scoring 12 through 16 had an effect size averaging -.14 (N=31). Finally, programs scoring 17 or more had a treatment effect of -.08 (N=25). In the successful completer sample, programs scoring lowest averaged a 10 percent increase in recidivism (N=30), programs scoring 12 through 16 had a rate of -.01 (N=26), and programs scoring highest showed a null effect (N=28). These differences were significant for both the full and successful completer samples. Cronbach's alphas for the 27 treatment items for males were .76 for the full sample and .80 for the successful completer sample, which suggests adequate to good internal consistency for this scale.

Figure 4.13: Treatment Characteristics Scale--Males



All Participants: F=4.375; p=.016; 27 items; α =.76
Successful Completers: F=4.913; p=.010; 27 items; α =.80

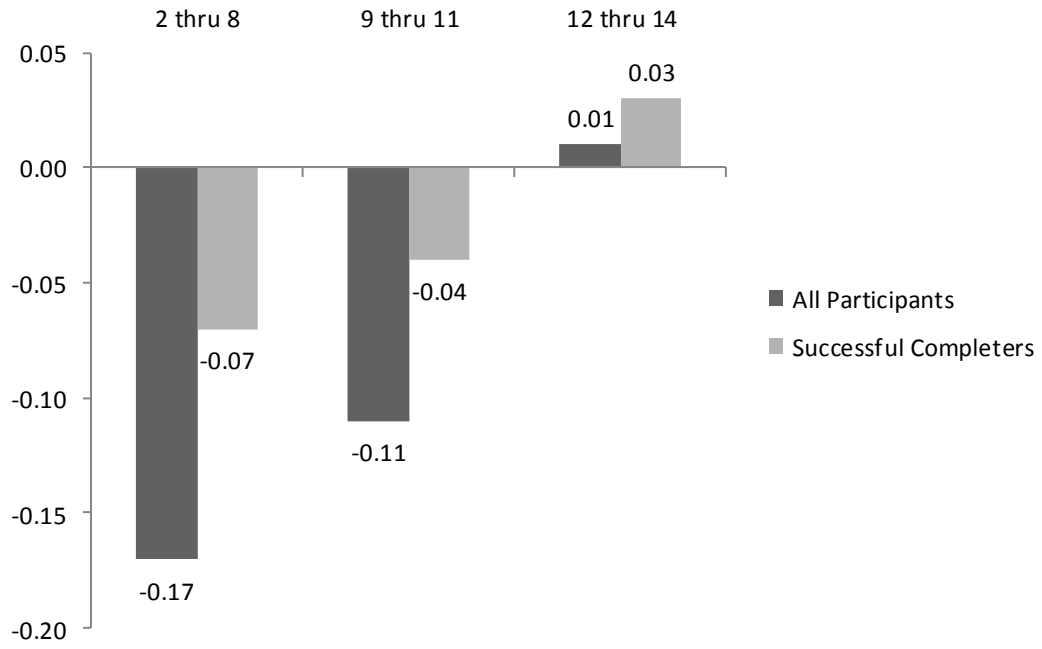
The next figure (Figure 4.14) shows the Treatment Characteristics scale for female programs. Items in this scale include:

1) Vocational achievement targeted;*
2) Educational achievement targeted;*
3) Family affection and communication targeted;*
4) Family problem solving targeted;*
5) Mental health group;*
6) Life skills group;
7) Groups assigned based on need;
8) Use of role play;*
9) Use of graduated practice;
10) Variation of treatment by risk level;
11) Offenders separated by risk level;*
12) Early release used as an incentive;*
13) Offenders are told why they are being reinforced;
14) Use of isolation as a punishment;
15) Removal of a pass as a punishment;*
16) Punishers individualized;
17) Aftercare provided internally;
18) Use of AA/NA for aftercare
19) Relapse prevention provided for aftercare; and
20) Case management provided as part of aftercare.*[62]

Similar items between the male and female scales include: targeting vocational and educational development, offering a mental health group, assigning to group based on need, use of role play and graduated practice, variation and separation by risk, use of early release as an incentive and isolation and pass removal as a punisher, individualizing punishers, and providing

[62] * These items lacked variation, but were included in the scale as theoretically relevant variables.

Figure 4.14: Treatment Characteristics--Females



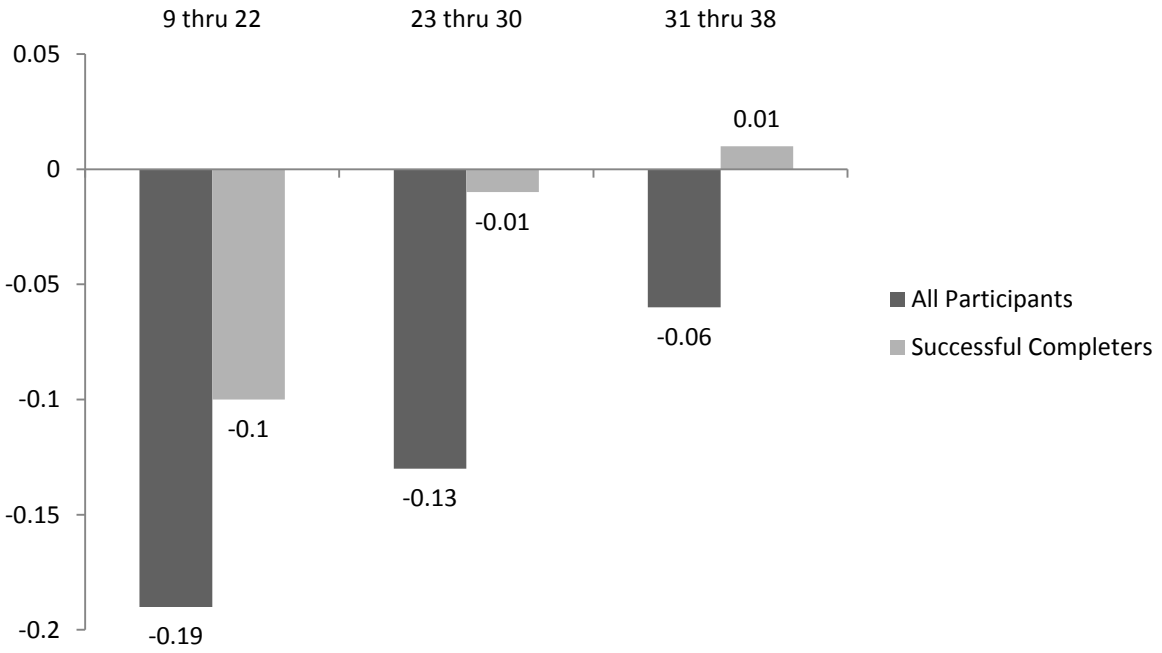
All Participants: F=9.794; p=.000; 20 items; α =.56
Successful Completers: F=2.678; p=.084; 20 items; α =.58

aftercare internally. Other items differ between the male and female treatment scales. The range on the female scale is 0 to 20, with actual scores falling between 2 and 14 points.

Figure 4.14 shows that in the all participants sample, programs with scores falling between 2 and 8 have an average effect size of -.17 (N=10). Programs that have a score falling between 9 and 11 increase their effect size to -.11 (N=18). Finally, programs with scores of 12 or more have an average .01 effect size (n=8). In the successful completer sample, programs with scores falling in the low range average a 7 percent increase in recidivism (N=10); programs falling in the mid-range average a -.04 percent effect size (N=17), and programs in the highest range decrease recidivism an average of 3 percent (N=9). Differences in both the samples were significant. Cronbach's alphas for the 20 treatment items for females were .56 for the full sample and .58 for the successful completer sample, which suggests poor internal consistency for this scale.

Finally the overall scales are presented for the male and female programs. These scales are simply a combination of the above four scales: Leadership/Design, Staff, Assessment and Treatment. Figure 4.15 depicts the overall scale for male programs. The possible range on this scale is 0 to 48, with programs scoring between 8 and 38 points. As expected, the same pattern of improved effect size with increased score can be found. In the full sample, programs scoring 22 or lower have an average treatment effect of -.19 (N=28). Programs scoring 23 through 30 have a treatment effect that averages -.13 (N=28). Programs scoring above 30 average a 6 percent increase in recidivism (N=20). In the successful completer sample, program scoring below 23 increase recidivism by an average of 10 percent (N=27), programs scoring mid-range have a -.01 average treatment effect (N=24), and program scoring highest decrease recidivism by 1 percent on average (N=23). These differences are significant. Cronbach's alphas for the 45

Figure 4.15: Overall Scale--Males



All Participants: F=8.973; p=.000; 45 items; α =.84
Successful Completers: F=5.039; p=.009; 45 items; α =.86

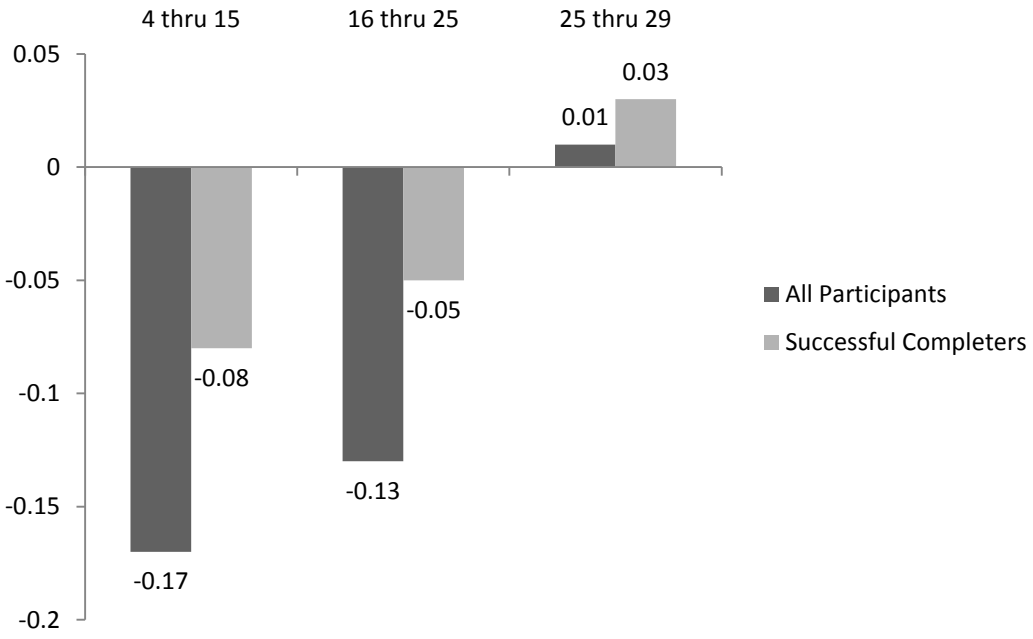
overall items for males were .86 for the full sample and .86 for the successful completer sample, which suggests good internal consistency for this scale.

The final Figure (Figure 4.16) demonstrates the overall scale for the female sample. The range on this scale is 0 to 40 with actual program scores falling between 4 and 29. In the full sample, programs increased recidivism by an average of 17 percent with scores between 4 and 15 (N=9); programs had an average treatment effect of -.13 with scores between 16 and 24 (N=14), and programs 25 or above had a one percent decrease in recidivism, on average (N=11). In the successful completer sample, programs in the low range have a -.08 average effect size (N=9); programs mid-range produce a -.05 percent treatment effect (N=12), and programs scoring in the highest range decrease recidivism by an average of 3 percent (N=12). Differences are significant in both the full and successful completer samples. Cronbach's alphas for the 37 overall items for females were .81 for the full sample and .82 for the successful completer sample, which suggests good internal consistency for this scale.

Table 4.9 examines the Pearson r values between each of the full scales and outcome in male and female programs for both the full and successful completer samples[63]. In the all participant male sample, correlations for each of the scales range between .34 and .43, with an overall correlation of .53. The female all participant scale produces similar correlation coefficients, with assessment being slightly lower, and staff and treatment being higher. The overall correlation for the female programs is similar to the male at .58. For successful completers, the male correlations ranged between .38 and .44, with the overall correlation falling at .50. The female staff and assessment correlations dropped out of significance, but the overall female correlation was still significant at .46.

[63] The full scale as opposed to the categorical scale was used to calculate the Pearson r values.

Figure 4.16: Overall Scale—Females



All Participants: F=13.159; p=.000; 37 items; α =.81
Successful Completers: F=3.447; p=.045; 37 items; α =.82

Table 4.9: Scale Correlations by Gender

Variable	All Participants		Successful Completers	
	Males r (N)	Females r (N)	Males r (N)	Females r (N)
Leadership/Design Scale	.38 (99)[a]	.36 (37)[c]	.38 (98)[a]	.36 (36)[c]
Staff Characteristics Scale	.37 (88)[a]	.45 (37)[b]	.44 (88)[a]	.18 (36)
Assessment Scale	.34 (96)[a]	.29 (34)[d]	.38 (95)[a]	.23 (34)
Treatment Scale	.43 (85)[a]	.56 (36)[a]	.42 (84)[a]	.48 (36)[b]
Overall Scale	.53 (78)[a]	.58 (34)[a]	.50 (77)[a]	.46 (33)[b]

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level

Finally, Table 4.10 examines correlations for each sample by state. Note that Ohio represents a larger proportion of the programs, particularly the female programs. The correlations for the overall scale remained fairly robust, other than the Ohio successful completer male sample, where the overall correlation was negative. Other overall correlations ranged between .26 and .76. There was, however, increased variability in correlations within each scale. In Ohio, most of the female scales continued to be strongly positively correlated with outcome, which is not surprising given they make up the bulk of the programs representing females. Other than the Staff Characteristics scale, scales appeared to hold up better for Pennsylvania male programs than Ohio male programs, particularly the leadership and treatment scales. Overall, the treatment scale appeared to most consistently predict outcome. Nonetheless, the variation on Pearson's r values by state suggest that results may have limited generalizability without further study.

Results Summary

This section attempted to answer how program treatment characteristics vary by program gender. Four primary domains were examined: Leadership and Design, Staff, Assessment and Treatment Characteristics. The bivariate relationship between the identified program characteristic and program effect size was studied by gender to determine what characteristics were important for one gender versus another. While there was variation in items that were predictive for one gender over another, for most variables, the direction of the relationship was similar, which as a whole supports the generalist perspective. Nonetheless, several gender specific items were explored, some of which showed results in support of the feminist literature. Hence, results are perhaps most supportive of those that merge the gender specific/RNR perspectives.

Table 4.10: Scale Correlations by Gender and State

Variable	OHIO				PENNSYLVANIA			
	All Participants		Successful Completers		All Participants		Successful Completers	
	Males r (N)	Females r (N)	Males r (N)	Females r (N)	Males r (N)	Females r (N)	Males r (N)	Females r (N)
Leadership/Design Scale	.10 (69)	.43 (32)[c]	-.16 (62)	.38 (31)[c]	.24 (30)	.09 (5)	.26 (36)	.09 (6)
Staff Characteristics Scale	.18 (61)	.57 (32)[a]	.16 (55)	.15 (31)	-.22 (27)	-.05 (5)	-.22 (33)	.00 (6)
Assessment Scale	.06 (69)	.37 (29)[c]	-.21 (62)	.15 (28)	.10 (27)	.05 (5)	.13 (33)	.05 (5)
Treatment Scale	.11 (65)	.68 (32)[a]	-.24 (59)[d]	.46 (31)[b]	.37 (21)	.24 (5)	.37 (26)[d]	.24 (5)
Overall Scale	.26 (59)[c]	.76 (29)[a]	-.14 (54)	.45 (28)[c]	.34 (19)	.30 (4)	.35 (23)[d]	.28 (5)

[a] p≤.001 level; [b] p≤.01 level; [c] p≤.05 level; [d] p≤.10 level

Scales were also developed to examine the aggregate contribution of program factors on recidivism. Separate scales were developed for males and females, as only those items positively related to outcome and/or theoretically relevant (either based on the RNR or gender-specific literature) were selected for each scale. Categories were developed for the scales and most showed a significant improvement in treatment effect as scores on the scale increased. The overall scale was also significantly related to improved incarceration rates for males and females. Finally, correlations between the scales and outcome were examined by state. While most of the overall scales were positively related to outcome, not all differences were significant. Also some of the individual scales became negatively correlated to outcome when examined by state. A possible explanation for this is that generally Ohio programs had better treatment effects than Pennsylvania programs, so together there was increased variation, which allows for improved predictability. Once scales were examined by state, the variation decreased and some scales failed to predict outcome. It also suggests that further studies are needed as generalizability comes into question. The final section of the dissertation will summarize the data presented in relation to the research questions posed. Also limitations will be discussed as well as implications on correctional policy.

CHAPTER 5
DISCUSSION

Introduction

Although men still dominate those served in the criminal justice system, there has been an influx of female offenders over the past two decades (BJS, 2010). Meanwhile, there is considerable debate in corrections about how treatment strategies should or should not vary according to a participant's gender. Researchers that follow the Canadian school of thought (Risk, Need and Responsivity Principles--RNR) view gender as a responsivity factor, arguing that while some adaptations might be made for females, the primary RNR principles should be followed (see Andrews and Bonta, 2010). Many feminists, on the other hand, argue that a female's pathway to crime is unique and that treatment should vary accordingly (see Bloom et al. 2003). Quantitative studies have been conducted to determine the applicability of RNR to females (see Blanchette and Brown, 2006). Nonetheless, there have been a limited number of comprehensive studies on program characteristics that are applied to both male and female offenders where differences in outcome can be examined. This dissertation empirically examines key tenants from both schools of thought to determine what specific treatment strategies appear important for each gender.

In exploring the question of how treatment factors might vary by gender, 138 programs were examined across two states; 99 of these programs served males and 39 programs served females. Effect sizes were calculated for each of these programs using a sample of 27,008 treatment and matched comparison cases. Treatment cases were matched with a non-treatment comparison group of parolees based on the following variables: gender, race, sex offender status, county size, and risk. Effect sizes were measured by the predicted probability of

incarceration. Effect sizes were calculated based on both a full sample of program participants and a successful completer sub-sample. Program characteristics were gathered via comprehensive evaluations of each of the sites, consisting of interviews with staff and participants, observation of treatment, and review of program files and material. Characteristics debated as important for correctional programming among the RNR and gender-specific literature were selected for analyses. Likewise, characteristics that support foundational theories from both camps (e.g., cognitive behavioral theory and relational theory) were included as well.

Bivariate analyses were conducted to determine which of the 96 selected treatment characteristics were associated with an improved effect size. Males and females were examined separately so that treatment effects could be compared. Bivariate analyses were used to develop scales in four program domains: Leadership/Design, Staff, Assessment and Treatment Characteristics. These domains are based on previous program evaluation literature (see Lowenkamp 2004). This chapter will summarize the primary findings from this research, discuss the limitations of the study, and review implications for both policy and future research.

Summary of Findings

Bivariate Results

Findings will be summarized by categorizing bivariate results according to each of the research questions posed. This will allow for a discussion of how effective program characteristics vary by gender. The research questions ask what program characteristics are 1) important for *both* males and females; 2) *More* important for one gender over another; 3) *Only* important for one gender over another; and 4) important for *neither* gender.

Factors fell into the 1) "Important for *both* males and females" when findings for the variable were significant in both programs serving males and those serving females. This could

pertain to either the full or successful completer sample. Factors determined 2) "*More* important" for one gender over another were determined when a significant difference in effect size was noted for only one gender, but the direction of the relationship was the same for the other gender. Characteristics were determined 3) "*Only* important" for one of the genders when a significant difference in effect size was noted for only one gender, and there was either no relationship or the direction of the relationship was opposite for the other gender. Finally, factors were determined 4) "important for *neither* males nor females" when significant differences were not found in any group. Note that for some items, while differences were significant for one gender group and not the other, that factor may have lacked variation for that group, affecting predictability of the item. Hence, items that did not significantly predict incarceration but had limited variation are noted with an asterisk.

RESEARCH QUESTION 1: What program characteristics are important for **both** adult **male and female** offenders?

--Characteristics significantly correlated with reduced recidivism for programs serving males and females:

- Avoid co-ed living units, educational programming*, and visitation time
- Use validated need assessment tools (i.e. tools that assess specific criminogenic needs)*
- Offer a group that targets mental health*
- Use of pass removal as a punishment/sanction*
- Ensure that punishers are individualized (i.e. avoiding group punishers)
- Avoid providing a program overview as the only source of family involvement
- Avoid the use of external providers for aftercare
- Avoid aftercare that exceeds 4 months

RESEARCH QUESTION 2: What program characteristics are **more** important for adult **male** offenders?

--Characteristics significantly correlated with reduced recidivism for programs serving males; improved effect size for programs serving females, but difference did not reach significance:

- The program has a higher program budget
- Individuals with experience in corrections/criminal justice are *not* sought as new employees
- Clinical meetings occur regularly, are multidisciplinary, and are spent reviewing cases
- Clinical supervision is conducted by a licensed social worker or counselor
- The bulk of employee evaluation items do not include soft skills such as creativity, openness to supervision and being a team player
- Evaluations *do* include communication and relationship skills such as crisis de-escalation skills, boundaries, avoiding negative interactions with clients, firm but fair, and empathic*
- A risk need assessment is in place*
- Use of the Level of Service Inventory Revised to assess risk and need level
- Avoid use of a biopsychosocial tool
- Avoid assessment of past abuse issues*
- Intake decisions are made by an intake coordinator or team rather than exclusively by the program director
- Written exclusionary criteria are in place
- Assessment training is *not* conducted by observation only
- Reassessment of risk and/or need areas are conducted
- Vocational achievement is targeted*
- Family affection and communication is targeted
- An eclectic group (e.g. process group) is *not* offered
- Groups are assigned based on need levels
- The program uses a structured curriculum*
- A higher number of groups are provided
- Use of role play* and graduated rehearsal of skills
- Offenders are separated by risk level*
- There is *not* intensive community monitoring while on pass
- Early release is used as a reinforcer/incentive for positive behavior*
- Isolation is used as punishment/sanction, but traditional therapeutic community strategies are avoided*

RESEARCH QUESTION 3: What program characteristics are **only** important for adult **male** offenders?

--Characteristics correlated with reduced recidivism for only programs serving males:

- The program director is involved in service delivery
- The bulk of employees are not seasoned in the corrections field (i.e., do *not* have 2 or more years of correctional experience)
- A validated substance abuse tool is used
- A small proportion of low risk offenders are served (less than 20%)
- Drug and alcohol issues are targeted*
- Family affection and communication is *not* targeted*
- Trauma/PTSD is targeted, where needed
- The Thinking for a Change curriculum is offered
- A substance abuse group is offered*
- An anger management group is offered
- A vocational/educational group is offered
- A cognitive-behavioral model is the primary treatment model*
- The program has over a 4 month length of stay
- Family treatment interventions are provided
- Avoid use of AA/NA or other self-help programs as the only source of aftercare

RESEARCH QUESTION 4: What program characteristics are **more** important for adult **female** offenders?

--Characteristics significantly correlated with reduced recidivism for programs serving females; improved effect size for programs serving males also, but difference did not reach significance:

- The program director delivers clinical training
- New staff are trained on the theory and philosophy of the program*
- An assessment tool that measures criminal attitude is used*

RESEARCH QUESTION 5: What program characteristics are **only** important for adult **female** offenders?

--Characteristics correlated with reduced recidivism for only programs serving females:

- The program has been in operation no more than 15 years
- The program seeks staff that believe treatment works
- The bulk of treatment staff do *not* have higher education levels
- Tier 1 criminogenic needs (cluster) are *not* targeted
- Gender specific needs (cluster) are *not* targeted
- Family problem solving skills are targeted*, but the offenders' relationship with their children* or the offenders' parenting skills are *not* targeted
- When the offender is reinforced, she is told why she is being reinforced
- Avoids family activities as the mechanism for involving the families in treatment

RESEARCH QUESTION 6: What program characteristics are important for **neither** adult **male nor female** offenders?

--Characteristics having minimal effect on outcomes (i.e. no significant differences found) for programs serving males or females:

- A high staff to resident ratio
- Support for the program from the community with which it is housed
- Harmony between the staff and managers of the program
- Staff evaluation of hard skills such as paperwork, attendance/tardiness, dress and productivity
- Staff delivering treatment value the treatment efforts (i.e. support rehabilitation)
- The security staff value the treatment efforts by the program
- Staff turnover
- Validated responsivity tool(s) are used
- Use of mental health assessments
- The program excludes offenders with mental illness
- The number of assessments conducted
- Tier 2 criminogenic needs are targeted (i.e. substance abuse, family, education, employment, and leisure)
- Educational achievement is targeted*
- Offenders' relationship with significant others is targeted
- Low self-esteem is targeted
- Mental health is targeted
- Social or economic needs are targeted
- Childhood abuse and neglect issues targeted
- General cog-based groups are offered (e.g. thinking errors)
- Family group is offered

- Gender specific group is offered*
- Life skills group is offered
- The program mixes genders in groups
- The program varies treatment dosage by risk level
- Cases are assigned by caseload size
- The programs' use of reinforcement*
- The programs' use of punishment*
- The reinforcement to punishment ratio
- More than 40 percent of offender families are involved with the treatment program
- Aftercare is provided by an internal provider
- Relapse prevention used as part of aftercare
- Case management used as part of aftercare*

**Lack of variation on item for one or both genders*

In answering the larger research question of whether findings side more with the RNR or gender specific literature, these categorizations first suggest that that there are generally a limited number of characteristics significantly correlated with outcome for female programs. As noted earlier, this is likely a function of the small sample size for female programs. Hence, to more broadly determine characteristics that are important, regardless of gender, one should look at those important for both genders as well as those *more* important for one gender over another, as the latter categories consider both significance as well as direction of the relationship. Items falling in these categories make up approximately 43 percent of the items examined. Items important for *only* one gender over another consist of approximately 28 percent of all items. Hence, there appears to be more similarities than differences in the application of treatment strategies by gender, which supports the generalist perspective.

Nonetheless, those items deemed important for *only* one gender do show clear differences in treatment effect by gender. Some of the items found *only* important for programs serving males support the RNR literature concerning males, but not in treating female offenders. For example, *avoiding low risk program participants*, and *use of a cognitive behavioral model*, key

RNR factors, were deemed *only* important for programs serving males:. Yet, other factors deemed important for *only* males stands in contrast to the gender responsive literature: *drug/alcohol issues targeted, trauma/PTSD targeted*, and *family treatment interventions provided*. Likewise, of the characteristics deemed *more* or *only* important for females, items such as *use of a criminal attitude assessment,* and ***not*** *targeting relationship with children* or *parenting skills* again stands in contrast with the gender specific approach, while ***not*** *targeting Tier 1 criminogenic needs* clearly supports a gender responsive approach and opposes the RNR model.

Also interesting is that the largest category concerns characteristics important for *neither* males nor females. Note that this does include items that may have approached, but failed to reach significance is predicting recidivism. Hence, the label of "unimportant" should be read with caution. When examining what fell into this category, several key items deemed important by the gender specific literature were not predictive of outcome (e.g., gender specific needs, low self-esteem, mental health, social/economic needs and childhood abuse and neglect targeted, use of case management in aftercare, and gender specific group offered). Another key consideration, likely to be argued by the gender-specific camp is that the outcome measure of decreased incarceration rate may not be capturing program characteristics that are important for reasons other than recidivism reduction.

Some of the items in this section that are deemed important by the RNR literature were also found to have no relationship with outcome (e.g., general cog-based groups offered, variation of treatment dosage by risk, use of reinforcement and punishment, and use of relapse prevention during aftercare). With these items, the lack of correlation may be a function of *how* these services are delivered. For example, a program might deliver a thinking errors group that

partially follows a cognitive model, but misses a cognitive restructuring component which teaches offenders how to think differently rather than just how their thinking is distorted. Hence, the details in the delivery of the item may be impeding a positive impact on outcome. This section also found that *use of validated responsivity tools* lacked a positive relationship with outcome. Like with the gender responsive items, use of responsivity tools may help with other program goals such as retention/successful completion rate, as opposed to decreasing likelihood of recidivism.

Scales

Scales for each of the four program areas (Leadership/Design, Staff, Assessment and Treatment) were developed, as well as an overall scale that combined the four domains. For programs serving males, in both the full and successful completer samples each of the scales were significantly correlated with outcome. This suggests that more items earned within a scale were related to improved effects. For programs serving females in the all participant sample, a higher score on the scale was significantly correlated with outcome for all scales except Assessment. In the successful completer sample, an improved score in the Treatment scale was associated with improved outcome while Leadership/Design, Staff, and Assessment scales were not significantly correlated to outcome. For both males and females in both the full and successful completer samples, the overall scale showed significant improvement in outcome with each higher category. Thus, the scales held up more strongly for males than females, likely because more individual items were correlated with outcome for programs serving males. Nonetheless, several of the scales developed specifically for females were related to improved program outcome, particularly in the full participant sample, and improved score on the overall scale did lead to reduced recidivism.

When examined by state, the overall scale was more predictive in Ohio than Pennsylvania. However, two of the four scales were negatively correlated with outcome in the Ohio successful completer sample, which was not true in Pennsylvania. The scales had limited predictability in Pennsylvania for programs serving females, attributable in part to the small sample size for female programs in this state.

Limitations

This study provides valuable information on gender-responsive program characteristics for both male and female offenders. Despite the importance of this type of research, several limitations to the current study should be noted. First, while this study included the examination of data on over 27,000 cases, the unit of analysis was the program, not the individual offender. Hence, the study's sample size was just 138 programs. While 138 programs is larger than many previous primary studies related to effective program characteristics, data was disaggregated by programs serving males and programs serving females. Hence, programs serving females (N=39) were particularly impacted by the problem of small sample size[64]. With a small sample size, non-significant differences may be more a function of the sample size than importance of a particular item. Also noteworthy is the variation in offender sample sizes within programs. The predicted probability of re-offense in some programs was based upon hundreds of cases, while others were based upon just a handful. This is why programs were weighted so that those with a larger sample size, which increases confidence in the findings, were given more weight.

Additional limitations related to the sample were that all cases came from just two mid-western states, which impacts generalizability. In fact, when data on the treatment scales were

[64] This is why p>.10, a less conservative estimate, was used to determine whether differences were significant.

examined by state, some scales were no longer predictive of outcome[65]. In addition to the location of the programs, facilities were also limited to serving an adult offender population. Hence, results could certainly look different for juvenile programs. Finally the data was taken only from community-based residential treatment facilities and half-way houses. Thus, for outpatient, day treatment, community supervision programs, as well as prison-based programs, results may also vary.

In addition to the small sample size and homogeneity of the group examined, there was limited variation on some of the key theoretical items. For example, in the female sample, nearly all programs indicated that the primary treatment model used in their program was a cognitive behavioral model (just 3 programs identified something other than CBT as the primary model). When items lack variation, it is difficult to compare programs with or without a particular characteristic. Hence, lack of variation may have deemed some items insignificant when a larger, less homogeneous sample may have yielded different results.

The other phenomenon with the data was that programs overwhelmingly produced a negative effect size, meaning that for most sites, matched comparison cases were less likely to be incarcerated than treatment cases. This was particularly true for the all participant sample. Hence, program variables deemed important were oftentimes not associated with a *positive* program effect, rather an *improved* program effect (albeit still negative). Reasons why most effect sizes were negative is in part tied to the outcome variable used. For example, offenders mandated to residential treatment programs have closer supervision, which can lead to a higher likelihood of incarceration (see Petersillia and Turner, 1993). Likewise, failure to meet the requirements of the treatment program (which are not imposed on the matched parolees) may

[65] This is also likely impacted by the drop in sample size for each state once divided, but particularly PA.

also lead to an increased likelihood of incarceration[66]. Nonetheless, *any incarceration* to a state facility was used as it was the only outcome variable consistent between the Ohio and Pennsylvania samples. In Ohio, incarceration for a new offense versus a parole violation could not be disaggregated. Examining only incarcerations for new crimes would have more accurately pinpointed those program characteristics associated with decreased criminal behavior. The use of any incarceration, however, does more broadly measure further involvement with the criminal justice system, be it from criminal behavior or failure to follow the conditions of community supervision. Regardless, use of an alternative outcome measure may have yielded different results.

A limitation, particularly for the Ohio sample was that the bulk of programs evaluated had been previously evaluated with the CPAI or CPC program assessment tools. Programs that have undergone multiple evaluations may have become savvy as to how to respond positively to questions related to program fidelity. While observation of treatment, interviews with participants and interviews with multiple program staff help to temper this, this may have affected how program characteristics were coded. Furthermore, program characteristics were coded by different research teams. Although these teams were trained, and steps were taken to ensure that program data was accurately and reliably coded, a certain amount of subjectivity is inevitable when conducting the process evaluations necessary in gathering the program-level data. Hence, the process of collecting the program-level data may have impacted the results.

Given that this study was designed to explore how treatment characteristics predict outcome for male versus female programs, a final limitation is that data collection on program characteristics did not begin with women in mind. Hence, program attributes were more general,

[66] This is one reason that both the full sample and successful completers were examined. The successful completer sample would omit nearly all offenders returned to prison due to failure to meet program expectations.

and while the current data allowed for exploration of many key gender specific variables (e.g. targeting mental health and past abuse, use of gender-specific groups, use of case management in aftercare), other areas could have been more thoroughly explored. Examples include more questions related to the relationship of program staff to participants or participants to other participants, frequency of family contact, services for children of participants, participant satisfaction with programming, and participant linkage with community resources. Expanding factors viewed as key by the gender specific advocates (many of which are also valued by generalists), may have broadened the study's findings.

Implications

Policy Implications

The current study represents one of the larger primary studies to explore the notion of gender responsivity by comparing data on both females and males. This method allows for examination of the impact on both sexes of constructs traditionally viewed as important for one gender over another (e.g. treating past abuse for females). Hence, including both sexes in the sample helps to determine whether a construct is in fact unique to one gender, or just presumed to be gender specific. Thus, findings from this study can have substantive impact on correctional policy and programming.

Correctional programs, particularly those serving females, have looked for direction on how to offer gender responsive services. As noted in Chapter 2, the National Institute of Corrections published a guide to help respond to this need (Bloom et al, 2003). However, in the current climate, a program's funding is increasingly dependent upon outcome. Programs must therefore be concerned about the use of evidence-based practices, specifically those tied to recidivism reduction. The NIC manuscript provided limited quantitative research on how

providing gender specific treatment can also meet a correctional program's goal of recidivism reduction. One goal of this study was to test whether core principles outlined by leaders in the gender specific/feminist field were effective at recidivism reduction, and how those same principles applied to men. Hence, findings from this study may be valuable for programs interested in being responsive to the unique needs of women (or men), but also in improving community safety by reducing offender recidivism.

Another key asset of the current study is that findings are detailed enough to provide practical guidelines for programs interested in evidence based practices. The study methods involved collecting detailed data on programs' current practices. Hence, sites can easily use the study results to design or redesign programming for adult male, female or co-ed programs.

Since many of the studies that support RNR come from meta-analyses, this study offers a somewhat unique contribution to the corrections literature as a large-scale primary study on effective program attributes. While meta-analyses provide a valuable synopsis of primary study findings, analyses are limited to the data reported by the original researchers. Data from the present study were specifically constructed to take an in-depth look at effective program attributes. Furthermore, the current study offers an important contribution to the feminist literature since this body of literature tends to be more qualitative. Hence, many of the assertions from the gender-specific literature tend to lack empirical backing. Therefore findings from this study offer a unique contribution to both fields of study within corrections.

Not only can findings from this study support programs in making decisions about the application of evidence based practices, but also those responsible for either funding or overseeing the fidelity of correctional programs. Such entities can consider this study a contribution to the body of literature on evidence-based practices and gender responsive

programming. Hence, findings from this study should be scrutinized and put within the context of the larger body of literature so that judicious decisions can be about program funding and support. This leads to the final section of this dissertation: implications related to future research.

Implications for Future Research

Although findings from this study provide useful information on attributes of effective programs, future research is needed, particularly as these attributes relate to participant gender. It will be important that future research attempt to correct some of the limitations noted in this study. Namely, findings should be replicated with programs within other jurisdictions. Likewise, expanding the sample population will be important to determine whether findings differ for youthful offenders or offenders in non-residential programming. In future studies, researchers should also consider examination of alternative recidivism measures. Incarceration is a conservative measure of recidivism; new arrest or conviction may yield different results. Researchers might also consider expanding the outcome measure(s) beyond recidivism. Alternative outcome measures might include examination of more specific criminogenic behaviors such as violence or substance abuse or measures of responsivity issues like general well-being or mental health. Perhaps some of the factors measured have limited impact on recidivism, but assist with other life areas or program goals.

This study aimed to examine how key tenants of RNR and the gender specific literature impacted program effectiveness. Certainly, as noted under the study limitations, more thorough examination of some of the gender specific variables or treatment strategies could be conducted. This might include expansion or refinement of some of the variables studied, or it may include a study of programs wholly designed to follow gender specific principles versus those not

designed to do so. Likewise, some treatment components could be examined in greater depth, such as aftercare services or group interventions.

While the bulk of findings were consistent with previous examinations of effective program characteristics, some of the findings from this study stand in contrast to that of previous studies. For example, previous studies have underscored the importance of adherence to the risk principle for both males and females (Andrews and Dowden 2006; Lovins et al. 2008). This study found that targeting higher risk offenders was important for males, but this finding did not hold for females. Some of the study's results also stand in contrast to logic. For example, this study found that less staff education and experience was correlated with improved outcome. Experience and or education, however, may be a proxy measure for staff burnout or inflexibility. Other study results explored areas not previously examined empirically, such as the effect of mixing genders in programs/groups. While mixing living areas in coed programs was problematic, mixed gender groups did not have a significant impact on outcome. Although there has been limited empirical exploration of mixed-gender groups, this finding does stand in contrast to recommendations from both the RNR and gender specific theorists. Given some of the contrasts to previous studies and current theory, results should be viewed in the context of the larger body of literature with this study's limitations in mind. Furthermore, replication will assist in determining both the validity and generalizability of the study results.

In conclusion, this study offers an empirical examination of how treatment characteristics might be varied by gender. In answering this question, both the feminist literature (a specificity perspective) and the literature on RNR (a generalist perspective) was explored. Some findings supported the RNR literature while other findings were in support of the gender specific approach. Overall, findings suggest many of the same factors positively affect outcome for

programs serving males and programs serving females, but that some key differences do exist. Future research should continue to empirically explore the question of whether treatment should vary by gender and if so, how? Meanwhile, this study further contributes to the literature on effective program characteristics, and offers one of the few comprehensive studies that will assist programs to develop empirically-driven, gender responsive services.

REFERENCES

Akers, R. (1998). *Social learning theory and social structure: A general theory of crime and deviance.* Boston: Northeastern University Press.

Andrews, D. (1999). Recidivism is predictable and can be influenced: Using risk assessments to reduce recidivism. *Forum on Correctional Research*, 12: 11-17.

Andrews, D. (2006). Enhancing adherence to risk-need-responsivity: Making quality a matter of policy. *Criminology and Public Policy,* 5: 595-602.

Andrews, D. and Bonta, J. (1995). *The Level of Service Inventory-Revised (LSI-R).* North Tonawanda, NY: Multi-Health Systems.

Andrews, D. and Bonta, J. (2006). *The Psychology of Criminal Conduct* (4th ed.). Cincinnati, OH: Anderson Publishing Co.

Andrews, D. and Bonta, J. (2010). *The Psychology of Criminal Conduct (5th ed.).* Cincinnati, OH: Anderson Publishing Co.

Andrews, D., Bonta, J. and Hoge, J. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior,* 17, 19-52.

Andrews, D. and Dowden, C. (2006). The risk principle of case classification in correctional treatment. *International Journal of Offender Therapy and Comparative Criminology,* 50: 88-100.

Andrews, D. and Dowden, C. (1999). What Works in Young Offender Treatment: A Meta-Analysis. *Forum on Correctional Research*, 11(2).

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.

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.

Bandura, A. (1973). *Aggression: A social learning analysis.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1979). Self-referent mechanisms in social learning theory. *American Psychologist, 34*, 439-441.

Barbaree, H., Seto, M., Langton, C. and Peacock, E. (2001). Evaluating the predictive accuracy of six risk assessment instruments for adult sex offenders. *Criminal Justice and Behavior,* 28(4): 490-521.

Baro, A. (1999). Effects of a cognitive restructuring program on inmate institutional behavior. *Criminal Justice and Behavior,* 26(4): 466-484.

Barnoski, R. (2004). Outcome evaluation of Washington State's research based programs for juvenile offenders. Olympia Wash: Washington State Institute for Public Policy.

Bechtel, K.A., Lowenkamp, C.T., and Latessa, E.J. (2007). Assessing the predictive validity of the youth level of service/case management inventory. *Journal of Offender Rehabilitation*, 45(2/4): 85-108.

Belknap, J. and Holsinger, K. (2006). The gendered nature of risk factors for delinquency. *Feminist Criminology,* 1(1): 48-71.

Blanchette, K. and Brown, S. (2006). *The assessment and treatment of women offenders: An integrative perspective.* West Sussex, England: John Wiley and Sons Ltd.

Bloom, B. (2000). Beyond recidivism: Perspectives on evaluation of programs for female offenders in community corrections. In M. McMahon (Ed.), *Assessment to assistance: Programs for women in community corrections* (pp. 107-138). Lanham, MD: American Correctional Association.

Bloom, B., B. Owen, and S. Covington. (2003). *Gender responsive strategies: Research, practice, and guiding principles for women offenders*. Washington DC: U.S. Department of Justice, National Institute of Corrections.

Blumstein, A. (1995). Youth violence, guns, and the illicit drug industry. *Journal of Criminal Law and Criminology***,** *Guns and Violence Symposium*, 86(1): 10.

Bonta, J. (2002). Offender risk assessment: Guidelines for selection and use. *Criminal Justice and Behavior,* 29(4): 355-379.

Bonta, J., Wallace-Capretta, S., and Rooney, J. (2000). A quasi-experimental evaluation of an intensive rehabilitation supervision program. *Criminal Justice and Behavior*, 27, 312-329.

Bonta, J., Wallace-Capretta, S., and Rooney, J. (1998). *Restorative justice: An evaluation of the restorative resolutions project.* Ottawa: Solicitor General Canada.

Braithwaite (1999). "Restorative justice: Assessing optimistic and pessimistic accounts." Pp. 1-127 in M. Tonry (ed.) *Crime and Justice: A review of Research,* vol. 25. Chicago: University of Chicago Press.

Braithwaite, J. (2002). *Restorative Justice and Responsive Regulation*. New York, Oxford University Press.

Brennan, T., Breitenbach, M., and Dieterich, W. (2010). Unraveling women's pathways to serious crime: New findings and Links to prior feminist pathways, *American Probation and Parole Association,* Spring 2010: 35-47.

Burgeon, G. and Armstrong, B. (2005). Transferring the principles of effective intervention into a "real world" prison setting. *Criminal Justice and Behavior,* 32(1): 3-25.

Cattell, R. (1973). *Personality and mood by questionnaire.* Jossey-Bass, San Francisco, CA.

Chesney-Lind, M. (2000). *Women and the criminal justice system: Gender matters*. Topics in Community Corrections. Annual issue 2000. U.S. Department of Justice, National Institute of Corrections.

Chesney-Lind, M. (1997). *The female offender: Girls, women, and crime.* Thousand Oaks, CA: Sage.

Chesney-Lind, M., and Shelden, R. G. (2004). *Girls, delinquency, and juvenile justice* (3rd ed.). Belmont, CA: Thompson Wadsworth.

Covington, S. (1999). *Helping women recover: A program for treating addiction*. San Francisco: Jossey-Bass.

Covington, S. (2003). *Beyond Trauma: A Healing Journey for Women.* Center City, MN: Hazelton Press.

Covington, S. (2004). *Voices: A Program of Self Discovery and Empowerment for Girls.* Carson City, NV: The Change Companies.

Covington, S. (2002). Helping Women Recover: Creating Gender-Responsive Treatment. In *The Handbook of Addiction Treatment for Women: Theory and Practice*, S.L.A. Straussner and Brown, (Eds.) Jossey-Bass.

Covington, S. Burke, C., Keaton, S., and Norcott, C. (2008). Evaluation of a Trauma-Informed and Gender-Responsive Intervention for Women in Drug Treatment. *Journal of Psychoactive Drugs, SARC Supplement 5, 387-398.*

Covington, S., and Bloom, B. (2007). Gender-responsive treatment and services in correctional settings. Chapter to appear in E. Leeder (Ed.), *Inside and out: Women, prison, and therapy* Binghamton, NY: Haworth Press.

Cullen, F.T. (2002). *Rehabilitation and corrections programs: Crime and public policy (2nd ed.)*. San Francisco, CA: ICS Press.

Cullen, F. and Gendreau, P. (1989). *The effectiveness of correctional rehabilitation: Reconsidering the "nothing works" debate.* In L. Goodstein and D. MacKenzie eds., American Prisons: Issues in Research and Policy. New York: Plenum.

Cullen, F. and Gendreau, P. (1999). *Assessing correctional rehabilitation: Policy, practice, and prospects.* In J. Horney ed., Criminal Justice 2000, Volume 3: Policies, Processes, and Decisions of the Criminal Justice System. Washington, DC: National Institute of Justice, U.S.

Cullen, F., Pratt, T., Micelli, S. and Moon, M. (2002). *Dangerous Liaison? Rational Choice Theory as the Basis for Correctional Intervention.* In A. R. Piquero and S. G. Tibbetts eds., Rational Choice and Criminal Behavior. New Brunswick, NJ: Transaction Publishers.

Daigle, L., Cullen, F., and Wright, J. (2007). Gender differences in the predictors of juvenile misconduct: Assessing the generality-specificity debate. *Journal of Youth Violence and Juvenile Justice*, 5, 254-286.

Daly, K. (1992). Women's pathways to felony court: Feminist theories of lawbreaking and problems of representation. *Southern California Review of Law and Women's Studies,* 2*,* 11-52.

Daly, K. (1994). *Gender, Crime and Punishment.* New Haven, CT: Yale University Press.

Davidson, W., Gottschalk, R., Gensheimer, L., and Mayer, J. (1984). *Interventions with juvenile delinquency; A meta-analysis of treatment efficacy.* Washington D.C.: National Institute of Juvenile Justice and Delinquency Prevention.

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 D. Andrews. (1999a). What works for female offenders: A meta-analytic review. *Forum on Corrections Research,* 11: 18-21.

Dowden, C. and D. Andrews. (1999b). What works in young offender treatment: A meta-analysis. *Forum on Correctional Research,* 11: 21-24.

Dowden, C. and D. Andrews. (2000). Effective correctional treatment and violent reoffending: A meta-analysis. *Canadian Journal of Criminology,* 449-467.

Dowden, C. and D. Andrews. (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.

Ellis, A., and Grieger, R. (1977). *Handbook of Rational- Emotive Therapy*. New York: Springer.

French, S. and Gendreau, P. (2006). Reducing Prison Misconducts. *Criminal Justice and Behavior,* 33(2), 185-218.

Flores, A., Lowenkamp, C., Smith, P and Latessa, E. (2006). Validating the Level of Service Inventory-Revised on a sample of federal probationers. *Federal Probation,* 70(2), 44-48.

Funk, S. (1999). Risk assessment for juveniles on probation: A focus on gender. *Criminal Justice and Behavior, 26,* 44-68.

Garrett, C. J. (1985). "Effects of Residential Treatment on Adjudicated Delinquents: A Meta-Analysis." *Journal of Research in Crime and Delinquency,* 22: 287-308.

Gehring, K., Van Voorhis, P., and Bell, V. (2010). "What Works" for Female Probationers?; An Evaluation of the Moving On Program. *Women, Girls, and Criminal Justice,* 11(1): 6-10.

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 CPAI.*

Gendreau, P., and Andrews, D. A. (2001). *Correctional Program Assessment Inventory – 2000 (CPAI-2000).* Saint John, New Brunswick, Canada: University of New Brunswick.

Gendreau, P. and Andrews, D. (1990). Tertiary prevention: What the meta-analysis of the offender treatment literature tells us about "what works." *Canadian Journal of Criminology*, 32: 173-184.

Gendreau, P., Andrews, D.A., Goggin, C., and Chanteloup, F. (1992). *The development of clinical and policy guidelines for the prediction of criminal behavior in criminal justice settings.* Unpublished manuscript, University of New Brunswick, St. John, New Brunswick.

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., Goggin, C. and Smith, P. (2002). Is the PCL-R Really the "Unparalleled" Measure of Offender Risk? A Lesson in Knowledge Cumulation. *Criminal Justice and Behavior,* 29(4), 397-426.

Gendreau, P., Goggin, C. and Smith, P. (1999). The forgotten issue of effective correctional treatment: Program implementation. *International Journal of Offender Therapy,* 43, 180-187.

Gendreau, P., Goggin, C., Cullen, F. and Andrews, D. (2000). The effects of community sanctions and incarceration on recidivism. *Forum on Corrections Research*, 12:10-13.

Gendreau, P., Little, T., and Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology,* 34, 575-607.

Gendreau, P. and Ross, R. (1979). Effective correctional treatment: Bibliography for cynics. *Crime and Delinquency,* 25: 463-489.

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

Giordano, P., Deines, J., and Cernkovich, S. (2006). In and Out of Crime: A Life Course Perspective on Girls' Delinquency. Pp. 17-40 in Karen Heimer and Candace Kruttschnitt (Eds.), *Gender and Crime: Patterns in Victimization and Offending*. New York: New York University Press.

Glaze, L. (2010). *Correctional Populations in the United States, 2009.* Bureau of Justice Statistics, Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Glaze, L. and Bonczar, T. (2007). *Probation and Parole in the United States, 2006.* Bureau of Justice Statistics, Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Glaze, L. and Bonczar, T. (2008). *Probation and Parole in the United States, 2007 Statistical Tables.* Bureau of Justice Statistics, Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Gordon, D., Arbuthnot, J., Gustafson K., and McGreen P. (1988). Home-Based Behavioral-Systems Family Therapy with Disadvantaged Juvenile Delinquents. *American Journal of Family Therapy* 16: 243-255.

Greenfeld, L. and Snell, T. (1999). *Women Offenders*. Bureau of Justice Statistics, Special Issue. Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Hannah-Moffat, K. (2004). Losing ground: Gender, responsibility and parole risk. *Social Politics.* 11(3): 363-385.

Hardyman, P. and Van Voorhis, P. (2004). *Developing Gender-Specific Classification Systems for*
Women Offenders. Washington, D.C.: National Institute of Corrections.

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.

Hanson, C. and Bussiere, M. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology,* 66(2): 348-362.

Harris, P. (1994). Client management classification and prediction of probation outcome. *Crime and Delinquency,* 40: 154-174.

Harris G., Rice M., and Quincy V. (1993). Violent recidivism of mentally disordered offenders: the development of a statistical prediction instrument. *Criminal Justice and Behavior,* 20: 315–35.

Hartney, C. (2007). The Nation's most Punitive States for Women. *National Council on Crime and Delinquency.*

Hien, D., Cohen, L., Litt, L. Miele, G., Capstick, C. (2004). Promising empirically supported treatments for women with comorbid PTSD and substance use disorders. *American Journal of Psychiatry*, 161:1426-1432.

Henggeler, S., Melton, G., Brondino, M., Scherer, D., Hanley, J. (1997). Multisystemic Therapy with violent and chronic juvenile offenders and their families: The role of treatment fidelity in successful dissemination. *Journal of Consulting and Clinical Psychology* 65: 821-833

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.

Hoffman, P. and Beck, J. (1980). Revalidating the Salient Factor Score - A research note. *Journal of Criminal Justice,* 8 (3): 185-188

Holsinger, A., Lowenkamp, C.T., and Latessa, E. (2006). Exploring the validity of the Level of Service Inventory-Revised with Native American offenders. *Journal of Criminal Justice*, 34(3), 331 - 337.

Holsinger, A., Lowenkamp, C.T., and Latessa, E. (2006a). Predicting institutional misconduct using the Youth Level of Service/Case Management Inventory. *American Journal of Criminal Justice*, 30(2), 267, 284.

Holtfretter, K. and Cupp, R. (2007). Gender and risk assessment: The empirical status of the LSI-R for women offenders. *Journal of Contemporary Criminal Justice,* 23: 363-382.

Holtfreter, K., Reisig, M., *and* Morash, M. *(*2004*).* Poverty, state capital, and recidivism among women offenders*. Criminology and Public Policy,* 3*:*185-208.

Hoge, J. and Andrews, D. (1996). *Assessing the youthful offender: Issues and techniques.* New York: Plenum Press.

Hubbard, D.J. and Matthews, B. (2007). Reconciling the differences between the "gender-responsive" and the "what works" literatures to improve services for girls. *Crime and Delinquency,* 54: 225.

Hubbard, D.J. and Pratt, T. (2002). A meta-analysis of the predictors of delinquency among girls. *Journal of Offender Rehabilitation*, 34(3):1-13.

Izzo, R. and Ross, R. (1990). "Meta-Analysis of Rehabilitation Programs for Juvenile Delinquents: A Brief Report". *Criminal Justice and Behavior,* 17: 134-142.

Irvine A. and Roa J. (2010). Title II Evaluation: Gender Specific Programs in Sonoma and Santa Cruz Counties 2007-2010. Santa Cruz, CA.: Ceres Policy Research.

Jesness, C. (1988). "Jesness Inventory Classification System." *Criminal Justice and Behavior*, 15(1):78-91.

Jesness, C. F. and R. F. Wedge (1983). *Classifying Offenders: The Jesness Inventory Classification System*. Sacramento: California Youth Authority.

Kohlberg, L. (1976). Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues* (pp.31-53). New York: Holt, Rinehart and Winston.

Koons, B., Burrow, J., Morash, M., and Bynum, T. (1997). Expert and offender perceptions of program elements linked to successful outcomes for incarcerated women. *Crime and Delinquency,* 43, 512-532.

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. (1998). *Public Protection Through Offender Risk Reductions: Putting Research Into Practice.* Washington, DC: National Institute of Corrections.

Latessa, E. and Allen H. (1982). Halfway houses and parole: A national assessment. *Journal of Criminal Justice,* 10: 153-163.

Latessa, E., Cullen, F. and Gendreau, P. (2002). Beyond Correctional Quackery: Professionalism and the Possibility of Effective Treatment. *Federal Probation,* 66(2), 43-49.

Latessa, E. and Holsinger, A. (1998). The importance of evaluating correctional programming: Assessing outcome and quality. *Corrections Management Quarterly.*

Latessa, E., Lovins, L. and Smith, P. (2010). *Follow-up Evaluation of Ohio's Community Based Correctional Facility and Half-way House Programs.* Retrieved from http://www.uc.edu/ccjr/reports.html#2010.

Latessa, E., Lovins, L., Smith, P. and Makarios, M. (2010). *Follow-up Evaluation of Ohio's Community Based Correctional Facility and Half-way House Programs: Program Characteristics Supplemental Report.* Retrieved from http://www.uc.edu/ccjr/reports.html#2010.

Latessa, E. and Travis, L. (1991). Halfway house or probation: A comparison of alternative dispositions. *Journal of Crime and Justice,* 14(1): 53-76.

Latessa, E., Travis, L., and Lowenkamp, C. (2005) "Halfway houses (updated)," in E. J. Latessa and A. M. Holsinger (eds.), *Correctional Contexts, 3rd ed.* (Los Angeles: Roxbury):367-374.

Latimer, J., Dowden, C., and Muise, D. (2005). The effectiveness of restorative justice practices: A meta-analysis. *Prison Journal*, 85: 127-145.

Laws, D. and Marshall, W. (2003). A Brief History of Behavioral and Cognitive Behavioral Approaches to Sexual Offenders: Part 1. Early Developments. *Sexual Abuse: A Journal of Research and Treatment, 15(2): 75-92.*

Lipsey, M. (2009). The Primary Factors that Characterize Effective Interventions with Juvenile Offenders: A Meta-Analytic Overview. *Victims and Offenders*, 4(2): 124-147.

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. (1999). Can rehabilitative programs reduce the recidivism of juvenile offenders? An inquiry into the effectiveness of practical programs. *The Virginia Journal of Social Policy and the Law,* 6: 611-641.

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.

Lipsey, M. and Derzon, J. (1998). Predictors of violent or serious delinquency in adolescence and early adulthood: A synthesis of longitudinal research. In R. Loeber and D.P. Farrington (Eds.). *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions.* (pp. 86-105). Thousand Oaks, CA: Sage.

Lipsey, M. and Wilson, D. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist,* 48: 1181-1209.

Lipsey, M. and Wilson, D. (1998). Effective Intervention for Serious Juvenile Offenders: A Synthesis of Research. Pp. 113-336 in R. Loeber and D.P. Farrington (eds.), *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions.* Thousand Oaks, CA: Sage Publications.

Lipsey, M. and Wilson, D. (2001). *Practical Meta-analysis*. Thousand Oaks: Sage Publications.

Llewellyn, J., and Howse, R. (1998). *Restorative justice: A conceptual framework*. Ottawa: Law Commission of Canada.

Losel, F. (1995). The efficacy of correctional treatment: a review and synthesis of meta-evaluations, in: J. McGuire (ed), *What Works: Reducing Reoffending*.

Loucks, A., and Zamble, E. (1999). Predictors of recidivism in serious female offenders: Canada searches for predictors common to both men and women. *Corrections Today,* 61: 26-31.

Lovins, L.B., Lowenkamp, C., Latessa, E. and Smith, P. (2007). Application of the Risk Principle to Female Offenders. *Journal of Contemporary Criminal Justice* 23: 383-398.

Lowenkamp, C. (2004). Correctional program integrity and treatment effectiveness: A multi-site, program level analysis. *Unpublished doctoral dissertation, University of Cincinnati, Cincinnati, OH.*

Lowenkamp, C., Flores, A., Holsinger, A., Makarios, M., Latessa, E. (2010). Intensive supervision programs: Does program philosophy and the principles of effective intervention matter? *Journal of Criminal Justice,* 38: 368-375.

Lowenkamp, C., Holsinger, A., Brusman-Lovins, L., Latessa, E. (2004). Assessing the Inter-Rater Agreement of the Level of Service Inventory Revised. *Federal Probation* 68: 34-38.

Lowencamp, C., Holsinger, A., and Latessa, E. (2001). Risk/need assessment, offender classification, and the role of childhood abuse. *Criminal Justice and Behavior*, 2 (5): 543-563.

Lowenkamp, C., Holsinger, A., and Latessa, E. (2001). Risk/need assessment, offender classification, and the role of childhood abuse. *Criminal Justice and Behavior*, 28(5): 543-563.

Lowenkamp, C., 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).

Lowenkamp, C. and Latessa, E. (2002). *Evaluation of Ohio's community based correctional facilities and halfway house programs.* Technical Report. University of Cincinnati.

Lowenkamp, C. and E. Latessa (2004). Increasing the effectiveness of correctional programming through the risk principle: Identifying offenders for residential placement. *Criminology and Public Policy* 4(1): 501-528.

Lowenkamp, C., Latessa, E., and 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., Latessa, E., and Lemke, R. (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., Latessa, E. and Smith, P. (2007). "Meta-analytic review of the relationship between the LSI-R and Recidivism for female offenders." Cincinnati, Ohio: Division of Criminal Justice, University of Cincinnati. Unpublished Data.

Lowenkamp, C., Latessa, E., and Smith, P. (2006). Does correctional program quality really matter? the impact of adhering to the principles of effective intervention 2006. *Federal Probation*, 5: 201-220.

Martinson, R. (1974). "What Works?—Questions and Answers about Prison Reform". *Public Interest,* 35: 22-54.

Mayer, J. P., Gensheimer, L. K., Davidson, W. S., and Gottschalk, R. (1986). Social Learning Treatment with Juvenile Justice: A Meta-analysis of Impact in the Natural Environment. In S. Apter and A. Goldstein (Eds.). *Youth Violence: Program and Prospects.* New York: Pergamon.

MacKenzie, D. (2000). "Evidence Based Corrections: Identifying What Works." *Crime and Delinquency*, 46: 457-471

MacKenzie, D., Wilson, D., and Kider, S. (2001). Effects of correctional boot camps on offending. *The Annals for the American Academy of Political and Social Science,* 578: 126-143.

Mathews, B., Jones-Hubbard, D., and Latessa, E. (2001). Making the next step: Using evaluability assessment to improve correctional programming. *The Prison Journal,* 81: 454-472.

McConnaughy, E., Prochaska, J. and Velicer, W. (1983). Stages of change in psychotherapy: Measurement and sample profiles. *Psychotherapy: Theory, Research and Practice,* 20(3): 368-375.

Megargee, E., Carbonnell, M., Bohn, M. and Sliger, G. (2001). *Classifying criminal offenders with the MMPI-2: The Megaree System.* Minneaplois, MN: University of Minnesota Press.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach*. Plenum Press: NY.

Morash, M. (2009). Establishment of a juvenile police record: The influence of individual and peer group characteristics. *Criminology,* 22(1): 97-111.

Najavits, L.M. (2002). Seeking Safety: A Treatment Manual for PTSD and substance abuse. New York, NY: Guilford.

Najavits L. (2002) Seeking Safety: *A new psychotherapy for posttraumatic stress disorder and substance use disorder.* In: Trauma and Substance Abuse: Causes, Consequences, and Treatment of Comorbid Disorders (P. Ouimette and P. Brown, Eds.). Washington, DC: American Psychological Association Press.

Najavits, L. M., Weiss, R. D., Shaw, S. R., and Muenz, L. R. (1998). Seeking safety": Outcome of a new cognitive-behavioral psychotherapy for women with posttraumatic stress disorder and substance dependence. *Journal of Traumatic Stress,* 11, 437-456.

Nesovic, A. (2004). *Psychometric Evaluation of the Correctional Program Assessment Inventory (CPAI).* Unpublished doctoral dissertation, Carleton University, Ottawa, Ontario.

Patternoster, R. (1987). The deterrence research at the outset of the twenty-first century. Pp. 1-42 in M. Tonry (ed.) *Crime and Justice: A Review of Research,* Volume 23. Chicago: University of Chicago Press.

Palmer, T. (1975). "Martinson Revisited". *Journal of Research in Crime and Delinquency,* 12: 133-152.

Palmer, T. (1995). Programmatic and non-programmatic aspects of successful intervention: New directions for research. *Crime and Delinquency*, 41, 100-131.

Pavlov, I. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex.* G. Anrep (trans.). London, UK: Lawrence and Wishart.

Pearson, F. S., and Lipton, D. S. (1999). A meta-analytic review of the effectiveness of corrections-based treatments for drug abuse. *The Prison Journal*, 79, 384−410.

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.

Petersilia, J. (1998). A decade of experimenting with intermediate sanctions: What have we learned? *Federal Probation,* 62(1): 3-9.

Petersilia, J. and Turner, S. (1993). Intensive probation and parole. Pp. 281-335 in M. Tonry (ed.), *Crime and Justice: A Review of Research,* Volume 17. Chicago: University of Chicago Press.

Reisig, M., K. Holtfreter and M. Morash (2006). "Assessing Recidivism Risk across Female Pathways to Crime." *Justice Quarterly* 23(3): 384-405.

Rosenthal, R. (1991). *Meta-analytic procedures for social research (rev. ed.).* Newbury Park, CA:Sage.

Ross, R. and Fabiano, E. (1985). *Time to Think: A Cognitive Model of Delinquency Prevention and Offender Rehabilitation.* Johnson City, TN: Institute of Social Science and Arts.

Salisbury, E. (2008). Gendered pathways: An empirical investigation of women offenders' unique paths to crime. *Unpublished doctoral dissertation, University of Cincinnati, Cincinnati, OH.*

Salisbury, E. and Van Voorhis, P. (2009). Gendered Pathways: A Quantitative Investigation of Women Probationers' Path to Incarceration. *Criminal Justice and Behavior* 36(6): 541-566.

Salisbury, E., Van Voorhis, P. and Spiropoulos, G. (2009). The predictive validity of a gender-responsive needs assessment: An exploratory study. *Crime and Delinquency* 55(4): 550-585.

Samenow, S. (1984). *Inside the criminal mind*. New York: Times Books.

Shields, I. and Simourd, L. (1991). Predicting predatory behavior in a population of young offenders. *Criminal Justice and Behavior,* 18: 180-194.

Simourd L. and D. Andrews (1994). Correlates of Delinquency: A Look at Gender Differences. Ottawa, ONT: Department of Psychology, Carleton University.

Skinner, B.F. (1953). *Science and Human Behavior.* New York: Macmillan.

Slavin, R. (1995). Best evidence synthesis: An intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology,* 48(1): 9-18.

Smith, M. and Glass, G. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist,* 32: 752-60.

Smith, P., Cullen, F. and Latessa, E. (2009). Can 14,737 Women be Wrong? A Meta-Analysis of the LSI-R and Recidivism for Female Offenders. *Criminology and Public Policy,* 8 (1).

Smith, P., Goggin, C., and Gendreau, P. (2002). The Effects of Prison Sentences and Intermediate Sanctions on Recidivism: General Effects and Individual Differences. *Public Works and Government Services, Canada.*

Sperber, K. (2003). Potential Applications of an existing typology to child molesting behaviors. *Unpublished doctoral dissertation, University of Cincinnati, Cincinnati, OH.*

Spiegler, M.D., and Guevremont, D.C. (2010). *Contemporary Behaviour Therapy.* (5th ed.). Pacific Grove, CA: Brooks/Cole.

Steffensmeier, D., and Allan, E. (1998). The nature of female offending: Patterns and explanation.
In R. T. Zaplan (Ed.), *Female offenders: Critical perspective and effective interventions* (pp. 5-29). Gaithersburg, MD: Aspen Publishers.

Taxman, F., (2000). *Effective Practices for Protecting Public Safety through Substance Abuse Treatment.* National Institute on Drug Abuse.

Toch, H., Adams, K., and Grant, D. (1989). Coping: Maladaptation in Prisons. New Brunswick, NJ: Transaction.

University of Cincinnati (2005). *Evidence-Based Correctional Program Checklist (CPC).*

Van Voorhis, P. (1987). Correctional effectiveness: The high cost of ignoring success. *Federal Probation,* 51(1): 56-62.

Van Voorhis, P., Braswell, M. and Lester, D. (2007). *Correctional Counseling and Rehabilitation,* 6th ed. Cincinnati, OH: Anderson Publishing Co.

Van Voorhis, P., Spruance, L., Ritchie, N., Listwan, S., Seabrook, R., and Pealer, J. (2002). *The Georgia Cognitive Skills Experiment Outcome Evaluation: Phase II.* Cincinnati, OH, University of Cincinnati (unpublished report).

Van Voorhis, P., Wright, E., Salisbury, E., and Bauman, A. (2010). Women's Risk Factors and their Contributions to Existing Risk/Needs Assessment: The Current Status of Gender Responsive Assessment. *Criminal Justice and Behavior*, 37(3): 261-288.

Van Voorhis, P. (2005). Classification of Women Offenders: Gender-Responsive Approaches to Risk/Needs Assessment. *Community Corrections Report,* 12(2): 19-20, 26-27.

Van Voorhis, P. and Brown, K. (1996). *Risk Classification in the 1990s.* Washington, DC: National Institute of Corrections.

Van Voorhis, P. and Presser. L. (2001). *Classification of Women Offenders: A National Assessment of Current Practices.* Washington, DC: U.S. Department of Justice, National Instituteof Corrections.

Visher, C. (1987). Incapacitation and crime control: Does a "lock 'em up" strategy reduce crime? *Justice Quarterly,* 4: 413-543.

Ward, T. and Brown, M. (2004). The good lives model and conceptual issues in offender rehabilitation. *Psychology, Crime, and Law*, *10*, 243-257.

Warren, M. Q. (1966). *Interpersonal Maturity Level Classification: Juvenile Diagnosis and Treatment of Low, Middle and High Maturity delinquents*. California Youth Authority.

Welsh, W. N. (2006). The need for a comprehensive approach to program planning, development and evaluation. *Criminology and Public Policy,* 5: 303-615.

West, H. (2010). *Prisoners at yearend 2009, advanced counts.* Bureau of Justice Statistics, Washington, DC: U.S. Department of Justice, Office of Justice Programs.

West, H. and Sabol, W. (2008). *Prisoners in 2007.* Bureau of Justice Statistics, Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Wilson, G. (1985). Halfway house programs for offenders. In Travis, L. (ed.) *Probation, Parole and Community Corrections,* pp. 151-164. Prospect Heights, IL: Waveland.

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.

Wilson, S.J., Lipsey, M. and Soydan, H. (2003). Are mainstream programs for juvenile delinquency less effective with minority youth than majority youth? A meta-analysis of outcomes research. *Research on Social Work Practice,* 13(1): 3-26.

Worrell, J. (2001). Feminist interventions: Accountability beyond symptom reduction. *Psychology of Women Quarterly*, 25(4), 335-343.

Worell, J. and Ramer, P. (2003). *Feminist perspectives in therapy; An empowerment model for women.* Chichester, UK: John Wiley and Sons.

Yochelson, S., and Samenow, S. E. (1976). *The criminal personality: Vol l: A profile for change*. New York: Jason Aronson.

Zajac G. and Bucklen, C. (2006). Special Issue: PA DOC's Parole Violator Study (Phase 2). *Pennsylvania Department of Corrections, Research in Review,* 9(4): 1:17.

Zamble, E. (1993). Expanding the recidivism inquiry: A look at dynamic factors. *Forum on Correctional Research,* 5: 27-30.

Zlotnick, C., L.M. Najavits, D.J. Rohsenow (2003). A Cognitive-behavioral treatment for incarcerated women with substance use disorder and posttraumatic stress disorder: Findings from a pilot study. *Journal of Substance Abuse Treatment,* 25(2):99-105.

APPENDIX A

Table A1: Program Leadership/Design Characteristics—Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Females t statistic (df)	p	Males t statistic (df)	p	Females t statistic (df)	p
Program budget	-2.53 (68)	.01	-0.81 (19)	.43	-3.21 (69)	.00	0.25 (19)	.80
Years of operation	0.31 (97)	.76	2.84 (34)	.01	0.60 (96)	.55	1.67 (34)	.10
Staff to resident ratio	0.77 (78)	.45	0.27 (28)	.79	0.13 (79)	.90	0.18 (27)	.86
Support rating from community-at-large	1.79 (96)	.08	-0.56 (35)	.58	1.63 (95)	.11	0.16 (34)	.87
PD delivers clinical training	-0.69 (97)	.49	-2.24 (35)	.03	-0.59 (96)	.56	-0.49 (34)	.63
PD delivers services	-1.60 (97)	.11	-0.20 (35)	.84	-2.38 (96)	.02	-0.09 (34)	.93
Harmony b/t staff and management	-0.70 (97)	.49	-0.97 (35)	.34	-0.55 (96)	.58	-0.16 (34)	.88
Co-ed—Shared living space	2.76 (38)	.01	1.83 (20)	.08	1.93 (40)	.06	2.25 (21)	.04
Co-ed—Shared education	1.62 (36)	.11	0.86 (20)	.40	2.21 (37)	.03	1.75 (21)	.09
Co-ed—Shared visitation	3.07 (36)	.00	1.23 (20)	.23	4.04 (37)	.00	1.87 (21)	.08

Table A2: Program Staff Characteristics—Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Females t statistic (df)	p	Males t statistic (df)	p	Females t statistic (df)	p
Selection of new Staff:								
Seeks belief that treatment works	-0.36 (97)	.72	-2.00 (35)	.05	-0.48 (96)	.63	-1.31 (34)	.20
Seeks any corrections experience	2.26 (97)	.03	1.64 (35)	.11	1.82 (96)	.07	0.37 (34)	.72
2 or more years relevant experience	4.74 (84)	.00	-1.48 (28)	.15	6.84 (90)	.00	-0.53 (28)	.60
Treatment staff higher education	-1.31 (67)	.19	1.86 (30)	.07	-0.93 (60)	.36	1.80 (29)	.08
Training on theory/philosophy of program	-1.54 (97)	.13	-2.10 (35)	.04	-0.94 (96)	.35	-0.75 (34)	.46
Clinical meeting scale	-2.24 (89)	.03	-0.36 (33)	.72	-2.53 (90)	.01	0.25 (32)	.81
Clinical supervision	-3.84 (77)	.00	-0.72 (35)	.48	-4.02 (84)	.00	0.47 (34)	.64
Staff Evaluation								
Soft skills	3.19 (94)	.00	0.88 (35)	.36	1.92 (93)	.06	0.03 (34)	.98
Hard skills	-0.64 (90)	.52	-0.21 (31)	.83	-1.31 (88)	.19	-1.09 (30)	.29
Communication/relationship skills	-0.02 (94)	.99	-0.41 (35)	.69	-2.10 (93)	.04	-1.27 (34)	.21
Treatment staff value treatment efforts	1.07 (94)	.29	-1.12 (35)	.29	0.74 (95)	.46	0.18 (34)	.86
Security staff value treatment efforts	0.65 (93)	.52	0.38 (33)	.71	1.66 (92)	.11	0.55 (32)	.59
Staff turnover	-0.16 (97)	.88	-0.74 (35)	.46	0.93 (96)	.36	0.84 (34)	.41

Table A3: Program Assessment Tools—Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Females t statistic (df)	p	Males t statistic (df)	p	Females t statistic (df)	p
Assessments Used by program:								
Risk/Need Assessment	-1.80 (97)	.08	-1.65 (35)	.11	-3.24 (96)	.00	-1.49 (34)	.15
Risk assessed using LSI-R	-1.52 (97)	.13	-0.78 (35)	.44	-2.09 (96)	.04	-0.75 (34)	.46
Validated need assessment tools	-2.07 (97)	.04	-1.12 (35)	.27	-1.14 (96)	.26	-1.83 (34)	.08
Validated substance abuse tool	-.68 (97)	.50	1.43 (35)	.16	-2.08 (96)	.04	.059 (34)	.56
Criminal attitude tool	-1.38 (97)	.17	-0.95 (35)	.35	-0.68 (96)	.50	-1.76 (34)	.09
Validated responsivity tools	-1.31 (97)	.19	-0.94 (35)	.36	-0.39 (96)	.70	0.71 (34)	.48
Biopsychosocial tool	1.92 (97)	.06	0.07 (35)	.95	2.30 (96)	.02	0.47 (34)	.64
Mental health assessment	0.75 (97)	.46	-1.16 (35)	.26	1.17 (96)	.25	-0.49 (34)	.63
Assess past abuse	2.33 (93)	.02	1.40 (34)	.17	3.22 (91)	.00	1.56 (33)	.13

Table A4: Program Assessment Characteristics—Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Females t statistic (df)	p	Males t statistic (df)	p	Females t statistic (df)	p
Who selects offenders for the program?								
Program Director	3.26 (97)	.00	0.33 (35)	.75	4.80 (96)	.00	1.18 (34)	.25
Intake Department	-3.17 (97)	.00	-0.99 (35)	.33	-3.51 (96)	.00	-1.26 (34)	.22
Admission Criteria								
Written exclusionary criteria	-1.58 (97)	.12	-1.50 (35)	.14	-2.10 (96)	.04	-1.30 (34)	.20
Excludes offender with mental illness	0.44 (96)	.66	-.026 (35)	.80	0.42 (95)	.68	0.73 (34)	.47
Proportion of population low risk	5.20 (97)	.00	-0.44 (35)	.66	6.71 (96)	.00	0.13 (34)	.90
Number of assessments conducted	-1.13 (97)	.26	-0.56 (35)	.58	-1.38 (96)	.17	-0.75 (34)	.46
Assessment training by observation only	2.02 (86)	.05	0.90 (34)	.38	1.53 (82)	.13	1.64 (33)	.11
Reassessment conducted	-2.21 (94)	.03	-1.22 (32)	.23	-0.94 (93)	.35	0.63 (32)	.53

Table A5: Program Treatment Targets—Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males		Females		Males		Females	
	t statistic (df)	p	t statistic (df)	p	t statistic (df)	p	t statistic (df)	p
Tier 1 Criminogenic Needs Targeted*	0.59 (96)	.57	2.58 (35)	.01	-0.02 (95)	.99	2.14 (34)	.04
Tier 2 Criminogenic Needs Targeted*	-0.81 (97)	.42	1.35 (35)	.19	-0.90 (96)	.37	0.56 (34)	.58
Gender Specific Needs Targeted*	-.049 (97)	.96	1.76 (35)	.09	-0.81 (96)	.42	0.64 (34)	.53
Individual Gender Responsive Needs								
Gender neutral targets:								
Drug/Alcohol	-1.31 (97)	.19	1.14 (35)	.26	-2.03 (96)	.05	0.72 (34)	.47
Vocational achievement	-2.03 (97)	.05	-0.49 (35)	.63	-2.75 (96)	.01	-1.28 (34)	.21
Educational achievement	-0.76 (97)	.45	-0.53 (35)	.60	-0.73 (96)	.47	-1.18 (34)	.25
Family affection/communication	2.93 (97)	.00	-1.27 (35)	.21	2.82 (96)	.01	-0.18 (34)	.86
Family problem solving	2.48 (97)	.02	-1.87 (35)	.07	3.66 (96)	.00	-1.29 (34)	.21
Offender relationship with sig other	0.59 (97)	.56	0.97 (35)	.34	-0.40 (96)	.69	-0.69 (34)	.50
Gender specific factors:								
Low self esteem	0.27 (97)	.79	0.23 (35)	.82	0.45 (96)	.66	0.21 (34)	.84
Mental health issues	-1.19 (97)	.24	-0.01 (35)	.99	-0.58 (96)	.56	-0.41 (34)	.68
Economic/social needs	0.23 (97)	.82	1.35 (35)	.19	-0.99 (96)	.32	-0.01 (34)	.99

Table A5 Continued: Program Treatment Targets—Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Females t statistic (df)	p	Males t statistic (df)	p	Females t statistic (df)	p
Childhood abuse/neglect	-0.63 (97)	.53	1.13 (35)	.27	-0.21 (96)	.84	0.83 (34)	.42
Trauma/PTSD	-2.51 (97)	.01	0.75 (35)	.46	-1.87 (96)	.06	1.28 (34)	.21
Offender relationship with children	0.25 (97)	.80	2.49 (35)	.02	-0.98 (96)	.33	1.12 (34)	.27
Parenting Skills	-.33 (97)	.74	1.87 (35)	.07	-0.06 (96)	.95	2.04 (34)	.05

*ANOVA used for categorical variables—F-statistic reported rather than t-statistic

Table A6: Program Group Interventions—Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Females t statistic (df)	p	Males t statistic (df)	p	Females t statistic (df)	p
Group Interventions Offered:								
Thinking for a Change	-2.49 (97)	.02	-0.17 (35)	.87	-2.84 (96)	.01	-0.36 (34)	.72
Other Cog -based group	0.25 (97)	.80	-0.25 (35)	.81	1.01 (96)	.31	-0.09 (34)	.93
Substance abuse	-1.99 (97)	.05	0.60 (35)	.55	-3.33 (96)	.00	-0.35 (34)	.73
Anger management	-3.10 (97)	.00	-0.48 (35)	.63	-2.15 (96)	.03	0.99 (34)	.33
Employment/vocational education	-1.70 (97)	.09	0.15 (35)	.88	-3.30 (96)	.00	0.08 (34)	.93
Family	-0.76 (97)	.45	0.43 (35)	.67	-1.20 (96)	.23	-0.48 (34)	.64
Gender-specific	1.27 (97)	.21	0.96 (35)	.34	-0.43 (96)	.67	-0.10 (34)	.92
Mental health	-1.34 (97)	.18	-1.97 (35)	.06	-1.69 (96)	.09	-0.99 (34)	.33
Life-skills	-0.06 (97)	.96	-0.56 (35)	.52	-0.37 (96)	.72	-1.16 (34)	.25
Eclectic	2.28 (97)	.03	-0.21 (35)	.84	1.48 (96)	.14	0.36 (34)	.72
Groups assigned based on need	-2.43 (93)	.02	-0.32 (35)	.75	-3.64 (92)	.00	-0.43 (34)	.67
Program mixes genders in groups	0.05 (86)	.96	0.89 (33)	.38	-1.39 (83)	.17	1.15 (32)	.26
Program uses structured curricula	-3.43 (90)	.00	0.42 (34)	.68	-3.80 (89)	.00	-0.61 (33)	.55

Table A6 Continued: Program Group Interventions—Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Females t statistic (df)	p	Males t statistic (df)	p	Females t statistic (df)	p
Number of groups provided	-2.08 (97)	.04	0.06 (35)	.96	-2.48 (96)	.02	-0.48 (34)	.63
Use of role play	-1.99 (96)	.05	-0.39 (35)	.70	-1.77 (96)	.08	-0.95 (34)	.35
Use of graduated rehearsal	-3.25 (96)	.00	-1.39 (35)	.17	-2.50 (96)	.01	-0.70 (34)	.49

Table A7: General Program Treatment Characteristics –Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Females t statistic (df)	p	Males t statistic (df)	p	Females t statistic (df)	p
Use of a CBT model	-2.25 (75)	.03	0.25 (22)	.80	-2.85 (75)	.01	0.27 (22)	.79
Average length of stay	-2.82 (97)	.01	0.88 (35)	.39	-3.74 (96)	.00	0.01 (34)	.99
Variation of treatment by risk level	-1.31 (96)	.19	-1.37 (35)	.18	-0.46 (96)	.65	-0.33 (34)	.75
Offenders separated by risk level	-1.59 (97)	.12	0.04 (35)	.97	-2.83 (96)	.01	-1.02 (34)	.31
Cases assigned by caseload size	-0.66 (96)	.51	1.31 (35)	.20	-0.70 (96)	.49	0.49 (34)	.62
Community monitoring while on pass	2.01 (75)	.04	-0.28 (27)	.78	0.79 (71)	.43	0.92 (27)	.37
Use of reinforcers and sanctions:								
Use of reinforcers	0.40 (97)	.69	0.26 (35)	.80	0.02 (96)	.99	0.14 (34)	.89
Use of early release as incentive	-2.29 (96)	.02	-1.18 (35)	.23	-1.18 (96)	.24	-1.30 (34)	.20
Offenders told why being reinforced	2.58 (95)	.01	-1.82 (35)	.08	2.99 (94)	.00	-1.64 (34)	.11
Use of punishers	1.37 (97)	.18	--	--	0.40 (96)	.69	--	--
Use of TC strategy as punishment	2.76 (97)	.01	0.36 (35)	.72	2.84 (96)	.01	0.91 (34)	.37
Use of isolation as a punishment	-3.30 (97)	.00	-0.49 (35)	.63	-2.58 (96)	.01	-1.08 (34)	.29

Table A7 Continued: General Program Treatment Characteristics –Detailed Bivariate Results

Variable	All Participants				Successful Completers			
	Males t statistic (df)	p	Males t statistic (df)	p	Males t statistic (df)	p	Males t statistic (df)	p
Removal of pass as a punishment	-2.10 (97)	.04	-1.27 (35)	.21	-1.84 (96)	.07	-1.92 (34)	.06
Punishers individualized	-1.33 (97)	.19	-1.73 (35)	.09	-2.20 (96)	.03	-1.49 (34)	.15
Reinforcement to punishment ratio	-1.07 (94)	.29	0.09 (35)	.93	0.27 (92)	.79	-0.06 (34)	.96
Family Interventions								
Program overview only	1.93 (97)	.06	1.48 (35)	.15	1.91 (96)	.06	2.13 (34)	.04
Family activities	0.87 (97)	.39	0.30 (35)	.77	0.80 (96)	.42	2.36 (34)	.02
Family treatment intervention	1.74 (97)	.09	0.57 (35)	.57	2.94 (96)	.00	1.44 (34)	.16
% of families involved in program	-0.57 (39)	.57	-0.92 (21)	.37	-0.16 (43)	.88	-1.31 (20)	.20
Aftercare								
External provider	1.59 (97)	.12	2.20 (35)	.03	2.83 (96)	.01	2.89 (34)	.01
Internal provider	-0.92 (97)	.36	-0.68 (35)	.50	-0.35 (96)	.73	0.20 (34)	.85
Self-help/AA only	1.85 (93)	.07	-1.57 (35)	.13	2.72 (93)	.01	-0.79 (34)	.43
Relapse prevention	0.56 (93)	.58	-1.47 (35)	.15	-0.01 (93)	.99	-0.50 (34)	.62
Case management	0.74 (93)	.46	-0.17 (35)	.87	0.66 (93)	.51	-0.51 (34)	.61
Aftercare length	1.10 (33)	.28	2.31 (16)	.04	1.90 (34)	.07	2.59 (16)	.02