

Outcome Evaluation of Intensive Residential Treatment in Colorado

Maureen L. O'Keefe
Kelli J. Klebe
Evie Fisher
Kasey Roebken

September 2003



Interagency Committee on Adult
and Juvenile Correctional Treatment

ACKNOWLEDGEMENTS

This project was made possible through a grant award from the Drug Control and System Improvement Program administered by the Division of Criminal Justice, Department of Public Safety (20-DB-19-8-2). These funds originate from the Edward Byrne Memorial State and Local Law Enforcement Formula Grant Program. Through the support of the Interagency Committee on Adult and Juvenile Correctional Treatment, matching funds were obtained from the Colorado Drug Offender Surcharge Fund.

The project staff would like to thank the IRT programs for their cooperation. Their willingness to participate in such an evaluation is essential to Colorado's ability to provide quality offender treatment programs. Their willingness to allow external research is a powerful statement about their interest in continually providing better services to offenders.

Several individuals provided invaluable assistance for the duration of the project. This project could not have been completed without the support of Suzanne Pullen and her staff at the State Probation Office. Additional thanks are extended to Glenn Tapia and Diane Pasini-Hill of the Division of Criminal Justice and to Joe Stommel of the Department of Corrections.

EXECUTIVE SUMMARY

IRT programs were targeted for research because of the challenging, high risk population served and the lack of information on program services and outcomes. Offenders in Colorado who participate in IRT programs are there primarily for a drug related violation; they are often referred to this modality in lieu of prison. They might also enter the program as a condition of their sentence, either as a diversionary program for less serious offenders or as a transition from prison.

Four IRT programs were studied in the present research:

- ✚ Community Intensive Residential Treatment (CIRT) program at the San Luis Valley Community Corrections Center
- ✚ Drug and Alcohol Residential Treatment (DART) program at Williams Street Center in Denver
- ✚ Residential Treatment Center (RTC) in Greeley
- ✚ Short-Term Intensive Remedial Residential Treatment (STIRRT) program in Denver.

The literature provides insight into client and program characteristics that produce successful outcomes. Nonetheless, at a local level, very little is known about the effectiveness of IRT programs for Colorado offenders—this study seeks to diminish that gap. Because random assignment is rarely possible in offender treatment settings, a quasi-experimental design was employed which incorporates comparison groups. As a consequence of the variations among the IRTs, each program had a unique comparison group. This study incorporated official record outcome measures as well as self-report outcomes obtained through follow-up interviews.

OUTCOMES

The findings clearly indicated a positive trend across IRT programs in comparison to control groups who received less intense treatment services or none at all. The primary outcome measures included misdemeanor and felony arrests, incarceration, and wages at a 6-month follow-up. Other outcomes are discussed within the larger report.

CIRT had a positive trend across all outcome measures, but only wages were statistically significant. DART participants had better wages and lower arrest rates than controls, but the incarceration rate was nearly twice that of controls. RTC participation resulted in higher earnings and fewer felony arrests, but not misdemeanor arrests or incarceration. Only STIRRT participants achieved statistically significant outcomes, relative to a comparison group, across multiple recidivism measures. These results indicate that the program is having an overall positive effect on supervision outcomes (i.e., arrests, incarceration, and wages). Implications are discussed herein.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	I
EXECUTIVE SUMMARY.....	II
OUTCOMES.....	II
TABLE OF CONTENTS	III
INTRODUCTION.....	1
IRT PROGRAM CHARACTERISTICS.....	1
SUBSTANCE ABUSE TREATMENT RESEARCH.....	2
IRT RESEARCH	2
METHOD	4
SAMPLE	4
MATERIALS.....	5
PROCEDURE	6
RESULTS AND CONCLUSIONS	8
CIRT.....	8
DART.....	11
RTC.....	15
STIRRT.....	19
FACTORS PREDICTING SUCCESS	22
DISCUSSION	23
PROGRAM FACTORS AND RECOMMENDATIONS.....	23
CLIENT FACTORS.....	23
STUDY LIMITATIONS	24
REFERENCES.....	25
APPENDIX A	27
APPENDIX B	29
APPENDIX C	32

INTRODUCTION

Criminal activity is strongly associated with substance abuse. According to a survey of 35 sites in the U.S. in 2000, 64% of male arrestees tested positive for drug use at the time of their arrest (National Institute of Justice). Because of this association, criminal justice institutions have developed programming to target offenders with substance abuse problems. These treatment alternatives were initially implemented in the 1970's, following an influx of inmates sentenced to prison for drug related crimes (Hiller, Knight, Devereux, & Hathcoat, 1996). Drug and alcohol treatment programs have remained in place since, continuing to respond to the elevated numbers of incarcerated substance abusers.

Treatment programs generally provide programming to address substance abuse issues and criminality. The goal is to help offenders develop skills to avoid relapse and lower recidivism. In order to meet these goals and ensure positive outcomes, it is essential that the treatment needs of each offender are matched appropriately to the available treatment services. The criminal justice system in Colorado has addressed this requisite by using assessments such as the Level of Supervision Inventory (LSI; Andrews, 1982). This and other measurement tools are used to evaluate the needs level of each offender and to provide an appropriate referral. Currently, Colorado recognizes seven levels of substance abuse treatment; these include: (1) no treatment, (2) education and increased urinalysis, (3) weekly outpatient therapy, (4) intensive outpatient therapy, (5) intensive residential treatment (IRT), (6) therapeutic community (TC), and (7) no treatment – assess for psychopathy. The emphasis of the present research is to evaluate outcomes of Colorado's IRT programs for offenders.

IRT PROGRAM CHARACTERISTICS

IRT programs were targeted for research because of the challenging, high risk population served and the lack of information on program services and outcomes. Offenders in Colorado who participate in IRT programs are there primarily for a drug related violation; they are often referred to this modality in lieu of prison. They might also enter the program as a condition of their sentence, either as a diversionary program for less serious offenders or as a transition from prison. Offenders participating in these programs are at a high risk of re-offending and/or relapsing in their substance use.

Four IRT programs were studied in the present research; these include the Community Intensive Residential Treatment (CIRT) program at the San Luis Valley Community Corrections Center, the Drug and Alcohol Residential Treatment (DART) program at Williams Street Center in Denver, the Residential Treatment Center (RTC) in Greeley, and the Short-Term Intensive Remedial Residential Treatment (STIRRT) program in Denver. Other Colorado IRTs were not included because they did not exist at the inception of this study in 1999.

Due to the varying nature of IRT programs, each has different treatment durations. An offender's length of stay in the residential program could range from 14 days at STIRRT to approximately 45 days at the other three sites, with differing amounts of continuing care following the residential component. During this time, participants are housed in a correctional-type facility. Of the four programs, two treat male offenders only, while the other two programs have beds available for both males and females.

Variability can also be found between the programs in relation to their treatment perspective (O'Keefe, Doffing, & Nugent, 2001). Some Colorado IRT programs endorse a mental health model of treatment, while others use an addictions model. The use of these different models is most obvious when looking at how individual treatment needs are handled. The mental health model of treatment prefers individualized treatment sessions, while the addiction model relies more heavily on group dynamics as the primary means of treatment. This distinction translates into how groups are conducted; some follow a lecture style format, while others use a group processing format.

Staffing patterns provide further evidence of different treatment perspectives. Consistent with the mental health model, some staff members have educational degrees in disciplines related to therapy. Others have less formal training and clinical experience, but are in recovery themselves and have deep personal knowledge of substance abuse treatment. Each programs' utilization of different treatment approaches and paradigms is reflected in the individuals hired.

Overall, the IRT clients are exposed to treatment aimed at developing positive coping skills, relapse prevention education, and skills used to identify thinking errors associated with substance abuse and criminal conduct. The extent to which clients receive 12-step treatment and cognitive-behavioral therapy depends on the preferred programming of each program.

These program variations are not unique to Colorado; similar differences have been found in IRT programs nationwide. It is not at all uncommon for IRT programs to use different titles as well. For example, an IRT program might be called a short-term inpatient treatment program (Broome, Simpson, & Joe, 2002) or a community residential facility (Moos, King, Burnett, & Andrassy, 1997).

IRT programs are commonly confused with TCs in the research literature; however, the two programs are actually very distinct from one another. For example, the intensity and duration of the TC (9-12 months) exceeds that of the IRT program. The most noticeable difference between the programs is their overall philosophy and approach to treatment. TCs are structured around the social learning model. The objective is to treat the whole person, focusing on psychological, social and behavioral aspects (Nielsen & Scarpitti, 1997). The goal of treatment is to restructure attitudes and provide offenders with relapse prevention and social skills (Wexler, 1995). The staff and clients in the TC are responsible for creating the context for change, whereby the community works together as the reinforcement for positive change (DeLeon, 1994).

An IRT may best be described by defining what it is not – a TC. Although both programs share similar treatment components, IRT programs are less homogeneous in their programming and treatment approach. Rather than focusing on a thorough restructuring of the entire person, the IRT programs were developed to rapidly address addiction issues and criminality through psycho-education and therapy (O’Keefe et al., 2001).

SUBSTANCE ABUSE TREATMENT RESEARCH

Most of the research investigating residential substance abuse treatment in criminal populations focuses on TC programs rather than other types of residential programs. Research exploring the success of IRT programs with criminal populations is very rare. Researchers have, nonetheless, been able to identify a few key factors that appear to influence treatment outcomes, regardless of modality.

Length of stay has been found to be highly correlated with the success or failure of participant outcomes (Condelli & DeLeon, 1993). Time in treatment has been identified as one of the most consistent factors in determining positive treatment outcomes, which include reductions in drug use, criminal behavior and unemployment. These results relate to offenders in both the community and in prison (Condelli & Hubbard, 1994; Wexler, DeLeon, George, Kressel, & Peters, 1999).

Another factor found to be predictive of treatment outcomes is retention. In order to maximize a client’s time spent in treatment, it must be ascertained what factors would preserve clients in the program. Retention studies have examined various client and program characteristics to determine whether they have any predictive qualities. Factors that have been identified in the literature include fixed and dynamic client characteristics and program characteristics.

Fixed variables are those client characteristics which cannot be changed; these include gender, age, education, family history, and drug history. Dynamic variables are client characteristics that are fluid; they include current legal trouble, level of motivation, and psychological status. Research has found that while fixed and dynamic variables can both predict retention, the predictive power of fixed variables is more sporadic, whereas dynamic variables tend to be stronger predictors of client retention (Condelli & DeLeon, 1993; Condelli & Dunteman, 1993).

Other research has not only assessed the predictability of client characteristics on retention, but has also looked at specific program characteristics that might encourage program retention. Program characteristics relating to whether the client feels accepted into the treatment milieu, his or her relationship with the staff, the perceived clinical abilities of the staff, and the level of difficulty required to conform to the treatment regimen are all important predictors that contribute to how long the client might stay in treatment (Condelli & DeLeon 1993; Simpson, 2001).

Simpson’s model for treatment process and outcomes (2001) has identified and integrated many of these key client and program characteristics that are necessary to sustain treatment retention and improve outcomes. This model suggests that client and program variables work together throughout the course of treatment to improve retention and outcomes. The model proposes that program participation must first be promoted by combining program and staff attributes to strengthen the therapeutic relationship between counselor and client; this will in turn encourage stronger psycho-social and behavioral changes in the client that have a direct impact on retention and the success of the treatment outcomes (Simpson).

IRT RESEARCH

Compared to these general findings, studies investigating the outcomes of IRT programs have found similar results. A review of a 6-month residential program indicated that those who transferred out of the program prematurely or were expelled had significantly higher rates of recidivism 6 months following their discharge. Results also found that those who completed the primary treatment program but did not go on to aftercare had a higher level of relapse than those who were in treatment for the full 6 months and continued with aftercare counseling (Hiller et al., 1996).

Similar results were found by Broome et al. (2002) who studied the impact of length of stay on clients participating in short-term residential programs. Results from 1 year follow-up interviews indicated that participation in the program significantly decreased the participants' drug and alcohol use. Those participants who completed 5 weeks of treatment had fewer relapses in one year than those who completed only 3 weeks of treatment.

It has been found that residential treatment settings have more success in retaining clients than outpatient or day treatment programs (Bell, Williams, Nelson, & Spence, 1994). These results were found after clients were randomly assigned to either short-term residential or outpatient treatment programs. This research project compared residential and outpatient programming while controlling for selection bias and curriculum. Retention rates were measured by program completion. The researchers speculated that the increased intimacy established between clients and staff in the residential program might have reinforced social relationships beyond what was established in the outpatient treatment. This could be what contributed to the clients' commitment to stay in the residential program longer than the outpatient program (Bell et al.).

Positive social support is an important client characteristic contributing to successful IRT outcomes. Broome et al. (2002) found that those individuals who had a negative support system or continued unhealthy relationships following treatment were more likely to relapse than those clients who had positive peer support and no longer had relationships with deviant peers. Those who had a positive environment were more successful in their recovery and had lower odds of relapse. Positive family and peer support have also been shown to be strong factors in increasing client retention in treatment (Hiller, Knight, & Simpson, 1999; Simpson, 2001)

Fixed client variables have been studied in IRT programs to determine their role in retaining clients. Men have been found to be more likely to complete short term residential treatment than women and those with high school diplomas were more likely to complete treatment compared to those who failed to graduate high school (Bell et al., 1994). Research has also identified age and marital status as factors that can influence retention and recidivism. Older adults were more likely to complete treatment, as were those who were married (Knight & Hiller, 1997).

A study of felony probationers, sentenced to short term treatment in lieu of prison, found that dynamic client characteristics such as a history of psychiatric treatment, previous unemployment, and increased levels of depression, anxiety and hostility were all predictive of premature discharges from treatment (Hiller et al., 1999).

Program characteristics have been found to be as important in retaining clients in short term programming as in long term programming. Moos et al. (1997) studied clients in a short term residential treatment program to see how they responded to different characteristics within the program. Information was collected about the client's perceptions of the program size, staffing, policies and services and how these correlated with the retention rates of the clients. Results indicated that there was a strong correlation between program completion and residential stability when there was a high percentage of the staff who were in recovery, when the program was highly structured, and when the residents were involved in some of the decision making. The research also found that the retention rate increased when clients were clear on the policies of the program and were given high expectations to engage in the recommended programming (Moos et al.).

The literature provides insight into client and program characteristics that produce successful outcomes. Nonetheless, at a local level, very little is known about the effectiveness of IRT programs for Colorado offenders—this study seeks to diminish that gap. Because random assignment is rarely possible in offender treatment settings, a quasi-experimental design was employed which incorporates comparison groups. As a consequence of the variations among the IRTs, each program had a unique comparison group. This study incorporated official record outcome measures (e.g., arrest, incarceration, wages) as well as self-report outcomes obtained through follow-up interviews.

METHOD

SAMPLE

Treatment participants at each of the four IRT programs were selected for this study. All participants admitted to CIRT, RTC, and DART within a 7-8 month timeframe, starting October 2000, were included in the sample. Data collection at STIRRT was extended to a whole year, also starting October 2000, due to delays with study approval to consent for follow-up interviews at that particular site. Because of the longer timeframe and the greater number of clients served by STIRRT, random selection procedures were engaged to limit the group size. The sample included 150 participants at CIRT, 172 at DART, 226 at RTC, and 278 at STIRRT.

Control participants were obtained after all treatment participants were selected, so as to identify individuals who were similar to the experimental group. Control participants were selected to represent the overall composition of the treatment sample based on two criteria. The first criteria categorized participants as being within Judicial or Department of Corrections (DOC) jurisdictions. Judicial clientele included probationers, drug court, and diversion offenders. DOC participants included parolees and transition offenders. The second matching dimension classified individuals as technical violators or condition of placement. Technical violators included offenders who needed IRT placement due to relapse while under supervision, as evidenced by a positive urinalysis or other indicator. Condition participants included substance abusers who were determined by a criminal justice authority as needing IRT placement as a diversion or transition from prison. Samples were downloaded by researchers at the State Probation Office and DOC.

Judicial controls included probationers who were issued a revocation for a drug-related violation with a reinstatement to Probation. Drug-related violations included positive urinalysis or an arrest for possession. Other Judicial controls, in need of IRT placement due to their history of addiction rather than a technical violation, included those assessed as level 5 upon intake into Probation.

DOC controls included inmates transitioning to the community who had a substance abuse problem, but went to a community corrections center without an IRT component. DOC controls also included parolees charged for a substance-related technical violation. All parolees included in the study were found guilty, and only those who were not returned to prison were eligible for the study. Many of the parole technical violators served 30-45 days in jail as opposed to residential treatment.

Criminal justice files on all control participants were hand-screened to determine their eligibility for IRT placements. Sex offenders and mentally ill offenders were excluded, to model the treatment participants. Additionally, anyone who attended residential treatment (level 5 or 6) in the 6 months prior to the study, or during the follow-up period, was excluded from the control group. These procedures yielded a control group of 215.

The target sample size for follow-up interviews was 50 per treatment group. In order to distribute the follow-up interviews over a longer period of time, participants were randomly selected from each treatment group at a rate established to accommodate the interviewers. Participants were over-sampled in anticipation of subject attrition.

Treatment participants selected for the follow-up interviews were approached within the week prior to their IRT discharge. Participants were informed they would receive \$30 per interview plus a \$10 bonus if the interviewee was able to meet within 10 days of their interview due date. Those who agreed to participate signed a consent form and provided contact information. Of the 246 randomly selected for follow-ups, 22 did not wish to participate or later withdrew their consent. In addition, the researchers were unable to locate nine treatment participants for follow-up interviews. The 87% return rate resulted in a final interview cohort of 49 from CIRT, 54 from DART, 58 from RTC, and 54 from STIRRT.

The target sample size for the control interview cohort was 100. One hundred thirteen were selected for interviews; 3 refused to participate and 12 were never located. It should be noted that those who were not found had never been consented for the study and were unaware of the study, unlike the treatment group. A control group was selected for each treatment group using the following as matching criteria: gender, jurisdiction (Judicial, DOC), and referral status (technical violators, condition of placement). It should be noted that the control groups were not discrete from each other; participants could be in a comparison group for more than one treatment group.

Agreeable control participants signed a consent form and were paid a rate similar to the treatment group. However, because the initial contact was made at the time of the first follow-up interview and incentives could not be used for timeliness, control participants were paid \$10 for their eligibility interview and \$30 for the follow-up interview. Those who were ineligible were not able to continue in the study. There were very few ineligible control participants because of the pre-screening procedures used by the researchers. Controls were paid under the same conditions as the treatment group at the second follow-up.

This study had eight groups, with a smaller interview cohort for each group. Table 1 clarifies the samples and interview cohorts for each group. Each treatment interview cohort was compared to their respective sample across 26 measures of demographics, criminal history, substance abuse, and motivation. For example, the 49 CIRT interviewees were compared to the 150 CIRT sample. One-way chi-square and *t*-tests were conducted, using a .05 significance level. Of the 104 analyses conducted, only one comparison was significant. DART interviewees had a higher mean Adult Substance Use Survey (ASUS; Wanberg, 1997) Social score than those in the sample. The same analyses were done with the control groups. Four sets of 26 comparisons across demographic, criminal history and substance abuse characteristics were conducted. Only one comparison indicated a difference; the RTC control interview cohort had a higher ASUS Social score than the RTC control sample. It was determined that the interview cohorts closely represented the groups.

Table 1. Group Sample Sizes

	Treatment Group		Control Group	
	Entire Sample	Interview Cohort	Entire Sample	Interview Cohort
CIRT	150	49	93	40
DART	172	54	120	47
RTC	226	58	145	66
STIRRT	278	54	70	30

MATERIALS

Several assessment instruments were used to measure participants' substance use, criminal recidivism risk, and motivation at baseline. Motivation measures were only used to determine whether each IRT interview cohort was similar to their respective sample. Additionally, the research protocol for follow-up interviews is discussed below.

ASUS. The ASUS (Wanberg, 1997) is a standardized self-report inventory to screen adults who indicate a history of substance use problems. The ASUS consists of five main subscales and a global scale. These subscales are designed to measure five domains: (1) involvement in ten common drug categories, (2) degree of disruption resulting from use of drugs, (3) antisocial attitudes and behavior, (4) emotional and mood adjustment difficulties, and (5) defensiveness and resistance to self-disclosure (Wanberg). Each ASUS subscale consists of 5 to 20 items set up on either a 4 or 5-point Likert-type scale. An overall, or Global, scale is obtained by combining the scores of the involvement, disruption, social, and mood subscales. This measure is part of Colorado's standardized offender assessment battery.

LSI. The LSI (Andrews, 1982) is a semi-structured interview administered by the supervising criminal justice agency to assess criminal risk. It consists of 54 items, rated either 0 or 1, across the following domains: criminal history, accommodation, companions, alcohol/drug problems, education/employment, financial, attitude/orientation, family/marital, leisure/recreation, and emotional/personal. The resulting LSI risk score can range from 0 to 54, with higher scores indicative of greater risk. Information obtained in the interview is verified whenever possible through official offender records and other sources. This total score is used to assign the level of supervision for the offender and to determine allocation of services. When used in the Colorado criminal justice system, treatment levels are set by combining the LSI total score (supervision) with the score on the Disruption subscale of the ASUS (substance abuse).

Circumstances, Motivation, Readiness, and Suitability Scale (CMRS). The CMRS (DeLeon, Melnick, Kressel & Jainchill, 1994) inventory assesses external pressures (circumstances), intrinsic pressures (motivation), readiness, and suitability for residential TC treatment. The 52 items on the CMRS are answered on 5-point Likert-type scales ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) or 9 (*not applicable*). The four subscales are: circumstances (C; 11 items), motivation (M; 17 items), readiness (R; 8 items), and suitability (S; 16 items). Internal consistency of the M, R, and S scales is adequate, with Cronbach's alphas ranging between .70 and .86; the reliability of the C scale was lower (approximately .34). The CMRS has limited predictive validity for retention in treatment. Validity coefficients for 30-day retention ranged from .19 to .31, whereas those for 10- and 12-month retention ranged from .16 to .21 (DeLeon et al.).

University of Rhode Island Change Assessment (URICA). The URICA is a 32-item inventory designed to assess an individual's placement along a theorized continuum of behavioral change (McConaughy, Prochaska, & Velicer, 1983). Items describe how a person might think or feel when starting therapy and elicit the level of agreement with the statements. Participants answer on 5-item Likert scales that range from 1 (*strongly disagree*) to 5 (*strongly agree*). Each stage of change, precontemplation, contemplation, action, and maintenance, is measured using an 8-item subscale. For each of the four subscales, Cronbach's alpha was .69, .75, .82, and .80, respectively (DiClemente & Hughes, 1990).

Follow-up Interview. The follow-up interview was designed by the researchers to address a broad scope of factors, including quality of life, substance use and continuing treatment, and criminal activity. Appendix A provides a replication of the scoring sheet completed by the field interviewers, based on the semi-structured interview. The interview protocol has ten quality of life subscales, with two to five items per scale. Each item is rated on a 6-point Likert-type scale; subscales are calculated from the mean of all corresponding items. Interviewers were given specific criteria for rating participants on each item. Additionally, there was an Interviewer Impressions subscale to address validity concerns surrounding the interview.

Four self-report measures regarding drug use and criminal activity were used in the analyses. These were coded as frequency of drug use, number of positive urinalyses, number of criminal activities, and number of criminal arrests. Drug use was coded on a 9-point scale, ranging from 0 (*none*) to 8 (*4 or more times per day*). Criminal activities included technical violations as well as misdemeanor and felonious behaviors.

To assess the inter-rater reliability of the interviews, 22 interviews were given with two researchers (1 interviewer, 1 observer). Appendix B describes the methodology used for the inter-rater reliability study as well as more detailed statistics on the item and subscale reliabilities. Inter-rater reliability across items was high ($M = .86$), with values ranging between .16 and 1.0. Of the 44 items, 57% had an inter-rater reliability coefficient greater than .90, and 89% had coefficients greater than .80. Of greater importance is the reliability of the quality of life indicators. These subscales had higher reliability coefficients than individual items with reliabilities ranging from .70 - .99 ($M = .90$), with only one score (*Interviewer Impressions*) falling below .80. Thus, inter-rater reliability was considered to be adequate.

PROCEDURE

Baseline participant data was collected from treatment files at each of the four program sites. Identifying information, demographics, substance use and treatment history, and criminal justice supervision was included in the baseline data. Electronic systems from DOC and Judicial were used to augment missing data and verify accuracy of information. Scores from the standardized offender assessment, the battery devised to make treatment placement decisions, were the most frequently missing data elements. Because the data was not available at either the treatment program or through the criminal justice databases, it was presumed that the assessments were not done for many offenders.

IRT participants were asked to complete two motivation questionnaires as part of the baseline data. Those who agreed completed a consent form prior to responding to the questionnaires.

The baseline data for the control group was collected entirely from the criminal justice databases. Researchers were able to augment baseline data with self-report from the control interview cohort. These data primarily included dynamic variables (e.g., marital status, education level) or items not consistently tracked in databases (e.g., number of children).

Baseline criminal history information was obtained from the Colorado Crime Information Center (CCIC) and National Crime Information Center (NCIC). Arrest records were used to obtain the following: age at first adult arrest and number of arrests for technical violations (not all technical violations result in arrest), failures to appear, misdemeanors, nonviolent felonies, violent felonies, drug-related arrests, and driving under the influence (DUI). Drug-related arrests included arrests for drug sales, possession (including narcotic equipment), or distribution. This baseline criminal history data was gathered for all controls but only for the IRT interview cohorts (not the entire treatment sample). Initially, it was planned to collect CCIC/NCIC baseline data for the interview cohorts only, but it was determined to be necessary to have a more complete data set for the entire control group.

Interviews were planned for 3- and 6-month follow-ups after treatment discharge or the controls' trigger date. Various tracking procedures were engaged to locate study participants, including postcards and phone calls to the participant, family members, and parole/probation officers. Researchers only contacted family members and employers when the participant gave them permission to do so. Additionally, searches on various criminal justice databases were routinely used to locate participants who had warrants or were in jail or prison. Researchers continued to track participants even if they moved out of state or had a warrant for their arrest. Phone interviews were conducted for participants who moved out of state.

Participants who were in prison at the time of the first interview were not interviewed again. Also, those who were in jail for the first interview were only interviewed a second time if they were released between the first and second interview dates. Despite the efforts of the field researchers to locate the participants and the added incentive for timely interviews, many interviews did not occur at the established time periods. For the purposes of this study, only the interview closest to the 6-month time period was used for outcome analyses.

Official outcome data was collected for all participants, regardless of whether they were in the interview cohort or not. Rearrest data for technical violations, misdemeanors, and felonies were obtained through the CCIC/NCIC database

system. Colorado incarceration outcomes were gathered through the DOC information system. Finally, wage data were obtained from the Colorado Department of Labor and Employment database.

RESULTS AND CONCLUSIONS

CIRT

Selection of CIRT control participants yielded 93 offenders who met the matching criteria. The control group excluded female offenders and was primarily composed of DOC substance abusers transitioning from prison to the community. CIRT participants were compared to controls across demographic, criminal history, and substance abuse factors (see Tables 2, 3, and 4). CIRT participants had more Hispanics and fewer African Americans than control participants. Controls were more likely to have a GED than were CIRT participants. No other comparisons produced significant differences.

Table 2. Comparison of Controls to CIRT Demographics

	CIRT (<i>N</i> = 150)	Controls (<i>N</i> = 93)	<i>p</i>
Gender			--
Male	100%	100%	
Female	0%	0%	
Ethnicity			<.05
Caucasian	58%	61%	
African American	11%	22%	
Hispanic	29%	17%	
Other	2%	0%	
Marital			n.s.
Single	49%	51%	
Married/ Common-law	29%	32%	
Divorced/Separated/Widowed	22%	17%	
Highest education level			<.05
Less than high school	32%	25%	
GED	26%	47%	
High school diploma	31%	18%	
Some college	11%	10%	
	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>
Age	32.83 (9.12)	35.10 (9.49)	n.s.
Number of children	1.56 (1.87)	1.27 (1.48)	n.s.
Number of dependents	1.44 (1.73)	1.31 (1.69)	n.s.

Table 3. Comparison of Controls to CIRT Baseline Criminal History

	CIRT (N = 150)	Controls (N = 93)	<i>p</i>
Supervising Agency			n.s.
Judicial	24%	24%	
Dept. of Corrections	76%	76%	
Referred for Technical Violation	34%	34%	n.s.
	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>
Level of Supervision Inventory	28.16 (6.87)	29.49 (7.00)	n.s.
Age at first adult arrest ^a	23.47 (5.91)	22.85 (6.97)	n.s.
# Prior technical violations ^a	0.77 (1.28)	1.15 (1.73)	n.s.
# Prior failure to appear arrests ^a	1.84 (2.95)	2.19 (2.70)	n.s.
# Prior misdemeanor arrests ^a	11.84 (10.43)	11.44 (8.46)	n.s.
# Prior nonviolent felony arrests ^a	5.42 (6.07)	7.23 (7.57)	n.s.
# Prior violent felony arrests ^a	0.81 (1.73)	0.87 (1.48)	n.s.
# Prior drug-related arrests ^a	2.86 (3.18)	2.87 (3.47)	n.s.
# Prior DUI arrests ^a	1.81 (2.90)	1.30 (2.20)	n.s.

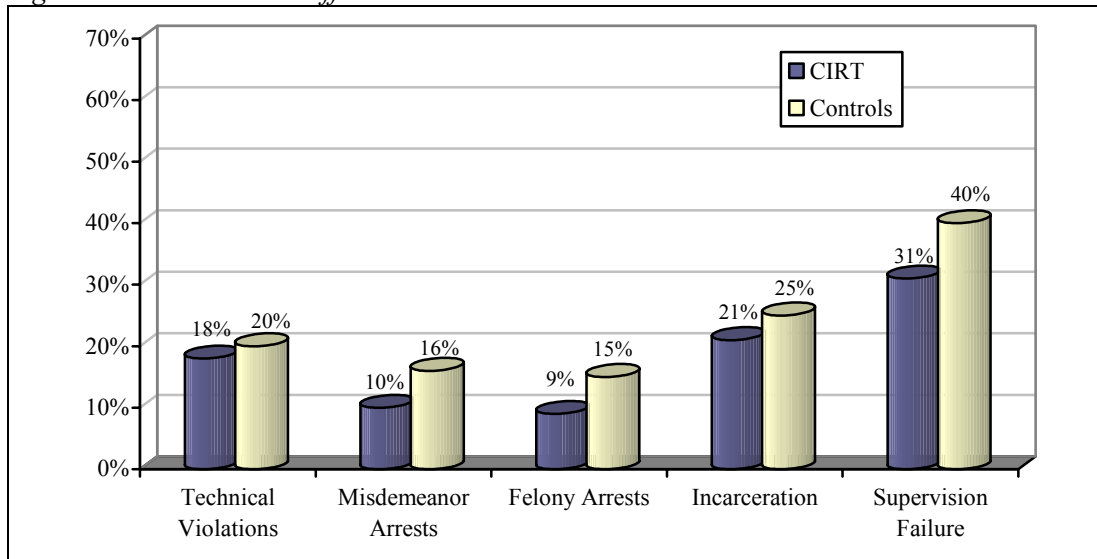
^aData collected for all controls, but only the CIRT interview cohort ($n = 49$).

Table 4. Comparison of Controls to CIRT Substance Abuse Needs

	CIRT (N = 150)	Controls (N = 93)	<i>p</i>
Previous IRT admissions	26%	20%	n.s.
Recommended treatment level			n.s.
1: No treatment	3%	4%	
2: Education	8%	5%	
3: Weekly Outpatient Therapy	20%	16%	
4: Intensive Outpatient Therapy	45%	29%	
5: Int. Residential Treatment	18%	38%	
6: Therapeutic Community	5%	8%	
7: Assess for Psychopathy	1%	0%	
	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>
Age first used drugs	13.47 (3.28)	14.45 (4.08)	n.s.
Adult Substance Use Survey (ASUS)			
Involvement	10.92 (7.45)	10.53 (7.40)	n.s.
Disruption	16.87(15.74)	21.17(17.18)	n.s.
Social	11.27 (5.03)	11.69 (5.42)	n.s.
Mood	5.43 (3.81)	6.28 (4.30)	n.s.
Defensive	8.58 (3.66)	7.86 (3.11)	n.s.

CIRT participants were compared to controls on outcomes gathered from official records. Six-month outcomes are presented in Figure 1. Ethnicity and education were not found to be significant covariates; therefore, chi-square analyses were conducted for each outcome measure, using a .05 significance level. Although there was a trend for lower recidivism rates for CIRT than controls across all outcomes, differences were not statistically significant. Twelve month outcomes did not result in significant differences between groups either (see Appendix C for a summary). However, 6-month earnings were higher for CIRT, $t(241) = 2.98$, $p < .05$. CIRT participants averaged \$5,168 compared to controls who earned only \$3,140.

Figure 1. CIRT 6-Month Official Record Outcomes



CIRT participants were compared to controls on self-report measures (see Table 5) using a series of *t*-tests. CIRT participants reported greater family satisfaction than controls. CIRT participants generally reported greater frequency of drug use and criminal behavior than controls, although only self-admitted crimes reached the significant level. No other comparisons yielded differences between groups.

The findings herein suggest a trend towards lower recidivism rates for CIRT participants, but not to a significant level. Conversely, CIRT clients tended to *report* higher rates of substance use and criminal activity. Overall, the findings indicate that CIRT participants do not have better outcomes than controls. It should be noted that the comparison group is not a “no treatment” control group. CIRT participants were compared to a large proportion of offenders who went to other community corrections centers that likely offered outpatient treatment services. Such a comparison group makes it even more challenging to demonstrate reduced recidivism rates, particularly given that outpatient services are generally longer in duration than an IRT without continuing care.

Table 5. CIRT Self-Report Outcomes

	CIRT (<i>n</i> = 49)		Controls (<i>n</i> = 40)		<i>p</i>
	<i>M</i>	(<i>SD</i>)	<i>M</i>	(<i>SD</i>)	
Quality of Life Ratings (0–5)					
Employment Stability	3.49	(1.55)	3.48	(1.16)	n.s.
Employment Meaningfulness	3.46	(1.20)	3.74	(0.82)	n.s.
Financial Stability	3.33	(1.67)	3.74	(1.40)	n.s.
Housing Stability	3.84	(1.29)	4.20	(0.78)	n.s.
Housing Satisfaction	3.79	(1.21)	3.55	(1.30)	n.s.
Partner Satisfaction	1.71	(2.08)	1.82	(2.08)	n.s.
Partner Support	3.37	(1.00)	3.54	(0.91)	n.s.
Family Satisfaction	3.82	(1.23)	3.04	(1.36)	<.01
Family Support	3.74	(1.20)	3.59	(0.91)	n.s.
Friends/Leisure	2.86	(1.05)	2.82	(0.57)	n.s.
Drug Use Frequency (0–8)	2.49	(4.94)	1.30	(2.53)	n.s.
# Positive Urinalyses	0.55	(1.14)	0.43	(0.84)	n.s.
# Self-Admitted Criminal Acts*	0.65	(1.15)	0.25	(0.49)	<.05
# Misdemeanor/Felony Arrests*	0.45	(1.26)	0.20	(0.41)	n.s.

* Includes technical violations, failures to appear, misdemeanors, and felonies.

DART

The three matching criteria yielded a control group for DART that included females, was primarily composed of Judicial clients, and approximately half were technical violators ($N = 120$). Tables 6, 7, and 8 present demographic characteristics, criminal history, and substance abuse history for DART and DART controls. Controls differed from DART participants across a number of measures. In comparison to controls, DART tended to have more African Americans, were older, had lower recommended treatment levels, and had more prior arrests for misdemeanors, violent felonies, and drug-related arrests. They also had fewer DUI arrests than controls. Taken together, it was not possible to discern that the control group was a less or more risky population than the treatment group. This supposition was further supported by the finding that none of the variables were significant covariates in the outcome analyses.

Table 6. Comparison of Controls to DART Demographics

	DART ($N = 172$)	Controls ($N = 120$)	p
Gender			n.s.
Male	74%	72%	
Female	26%	28%	
Ethnicity			<.01
Caucasian	28%	65%	
African American	43%	13%	
Hispanic	25%	20%	
Other	4%	2%	
Marital			n.s.
Single	49%	49%	
Married/ Common-law	24%	33%	
Divorced/Separated/Widowed	27%	18%	
Highest education level			n.s.
Less than high school	28%	37%	
GED	27%	28%	
High school diploma	28%	22%	
Some college	17%	13%	
	$M (SD)$	$M (SD)$	p
Age	35.41 (9.00)	32.88 (8.42)	<.05
Number of children	1.74 (1.70)	1.42 (1.40)	n.s.
Number of dependents	1.34 (1.60)	1.13 (1.21)	n.s.

Table 7. Comparison of Controls to DART Baseline Criminal History

	DART (N = 172)		Controls (N = 120)	p	
Supervising Agency				n.s.	
Judicial	68%		67%		
Dept. of Corrections	32%		33%		
Referral for Technical Violation	57%		54%	n.s.	
	<i>M (SD)</i>		<i>M (SD)</i>	<i>p</i>	
Level of Supervision Inventory	29.86	(7.14)	29.79	(6.97)	n.s.
Age at first adult arrest ^a	25.48	(7.10)	25.36	(7.94)	n.s.
# Prior technical violations ^a	1.08	(2.48)	0.76	(1.30)	n.s.
# Prior failure to appear arrests ^a	2.51	(2.45)	2.31	(2.96)	n.s.
# Prior misdemeanor arrests ^a	12.68	(11.80)	9.53	(8.21)	<.05
# Prior nonviolent felony arrests ^a	5.11	(5.63)	5.42	(5.50)	n.s.
# Prior violent felony arrests ^a	0.73	(1.30)	0.36	(0.84)	<.05
# Prior drug-related arrests ^a	5.32	(6.41)	2.23	(2.74)	<.01
# Prior DUI arrests ^a	0.24	(0.73)	1.06	(1.98)	<.01

^aData collected for all controls, but only the DART interview cohort ($n = 54$).

Table 8. Comparison of Controls to DART Substance Abuse Needs

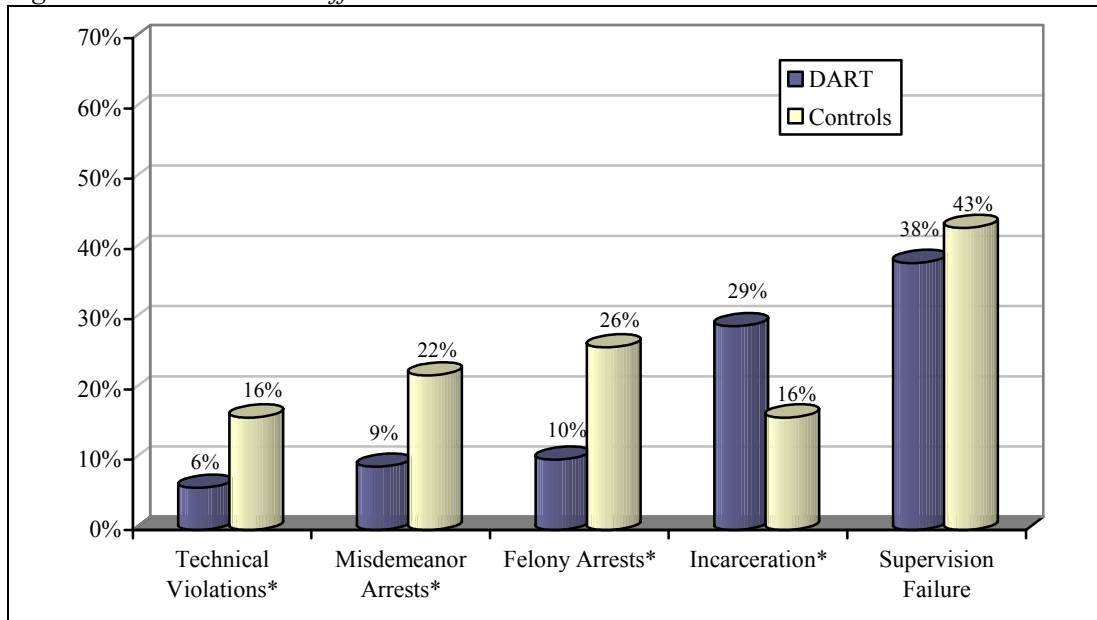
	DART (N = 172)		Controls (N = 120)	p	
Previous IRT admissions	18%		14%	n.s.	
Recommended treatment level				<.01	
1: No treatment	0%		4%		
2: Education	5%		4%		
3: Weekly Outpatient Therapy	21%		12%		
4: Intensive Outpatient Therapy	46%		26%		
5: Int. Residential Treatment	17%		47%		
6: Therapeutic Community	9%		7%		
7: Assess for Psychopathy	2%		0%		
	<i>M (SD)</i>		<i>M (SD)</i>	<i>p</i>	
Age first used drugs	13.96	(4.00)	14.48	(4.03)	n.s.
Adult Substance Use Survey (ASUS)					
Involvement	9.90	(7.00)	10.34	(6.02)	n.s.
Disruption	16.19	(16.64)	23.23	(14.94)	<.01
Social	10.06	(4.00)	10.10	(5.34)	n.s.
Mood	7.19	(5.56)	7.60	(5.35)	n.s.
Defensive	7.30	(3.26)	6.96	(3.67)	n.s.

Official record outcomes yielded interesting findings for the DART program. The pattern of results was the same at both 6- and 12- month follow-ups (see Figure 2 and Appendix C). Because none of the baseline differences in groups were significant covariates, all comparisons were conducted using chi-square analyses. DART had lower rates of arrest for technical violations, misdemeanors, and felonies. However, they had nearly double the rate of incarceration as the control group. The supervision failure variable, which takes into account misdemeanor or felony arrests as well as incarceration, revealed no differences between the groups. Indeed, further explorations revealed that the lower arrest rates were due to a greater number of DART participants who were incarcerated.

Additional exploratory analyses were conducted to understand the higher rate of incarceration for DART participants. An analysis of covariance (ANCOVA), using drug court (yes, no) and jurisdiction (Judicial, DOC) as covariates, revealed that Judicial clients in DART were at greater risk of incarceration than DOC clients. When examining DOC clients alone, DART had equivalent incarceration rates to the controls. However, the Judicial DART clients had higher rates of

recidivism than Judicial controls. This difference was not attributable to drug court cases, despite the high rate of drug-court referrals at DART.

Figure 2. DART 6-Month Official Record Outcomes



* $p < .05$

DART participants had higher reported wage earnings than did the controls, $t(289) = 2.32, p < .05$. The DART group averaged \$3,301 whereas the control group averaged \$2,344. There were virtually no differences between DART and controls on self-report outcomes (see Table 9). DART reported greater family support for a drug- and crime-free lifestyle, but no other differences were significant.

Table 9. DART Self-Report Outcomes

	DART ($n = 54$)		Controls ($n = 47$)	p
	M (SD)	M (SD)	M (SD)	
Quality of Life Ratings (0–5)				
Employment Stability	3.03 (1.43)	3.02 (1.68)		n.s.
Employment Meaningfulness	3.43 (1.09)	3.03 (1.69)		n.s.
Financial Stability	3.50 (1.42)	2.98 (1.76)		n.s.
Housing Stability	3.76 (1.18)	3.73 (1.01)		n.s.
Housing Satisfaction	3.27 (1.25)	3.56 (1.14)		n.s.
Partner Satisfaction	1.36 (1.91)	1.57 (1.94)		n.s.
Partner Support	3.34 (0.90)	3.34 (0.72)		n.s.
Family Satisfaction	3.51 (1.08)	3.23 (1.19)		n.s.
Family Support	4.00 (1.01)	3.58 (0.92)		<.05
Friends/Leisure	2.70 (0.90)	2.72 (0.62)		n.s.
Drug Use Frequency (0–8)	2.65 (5.81)	1.43 (2.95)		n.s.
# Positive Urinalyses	0.31 (0.61)	0.60 (1.35)		n.s.
# Self-Admitted Criminal Acts*	0.74 (1.67)	0.43 (0.83)		n.s.
# Misdemeanor/Felony Arrests*	0.37 (0.94)	0.38 (0.80)		n.s.

* Includes technical violations, failures to appear, misdemeanors, and felonies.

The pattern of findings for DART was different than expected. DART’s lower arrest rates were counteracted by their higher incarceration rate, resulting in an overall finding that DART did not have better outcomes than controls. In fact, the higher incarceration rate is in direct conflict with the goal of IRT programs. There was nothing within the data to indicate why DART clients, specifically Judicial offenders, would be incarcerated at a higher rate, although there is one plausible

explanation based on anecdotal information. Judicial clients sent to DART may have their sentence reconsidered following treatment completion so that they do not have to stay in community corrections. This would preclude them from returning to community corrections in the event of a later technical violation, leaving the judge few alternatives to a prison sentence.

In conclusion, findings indicate that DART does not reduce recidivism, and in fact, there is an increased risk of incarceration for program completers. Further exploration is required to determine whether these outcomes are truly related to the quality of treatment services or whether they are a result of misguided use by the criminal justice system.

RTC

The control group for RTC participants was selected to match on gender, jurisdiction, and technical violations ($N = 145$). Controls were similar to RTC on most measures (see Tables 10, 11, and 12). However, controls were more likely to be single and less likely to be divorced, separated or widowed. Controls were more likely to have a GED as compared to RTC clients who were more likely to either have a high school diploma or no certificate at all. On average, RTC clients were a year and a half younger than controls at the time of their first drug use. None of these differences were found to be related to outcomes, and therefore, they were not used as covariates in subsequent analyses.

Table 10. Comparison of Controls to RTC Demographics

	RTC ($N = 226$)	Controls ($N = 145$)	p
Gender			n.s.
Male	78%	77%	
Female	22%	23%	
Ethnicity			n.s.
Caucasian	58%	59%	
African American	11%	17%	
Hispanic	28%	22%	
Other	3%	2%	
Marital			<.05
Single	40%	46%	
Married/ Common-law	31%	37%	
Divorced/Separated/Widowed	29%	17%	
Highest education level			<.05
Less than high school	30%	23%	
GED	27%	43%	
High school diploma	29%	19%	
Some college	14%	15%	
	$M (SD)$	$M (SD)$	p
Age	33.35 (8.81)	33.49 (8.59)	n.s.
Number of children	1.64 (1.70)	1.42 (1.32)	n.s.
Number of dependents	1.45 (1.55)	1.40 (1.40)	n.s.

Table 11. Comparison of Controls to RTC Baseline Criminal History

	RTC (N = 226)	Controls (N = 145)	<i>p</i>
Supervising Agency			n.s.
Judicial	43%	43%	
Dept. of Corrections	57%	57%	
Referral for Technical Violation	66%	66%	n.s.
	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>
Level of Supervision Inventory	31.06 (7.29)	29.55 (7.54)	n.s.
Age at first adult arrest ^a	24.20 (6.74)	24.04 (7.12)	n.s.
# Prior technical violations ^a	0.86 (1.19)	1.17 (1.65)	n.s.
# Prior failure to appear arrests ^a	1.75 (2.83)	2.14 (2.70)	n.s.
# Prior misdemeanor arrests ^a	10.28 (9.73)	10.54 (8.23)	n.s.
# Prior nonviolent felony arrests ^a	4.55 (3.36)	5.76 (5.48)	n.s.
# Prior violent felony arrests ^a	0.56 (1.15)	0.52 (1.09)	n.s.
# Prior drug-related arrests ^a	2.84 (3.05)	2.60 (2.97)	n.s.
# Prior DUI arrests ^a	1.30 (2.24)	1.18 (2.13)	n.s.

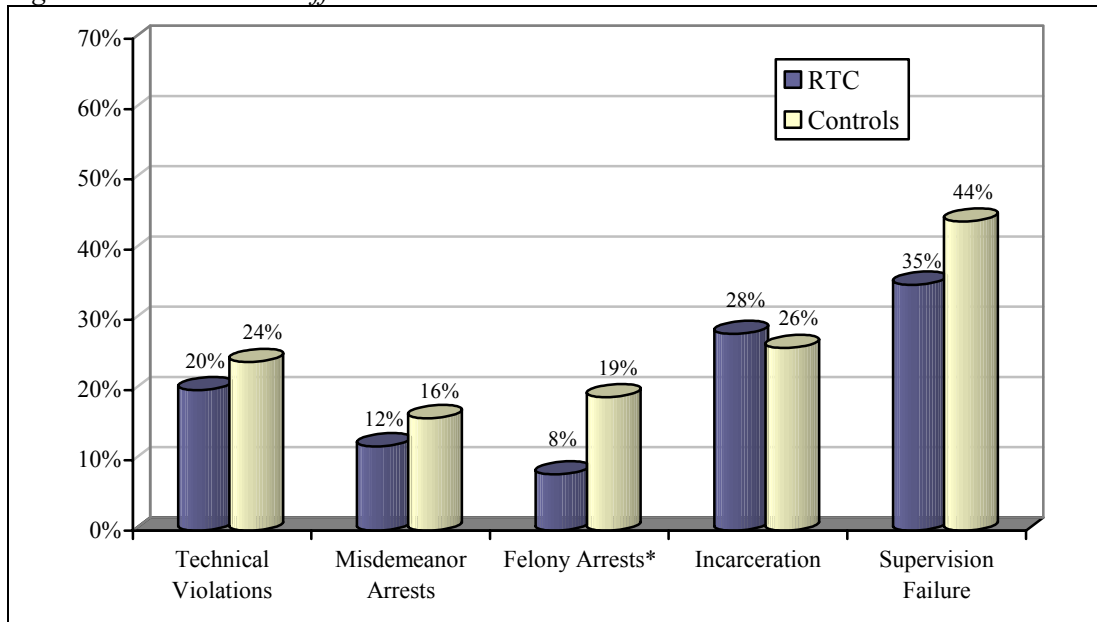
^aData collected for all controls, but only the RTC interview cohort (*n* = 58).

Table 12. Comparison of Controls to RTC Substance Abuse Needs

	RTC (N = 226)	Controls (N = 145)	<i>p</i>
Previous IRT admissions	11%	17%	n.s.
Recommended treatment level			n.s.
1: No treatment	2%	4%	
2: Education	9%	5%	
3: Weekly Outpatient Therapy	20%	20%	
4: Intensive Outpatient Therapy	41%	34%	
5: Int. Residential Treatment	20%	32%	
6: Therapeutic Community	7%	5%	
7: Assess for Psychopathy	1%	0%	
	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>
Age first used drugs	13.09 (3.59)	14.74 (3.89)	<.01
Adult Substance Use Survey (ASUS)			
Involvement	10.75 (6.50)	9.90 (6.11)	n.s.
Disruption	19.52 (16.97)	18.97 (15.23)	n.s.
Social	10.87 (4.88)	10.54 (5.52)	n.s.
Mood	6.34 (5.16)	6.52 (5.05)	n.s.
Defensive	8.33 (3.75)	7.68 (3.83)	n.s.

The RTC group generally had improved outcomes over the control group, except on the incarceration variable (see Figure 3). The only comparison found to be significantly different between groups was felony arrests, with controls nearly two and a half times more likely to have a felony arrest than RTC clients. This finding held constant for the 12-month follow-up period (see Appendix C). At 12 months, the RTC group had a 16% felony arrest rate compared to 26% for controls, $P^2(1, 371) = 5.14$, $p < .05$. RTC participants earned double the wages of controls for the 6-month follow-up period, $t(368) = 5.11$, $p < .01$. The RTC groups averaged \$4,952 while controls averaged \$2,462.

Figure 3. RTC 6-Month Official Record Outcomes



* $p < .05$.

Self-report outcomes indicated similarities across eight quality of life ratings scales and the four measures of drug and crime activities (see Table 13). However, RTC clients had better outcomes on two quality of life scales. They were found to have greater financial stability, congruent with the wage data, and greater family satisfaction than controls.

Outcomes for the RTC program were in the positive direction for several variables, most notably felony arrests at 6- and 12-month follow-ups. Reduction of felonies is one of the most significant outcomes, as it suggests an important impact on public safety. Unfortunately, the overall results lacked the strength to conclude that RTC produced greater supervision outcomes than the control group. The supervision failure variable, the most comprehensive measure of outcome, does not indicate a significant reduction in recidivism for RTC over controls. Regardless, the majority of outcomes indicate a successful, if not significant, trend.

Table 13. RTC Self-Report Outcomes

	RTC (<i>n</i> = 58)	Controls (<i>n</i> = 66)	<i>p</i>
	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)	
Quality of Life Ratings (0–5)			
Employment Stability	3.46 (1.54)	3.19 (1.50)	n.s.
Employment Meaningfulness	3.62 (1.29)	3.18 (1.43)	n.s.
Financial Stability	3.60 (1.60)	2.92 (1.81)	<.05
Housing Stability	3.99 (1.09)	3.73 (1.08)	n.s.
Housing Satisfaction	3.89 (1.04)	3.45 (1.12)	n.s.
Partner Satisfaction	1.83 (2.12)	1.81 (1.99)	n.s.
Partner Support	3.33 (0.95)	3.49 (0.93)	n.s.
Family Satisfaction	3.74 (1.04)	3.13 (1.08)	<.01
Family Support	3.83 (1.01)	3.59 (0.97)	n.s.
Friends/Leisure	2.76 (0.88)	2.73 (0.62)	n.s.
Drug Use Frequency (0–8)	2.28 (3.97)	1.35 (2.74)	n.s.
# Positive Urinalyses	0.38 (0.72)	0.44 (0.86)	n.s.
# Self-Admitted Criminal Acts*	0.34 (0.78)	0.36 (0.69)	n.s.
# Misdemeanor/Felony Arrests*	0.16 (0.45)	0.32 (0.64)	n.s.

* Includes technical violations, failures to appear, misdemeanors, and felonies.

STIRRT

The control group ($N = 70$), like the treatment group, excluded females and was composed almost entirely of technical violators. The majority were under Judicial's jurisdiction. Tables 14, 15 and 16 provide demographic, criminal history and substance abuse comparisons between groups. They differed from controls on four baseline measures; controls were more likely to be married and had lower LSI scores, ASUS Social scores, and recommended treatment level. The direction of these differences indicated that STIRRT clients were a riskier population than the controls. Indeed, LSI proved to be a significant covariate on several outcomes measures, meaning that individuals with higher scores tended to have worse outcomes. Therefore, ANCOVAs were conducted for technical violations, incarceration, and supervision failure variables, and chi-square tests were conducted for misdemeanor and felony arrests.

Table 14. Comparison of Controls to STIRRT Demographics

	STIRRT ($N = 278$)	Controls ($N = 70$)	p
Gender			--
Male	100%	100%	
Female	0%	0%	
Ethnicity			n.s.
Caucasian	56%	59%	
African American	19%	17%	
Hispanic	22%	23%	
Other	3%	1%	
Marital			<.05
Single	49%	43%	
Married/ Common-law	25%	43%	
Divorced/Separated/Widowed	26%	14%	
Highest education level			n.s.
Less than high school	28%	28%	
GED	32%	33%	
High school diploma	21%	24%	
Some college	19%	15%	
	$M (SD)$	$M (SD)$	p
Age	32.91 (9.03)	32.13 (9.19)	n.s.
Number of children	1.50 (1.77)	1.22 (1.29)	n.s.
Number of dependents	1.40 (1.67)	1.43 (1.29)	n.s.

Table 15. Comparison of Controls to STIRRT Baseline Criminal History

	STIRRT (N = 278)	Controls (N = 70)	<i>p</i>
Supervising Agency			n.s.
Judicial	57%	59%	
Dept. of Corrections	43%	41%	
Referral for Technical Violation	95%	94%	n.s.
	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>
Level of Supervision Inventory	32.48 (4.54)	27.94 (7.56)	<.01
Age at first adult arrest ^a	24.36 (8.25)	23.46 (7.16)	n.s.
# Prior technical violations ^a	0.79 (1.31)	1.24 (1.67)	n.s.
# Prior failure to appear arrests ^a	1.95 (2.40)	1.82 (2.39)	n.s.
# Prior misdemeanor arrests ^a	10.97 (8.94)	11.34 (9.13)	n.s.
# Prior nonviolent felony arrests ^a	4.74 (4.25)	6.08 (6.24)	n.s.
# Prior violent felony arrests ^a	0.83 (1.57)	0.39 (0.99)	n.s.
# Prior drug-related arrests ^a	2.45 (2.72)	2.44 (3.22)	n.s.
# Prior DUI arrests ^a	1.47 (1.98)	0.97 (2.28)	n.s.

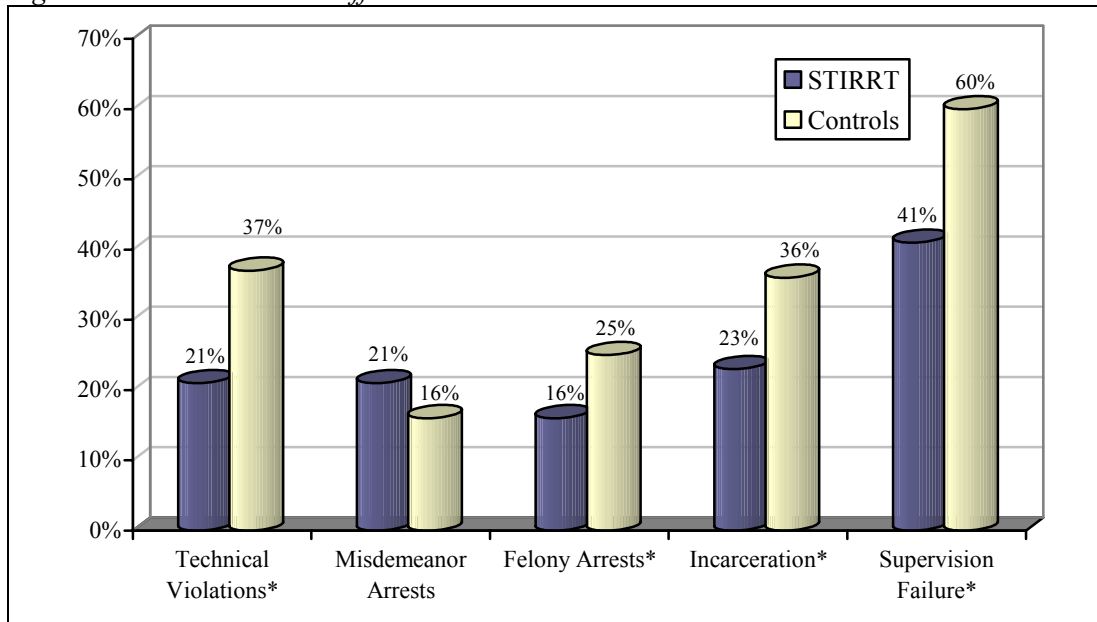
^aData collected for all controls, but only the STIRRT interview cohort (*n* = 54).

Table 16. Comparison of Controls to STIRRT Substance Abuse Needs

	STIRRT (N = 278)	Controls (N = 70)	<i>p</i>
Previous IRT admissions	14%	13%	n.s.
Recommended treatment level			<.01
1: No treatment	0%	10%	
2: Education	2%	8%	
3: Weekly Outpatient Therapy	16%	24%	
4: Intensive Outpatient Therapy	50%	37%	
5: Int. Residential Treatment	23%	18%	
6: Therapeutic Community	7%	3%	
7: Assess for Psychopathy	2%	0%	
	<i>M (SD)</i>	<i>M (SD)</i>	<i>p</i>
Age first used drugs	13.35 (3.09)	14.74 (4.95)	n.s.
Adult Substance Use Survey (ASUS)			
Involvement	10.27 (6.57)	8.52 (6.14)	n.s.
Disruption	17.55 (16.32)	14.29 (13.41)	n.s.
Social	11.38 (5.30)	9.57 (5.41)	<.05
Mood	6.30 (4.61)	5.75 (4.47)	n.s.
Defensive	8.69 (3.84)	8.08 (4.05)	n.s.

The STIRRT group had lower recidivism rates than the control group, as measured by arrests for technical violations, felony arrests, incarceration, and overall failure (see Figure 4). There was not a significant difference between groups for misdemeanor arrests. The successful trend for STIRRT participants was evidenced at 12 months (see Appendix C), but the groups were only statistically different on technical violations at that follow-up period.

Figure 4. STIRRT 6-Month Official Record Outcomes



* $p < .05$

Six-month earnings produced positive outcomes as well; STIRRT clients earned more than twice as much as the controls, $t(347) = 3.31, p < .01$. STIRRT clients averaged \$3,761 earnings for the 6-month period as compared to \$1,584 for the controls. Contrary to official record data, there were no differences between groups on any self-reported outcome measure. Table 17 summarizes self-report results.

The findings indicate that the STIRRT program successfully reduces recidivism across several outcome measures, including the comprehensive ‘supervision failure’ variable which takes into account multiple measures. STIRRT participants have fewer arrests for technical violations and felonies, are incarcerated less often, and earn higher wages than a matched comparison group. These findings do not hold constant at a 12-month follow-up period; however, these findings correspond to other research that shows a significant decline in the effectiveness of residential treatment programs for longer follow-up periods (Martin, Butzin, Saum, & Inciardi, 1999; Wexler, Melnick, Lowe, & Peters, 1999; Knight, Simpson, & Hiller, 1999).

Table 17. STIRRT Self-Report Outcomes

	STIRRT ($n = 54$)		Controls ($n = 30$)	p
	M (SD)	M (SD)	M (SD)	
Quality of Life Ratings (0–5)				
Employment Stability	3.15 (1.73)	3.33 (1.45)		n.s.
Employment Meaningfulness	3.46 (1.35)	3.34 (1.30)		n.s.
Financial Stability	2.94 (1.77)	3.10 (1.56)		n.s.
Housing Stability	3.58 (1.07)	3.88 (1.04)		n.s.
Housing Satisfaction	3.47 (1.21)	3.51 (1.22)		n.s.
Partner Satisfaction	2.32 (1.92)	1.88 (2.07)		n.s.
Partner Support	3.53 (0.85)	3.44 (0.85)		n.s.
Family Satisfaction	3.33 (0.94)	3.14 (1.33)		n.s.
Family Support	3.62 (0.97)	3.49 (0.92)		n.s.
Friends/Leisure	2.72 (0.99)	2.65 (0.63)		n.s.
Drug Use Frequency (0–8)	3.22 (4.74)	1.57 (2.94)		n.s.
# Positive Urinalyses	0.98 (2.56)	0.37 (0.93)		n.s.
# Self-Admitted Criminal Acts*	0.70 (1.64)	0.33 (0.61)		n.s.
# Misdemeanor/Felony Arrests*	0.31 (0.80)	0.23 (0.43)		n.s.

* Includes technical violations, failures to appear, misdemeanors, and felonies.

FACTORS PREDICTING SUCCESS

Correlation coefficients between predictor variables and success at 6 months and 12 months were calculated for treatment and control groups separately. Successful outcome was defined as supervision success, meaning no misdemeanor arrest, felony arrest or incarceration during the follow-up period. The interview data is reported for the first interview and significant correlations with the second interview data are noted with a superscript. Table 18 gives the correlations between the predictors and outcomes at both 6 and 12 months for the treatment and control groups.

Table 18. Predictors of Successful Outcomes

	Treatment		Control	
	6 months	12 months	6 months	12 months
Employment Stability	.21** ²	.17*	.01	-.01
Employment Meaningfulness	.22**	.22**	.00	-.09
Financial Stability	.25** ²	.27** ²	.07	-.01
Housing Stability	.17	.16	-.01	-.10
Housing Satisfaction	.19*	.22*	.00 ²	-.10 ²
Partner Satisfaction	-.03	.04	-.05	-.04 ²
Partner Support	.05	.01	.01	.05 ²
Family Satisfaction	.12	.17* ²	.33** ²	.33** ²
Family Support	.26**	.26** ²	.34** ²	.34** ²
Friends/Leisure	.30** ²	.33**	.09	-.02
Age at 1 st arrest	.13	.17*	.04	.20**
Number of Children	.03	-.07	.12	.06
LSI	.10	-.11	-.16*	-.18*
Earnings	.31**	.25**	.07	-.02
Age	.13	.20**	-.09	-.03

* $p < .05$; ** $p < .01$; ²significant using the second interview data.

Significant predictors varied between control group and treatment group participants. Employment stability and meaningfulness, financial stability and earnings were significant predictors for the treatment group but not for the control group. Higher levels of income, stability and meaningfulness were related to successful outcomes. Housing satisfaction and friends/leisure were also significant predictors of success for the treatment group with higher levels of satisfaction and positive ratings of friends/leisure corresponding to positive outcomes. Family satisfaction had a significant although small correlation with successful outcomes at 12 months, but not at 6 months for the treatment group. Family support was a statistically significant predictor of successful outcomes at both time periods for both the treatment and control groups with higher ratings of support being associated with more positive outcomes. Family satisfaction was positively correlated with successful outcomes at both time periods for the control group. Criminal risk scores had low but statistically significant negative correlations with outcomes for the control, but not the treatment, group with lower risk scores associated with successful outcomes. Age at first arrest was positively correlated to 12 month outcomes for both groups with older ages associated with successful outcomes; however, there was not a relationship of this variable with outcomes at 6 months.

Chi-square analyses were completed to assess the relationship between categorical variables with outcomes. Marital status was significantly related to outcomes for the control group, but not for the treatment group, with married individuals having poorer outcomes than single or divorced/separated/widowed individuals. Gender was related to outcomes for both the treatment and control groups, with females having a higher percentage of successful outcomes than males. Education level and ethnicity were not related to outcomes. Referral reason (i.e., technical violation, supervision condition) was not significantly related to outcomes at either time but referring agency (DOC or probation) was significantly related to outcomes at both 6 and 12 months with probation clients having more positive outcomes than DOC clients.

DISCUSSION

This study represents the first comprehensive outcome evaluation of Colorado's IRT programs for offenders. It is challenging for a treatment program to demonstrate statistically better outcomes than a comparison group, particularly when the comparison group also receives treatment. Because of the proliferation of weekly outpatient offender programs, it can be assumed that the majority of control participants received some form of non-residential treatment. Thus, in actuality, the research question addressed in this study in actuality was whether residential treatment provided an added benefit over outpatient treatment.

PROGRAM FACTORS AND RECOMMENDATIONS

The findings clearly indicate a positive trend across IRT programs, but only STIRRT participants achieved statistically significant outcomes, relative to a comparison group, across multiple measures. Even so, these findings did not hold constant over a longer follow-up period, probably due to the program's brief duration (STIRRT's continuing care component was not well developed during this study period). RTC evidenced a reduction in felony arrests at both follow-up periods, but did not impact overall supervision success. Also, DART had lower arrest rates, but those were counteracted by a higher incarceration rate.

STIRRT participants had the highest supervision failure rate of any program, but correspondingly, their target population includes the most serious, high risk offenders of all, as evidenced by their high LSI scores and the very high recidivism rate of the comparison group. This finding that STIRRT had the greatest reduction relative to their control group indicates it may be beneficial to the criminal justice system to provide intensive services to the highest risk population in lieu of prison.

Study findings lead to the question of whether the IRT model is only modestly successful or whether the programs do not have good fidelity to the model. There seems to be evidence to indicate it being a mixture of both.

The significant body of literature emphasizing the relationship between treatment duration and outcomes would suggest there is a flaw in the IRT model (Knight, Hiller, Broome, & Simpson, 2000; Condelli, 1994). The residential portion of an IRT is not nearly long enough to elicit change in this population. Colorado's criminal justice system has done very little, until recently, with its assessment and treatment system to promote the importance of continuing care for IRT completers. Without a continuing care component of 6 months or longer following residential stays, the IRT modality will have limited effectiveness. Promoting this model, as some have taken special effort to do with the STIRRT program recently, is crucial to the success of this modality.

The process evaluation of these four IRTs (O'Keefe et al., 2001) is seemingly predictive of the outcome findings, thereby suggesting issues exist regarding implementation of the model. The program that demonstrated the greatest fidelity to the model with the most intense and high quality services is the same one that had the most successful program outcomes.

Several of the conclusions drawn from the process evaluation seem to be relevant for this study as well. Criminal justice personnel consistently make inappropriate referrals and IRT programs continually enroll offenders who are not assessed as level 5. Colorado's substance abuse assessment and treatment system is predicated on the evidence that matching clients' risk and needs to treatment modality is essential. Lower risk/need clients will not benefit from the intensity afforded in an IRT, while higher risk/need clients may be better suited to a TC environment. Not even a quarter of the program enrollments, for those who were assessed, were recommended to level 5 services.

Treatment intensity is likely another essential factor impacting program outcomes. Treatment intensity was found to be minimal for three of the programs, both in terms of content and frequency of contact (O'Keefe et al., 2001). Furthermore, solid clinical training and supervision is necessary for skilled, qualified staff to deliver treatment programming. Oversight from the regulatory body is a key component to ensuring quality treatment services. It makes inherent sense that better quality services will positively impact offender outcomes.

CLIENT FACTORS

Factors that might predict successful outcomes were explored for both the treatment and control groups separately. These factors included fixed participant characteristics such as gender, criminal risk, and ethnicity as well as dynamic factors such as social support and employment and financial factors. Some variables were predictors of outcomes in both groups, but there were more differences across the groups. These differences may be indicators of the types of change that may be afforded by participating in treatment.

Social support has been shown by Broome et al. (2002) to be a predictor for treatment retention. Similarly in this study family support was a significant predictor of positive outcomes for both groups. Family satisfaction was a strong predictor for the control group but was a weaker predictor for the treatment group. Having friends who supported clean living and having positive leisure activities was also a predictor of successful outcomes, but only for the treatment group. These differences in significant predictors across the groups are noteworthy because of the lack of mean differences between the treatment and control groups. Treatment may be increasing the ability of factors to protect against failure rather than just increasing the frequency of those factors.

An interesting finding was that all four financial and employment variables were predictors of successful outcomes for the treatment group but not the control group. This may correspond to program activities to encourage and support clients in obtaining and keeping employment as a condition of treatment. Participants who are more invested in the treatment may show an increase in these specific concrete behaviors (e.g., employment) which are then impacting success in clean living (Hiller et al., 1999).

In contrast to other research (Bell et al., 1994; Knight & Heller, 1997), education level and age were not significant predictors of successful outcomes. Marital status was related to positive outcomes for the control group but in a different way than has been found in past research. Married participants had worse outcomes than single or separated/ divorced/ widowed participants, which is in direct contradiction with other research that has shown a protective factor of marriage (Knight & Hiller, 1997). Women had more positive outcomes than men. Criminal risk factors were negatively related to successful outcomes for the control group, but the effects were small in nature. The weaker and non-significant relationship for the treatment group may be an indication of the effects that treatment can have on the dynamic factors associated with risk.

STUDY LIMITATIONS

Several factors contributed to the complexity of this research design. Identifying control participants has historically challenged many researchers within the criminal justice system. However, multiple evaluation sites and multiple criminal justice jurisdictions further compounded the complexity. Even though this study was born out of interagency collaboration and had ongoing support from other state agencies, learning the intricacies of other systems was formidable.

Collecting interview data was undeniably the most difficult aspect of this study. Many researchers do not possess the clinical skills necessary to conduct interviews, and many clinicians struggle with the minutia and objectivity inherent to research. Even more so, the field researchers were frequently frustrated by their inability to locate research participants, resulting in a high turnover rate. In turn, continual training and quality assurance for the interviews was time-consuming and costly.

The researchers on this project would not endorse follow-up interviews for subsequent studies. In fact, in a separate study, the researchers discovered that offenders tend to not even report criminal activities for which they have an official record, much less undetected criminal behavior (Klebe, Fisher, & O'Keefe, 2002). Resources would probably be better invested in collecting more official data, even information that is only available in narrative form or hard files.

The breadth of data collected was substantial. However, two pieces of data seemed to be inadequate at the conclusion of the study. These were amount of time spent in a jail or residential (e.g., community corrections, other treatment) setting during the follow-up period and other treatment during the follow-up period. This study collected that data during the follow-up interviews, but the accuracy and consistency of the data remains questionable. Future research needs to explore the interaction of residential supervision and continuing care with treatment outcomes.

REFERENCES

- Andrews, D. A. (1982). *The Level of Supervision Inventory (LSI): The first follow-up*. Toronto: Ontario Ministry of Correctional Services.
- Bell, D. C., Williams, M. L., Nelson, R., & Spence, R. T. (1994). An experimental test on retention in residential and outpatient programs. *American Journal of Drug and Alcohol Abuse, 20*, 331- 341.
- Broome, K. M., Simpson, D. D., & Joe, G. W. (2002). The role of social support following short-term inpatient treatment. *The American Journal on Addiction, 11*, 57-65.
- Condelli, W. S. (1994). Domains of variables for understanding and improving retention in therapeutic communities. *The International Journal of the Addictions, 29*, 593-607.
- Condelli, W. S. & DeLeon, G. (1993). Fixed and dynamic predictors of client retention in therapeutic communities. *Journal of Substance Abuse Treatment, 10*, 11-16.
- Condelli, W.S. & Duntzman, G. H. (1993). Issues to consider when predicting retention in therapeutic communities. *Journal of Psychoactive Drugs, 25*, 239-244.
- Condelli, W. S. & Hubbard, R. L. (1994). Relationship between time present in treatment and client outcomes from therapeutic communities. *Journal of Substance Abuse Treatment, 11*, 25-33.
- DeLeon, G. (1994). The therapeutic community: Toward a general theory and model. *NIDA Research Monograph Series, 144*, 16-53.
- DeLeon, G., Melnick, B., Kressel, D., & Jainchill, N. (1994). Circumstances, motivation, readiness, and suitability (the CMRS scales): Predicting retention in therapeutic community treatment. *American Journal of Drug & Alcohol Abuse, 20*, 495-515.
- DiClemente, C. C., & Hughes, S. (1990). Stages of change profiles in outpatient alcoholism treatment. *Journal of Substance Abuse, 2*, 217-235.
- Hiller, M. L., Knight, K., Devereux, J., & Hathcoat, M. (1996). Post-treatment outcomes for substance abusing probationers mandated to residential treatment. *Journal of Psychoactive Drugs, 28*, 91-296.
- Hiller, M. L., Knight, K., & Simpson, D. D. (1999). Risk factors that predict dropout from correction-based treatment for drug abuse. *The Prison Journal, 79* (3), 411-430.
- Klebe, K. J., Fisher, E., & O'Keefe, M. (2002). *A Comparison of Outcome Data from Self-reports and Official Records*. Poster presented at the annual meeting of the American Evaluation Association, Washington, DC.
- Knight, K., Hiller, M. L., Broome, K. M., & Simpson, D. (2000). Legal pressure, treatment readiness, and engagement in long term residential programs. *Journal of Offender Rehabilitation, 31*, 101-115.
- Knight, K. & Hiller, M. L. (1997). Community-based substance abuse treatment: A 1-year outcome evaluation of the Dallas county judicial treatment center. *Federal Probation, 61*, 61-68.
- Martin, S. S., Butzin, C. A., Saum, C. A., & Inciardi, J. A. (1999). Three-year outcomes of therapeutic community treatment for drug-involved offenders in Delaware: From prison to work release to aftercare. *The Prison Journal, 79* (3), 294 - 320.
- McConaughy, E., Prochaska, J., & Velicer, (1983). Stages of change in psychotherapy: Measurement and sample profiles. *Psychotherapy: Theory, Research, and Practice, 20*, 368-375.
- Moos, R. H., King, M. J., Burnett, E. B., & Andrassy, J. M. (1997). Community residential program policies, services, and treatment orientations influence patients' participation in treatment. *Journal of Substance Abuse, 9*, 171-187.
- National Institute of Justice. (2000). ADAM: 2000 annual report on adult and juvenile arrestees. Washington DC: US Department of Justice.
- Nielsen, A. L. & Scarpitti, F. R. (1997). Changing the behavior of substance abusers: Factors influencing the effectiveness of therapeutic communities. *Journal of Drug Issues, 27*, 279-298.
- O'Keefe, M. L., Doffing, M. A., & Nugent, N. R. (2001). *Process evaluation of intensive residential treatment*. Colorado Springs, CO: Department of Corrections.
- Simpson, D. D. (2001). Modeling treatment process and outcomes. *Addiction, 96*, 201-211.
- Wanberg, K. W. (1997). *User's guide to the Adult Substance Use Survey – ASUS*. Arvada, CO: Center for Addictions Research and Evaluation.
- Wexler, H. K. (1995). The success of therapeutic communities for substance abusers in American prisons. *Journal of Psychoactive Drugs, 27*, 57-66.
- Wexler, H. K., DeLeon, G. T., Kressel, G., & Peters, D. (1999). The amity prison TC Evaluation: Re-incarceration outcomes. *Criminal Justice & Behavior, 25*, 147-167.

Wexler, H. K., Melnick, G., Lowe, L., & Peters, J. (1999). Three-year reincarceration outcomes for Amity in prison therapeutic community and aftercare in California. *The Prison Journal*, 79(3), 321-336.

APPENDIX A

Outcome Interview Score Sheet									
Subject									
#: _____		Interviewer: _____			Observer: _____				
Date: _____		Start time: _____			Finish time: _____				
Participant Supervision Status: Please list all supervisory agencies and movement dates below.									
<input type="checkbox"/> Residential supervision <input type="checkbox"/> Non-residential supervision: <input type="checkbox"/> Parole <input type="checkbox"/> Probation <input type="checkbox"/> No supervision (Off Paper)									
Employment/Education					Employment Status				
Indicate area rating: <input type="checkbox"/> Employment					<input type="checkbox"/> F/T				
<input type="checkbox"/> Education					<input type="checkbox"/> Disabled				
					<input type="checkbox"/> P/T				
					<input type="checkbox"/> Temporary				
					<input type="checkbox"/> Other (list): _____				
Stability of Employment/Education					<input type="checkbox"/> Seeking employment				
1	Type of employment/ education	0	1	2	3	4	5		
2	Length of employment/ education	0	1	2	3	4	5		
3	Continued employability/ student status	0	1	2	3	4	5		
4	Relationships with em- ployer/boss/ instructors	0	1	2	3	4	5		
5	Is seeking employment/ education	0	1	2	3	4	5		
					Meaningfulness of Employment/Education				
1	Satisfaction with job/ education as career	0	1	2	3	4	5		
2	Misses work/classes	0	1	2	3	4	5		
3	Looking for other work/ education options	0	1	2	3	4	5		
4	Relationships with co- workers/classmates	0	1	2	3	4	5		
Financial					Legal Requirements				
Stability					<input type="checkbox"/> No legal requirements				
1	Source of income/ reliance on other	0	1	2	3	4	5	<input type="checkbox"/> Legal requirements are garnished from wages first	
2	Basic requirements of living	0	1	2	3	4	5	<input type="checkbox"/> Legal requirements are not garnished but are being paid regularly (no delinquent payments)	
					<input type="checkbox"/> Legal requirements are not garnished and are not being paid regularly				
					<input type="checkbox"/> Legal requirements are not being paid				
Housing # of weeks ratings below apply to: _____									
Stability					Satisfaction				
1	Where living	0	1	2	3	4	5	N	1
2	Number of places have lived	0	1	2	3	4	5	N	2
3	Missed or late pay- ments	0	1	2	3	4	5	N	3
4	Eviction	0	1	2	3	4	5	N	4
Family/Committed Relationships					<input type="checkbox"/> Single				
Satisfaction with current partner					<input type="checkbox"/> In committed relationship _____ Number				
1	Level of support and trust	0	1	2	3	4	5	<input type="checkbox"/> Married _____ Number actively	
2	Long-term future	0	1	2	3	4	5	<input type="checkbox"/> Divorced _____ involved with	
3	Length of time	0	1	2	3	4	5	<input type="checkbox"/> Other:	

Support for Crime-free & Drug-free Lifestyle from Partner					
1	Support for crime-free lifestyle	0	1	2	3 4 5
2	Support for drug-free lifestyle	0	1	2	3 4 5
3	Crime-free lifestyle of partner	0	1	2	3 4 5
4	Substance use lifestyle of partner	0	1	2	3 4 5
Satisfaction with Immediate Family Members			Support for Crime-free & Drug-free Lifestyle from Family Members		
1	Quality/Level of support	0	1	2	3 4 5
2	Frequency	0	1	2	3 4 5
3	Percent of family have interactions with	0	1	2	3 4 5
1	Support for crime-free lifestyle	0	1	2	3 4 5
2	Support for drug-free lifestyle	0	1	2	3 4 5
3	Crime-free lifestyle of family members	0	1	2	3 4 5
4	Substance use lifestyle of family members	0	1	2	3 4 5
Friends/Leisure Activities					
1	Does prosocial activities	0	1	2	3 4 5
2	Enjoys prosocial activities	0	1	2	3 4 5
3	Activities with friends	0	1	2	3 4 5
4	Friends support for drug-free lifestyle	0	1	2	3 4 5
5	Friends support for crime-free lifestyle	0	1	2	3 4 5
Treatment (since consent or last interview)					
Check all that apply		Where:	Frequency (hrs/wk)	Duration (# wks)	
<input type="checkbox"/>	12-Step	_____	_____	_____	
<input type="checkbox"/>	SSC	_____	_____	_____	
<input type="checkbox"/>	Relapse Prevention	_____	_____	_____	
<input type="checkbox"/>	Anger Management	_____	_____	_____	
<input type="checkbox"/>	Domestic Violence	_____	_____	_____	
<input type="checkbox"/>	Cognitive Skills	_____	_____	_____	
<input type="checkbox"/>	Parenting	_____	_____	_____	
<input type="checkbox"/>	Sex Offender Tx	_____	_____	_____	
<input type="checkbox"/>	Money Management	_____	_____	_____	
<input type="checkbox"/>	Antabuse/Naltrexone:	_____	_____	_____	
<input type="checkbox"/>	Other:	_____	_____	_____	
Interviewer Impressions					
1	Openness	0	1	2	3 4 5
2	Honesty	0	1	2	3 4 5
3	Psychological Awareness	0	1	2	3 4 5

APPENDIX B

In order to assess the quality of the interview data, the inter-rater reliability of the interviews was determined. Interviews were given to 22 participants using two researchers: one was the interviewer and the other was the observer. Both researchers were in the same room during the interview. Interviews followed the same procedure as described in the method section describing the study. Following the interview, both researchers completed the *Outcome Interview Score Sheet*, without discussing the interview with each other. For individual items, inter-rater reliability was calculated with either the correlation coefficient between the interviewer and the observer ratings when the item was a quantitative variable or with Cohen's kappa if the item was categorical. For the quality of life indicator subscales, the items for the subscale were summed and the correlation between the interviewer and observer ratings was calculated. The following tables present the means for the items and subscales for the interviewer and observer along with the reliability coefficient.

Employment and Financial Variables

Variable	Interviewer Mean	Observer Mean	Inter-rater Reliability
Supervision Status	--	--	.59
Employment Status	--	--	1.00
Stability of Employment (QOL)	15.27	14.86	.97
Type	3.95	3.95	1.00
Length	2.95	2.91	.99
Employability	2.95	3.18	.995
Relationships with employer	3.18	3.05	.94
Seeking employment	2.23	2.18	.78
Meaningfulness of Employment (QOL)	13.59	13.68	.98
Satisfaction	2.73	2.73	.92
Misses work*	3.77	3.59	.97
Looking for other work	3.82	3.95	.98
Relationships with co-workers	3.27	3.41	.92
Financial Status (QOL)	6.23	6.27	.95
Source of income/reliance on others	3.41	3.41	.93
Basic requirements of living	2.95	2.86	.94
Legal requirements	--	--	.86

*Interviewer and observer means are significantly different from each other at the .05 significance level.

Housing Information

Variable	Interviewer Mean	Observer Mean	Inter-rater Reliability
Housing Stability (QOL)	16.36	15.50	.82
Where living	3.21	3.25	.95
Number of places have lived	4.64	4.25	.16
Missed or late payments	4.21	3.83	.95
Eviction	4.29	4.17	.26
Housing Satisfaction (QOL)	10.71	9.83	.86
Crime/Safety	3.64	3.58	.89
Desire to stay	3.86	3.50	.87
Comfort	3.21	2.75	.54

Significant Other Relationship

Variable	Interviewer Mean	Observer Mean	Inter-rater Reliability
Partner Satisfaction (QOL)	6.86	7.00	.99
Support and trust	2.09	2.14	.98
Long-term future*	2.45	2.64	.98
Length of time	2.32	2.23	.97
Marital status	--	--	.86
Number of children	1.59	1.68	.97
Number of children involved	.77	.73	.91
Partner Support (QOL)**	12.68	9.23	.86
Support for crime-free	3.18	3.14	.88
Support for drug-free	3.09	2.85	.87
Partner's crime-free lifestyle	3.59	3.68	.97
Partner's drug-free lifestyle	2.95	3.00	.93

*Interviewer and observer means are significantly different from each other at the .05 significance level.

**Interviewer and observer means are significantly different from each other at the .01 significance level.

Family Relationships

Variable	Interviewer Mean	Observer Mean	Inter-rater Reliability
Family Satisfaction (QOL)	10.04	10.09	.95
Quality/level of support	3.09	2.95	.90
Frequency	3.59	3.68	.97
Percent of family interactions	3.36	3.45	.88
Family Support (QOL)*	13.50	10.09	.88
Support for crime-free	3.23	3.41	.89
Support for drug-free	3.27	3.41	.71
Family's crime-free lifestyle	3.59	3.42	.92
Family's drug-free lifestyle	3.41	3.38	.89

*Interviewer and observer means are significantly different from each other at the .01 significance level.

Friends and Leisure

Variable	Interviewer Mean	Observer Mean	Inter-rater Reliability
Friends and Leisure (QOL)	13.00	12.64	.96
Engages in prosocial activities	1.55	1.50	.94
Enjoys prosocial activities	2.91	2.73	.83
Activities with friends	2.32	2.52	.84
Friends support for drug-free lifestyle	3.09	3.00	.94
Friends support for crime-free lifestyle	3.14	3.00	.96

Interviewer Impressions

Variable	Interviewer Mean	Observer Mean	Inter-rater Reliability
Interviewer Impressions	8.18	8.43	.70
Openness	3.14	3.05	.55
Honesty	2.73	2.81	.67
Psychological awareness	2.32	2.57	.79

APPENDIX C

Official outcomes for each program were tracked at a 12-month follow-up period. The samples are the same as those found within the body of this report for 6-month official outcomes. The table below lists the percent of offenders who were arrested or incarcerated at any point during the follow-up period.

12-Month Official Outcomes

	Treatment	Controls	<i>p</i>
CIRT			
Technical Violations	27%	28%	n.s.
Misdemeanor Arrests	17%	26%	n.s.
Felony Arrests	16%	19%	n.s.
Incarceration	37%	42%	n.s.
Failure	49%	58%	n.s.
DART			
Technical Violations	11%	19%	<.05
Misdemeanor Arrests	19%	29%	<.05
Felony Arrests	18%	30%	<.05
Incarceration	41%	28%	<.05
Failure	56%	54%	n.s.
RTC			
Technical Violations	31%	30%	n.s.
Misdemeanor Arrests	20%	27%	n.s.
Felony Arrests	16%	26%	<.05
Incarceration	45%	40%	n.s.
Failure	57%	59%	n.s.
STIRRT			
Technical Violations	26%	41%	<.05
Misdemeanor Arrests	30%	23%	n.s.
Felony Arrests	26%	30%	n.s.
Incarceration	34%	46%	n.s.
Failure	58%	68%	n.s.