

The Georgia Cognitive Skills Experiment

Outcome Evaluation

Phase II

Final Report Submitted by:

Patricia Van Voorhis, Ph.D.
Division of Criminal Justice
University of Cincinnati

Lisa M. Spruance, M.S.
Division of Criminal Justice
University of Cincinnati

P. Neal Ritchie, Ph.D.
Department of Sociology
University of Cincinnati

Shelley Johnson-Listwan, Ph.D.
Department of Criminal Justice
University of Nevada-Las Vegas

Renita Seabrook, M.S.
Georgia Board of Pardons and Paroles

Jennifer Pealer, M.S.
Division of Criminal Justice
University of Cincinnati

This study was funded through the Office of Justice Programs for the National Institute of Justice (Grant #98-CE-VX-0013). The report reflects conclusions drawn by the authors and not the Georgia Board of Pardons and Paroles or the National Institute of Justice.

Table of Contents

Acknowledgements	i
Executive Summary	iii
Chapter 1: Introduction _____	2
Background	3
Appropriate Clients	5
Treatment Integrity	8
Description of Georgia Cognitive Skills Program	11
Organization of Report	12
Chapter 2: Methodology _____	13
Research Design	13
Sample.....	17
Parole Participants	17
Pre-release Participants	22
Data Collection	32
Intake Data Sources	34
Program Data Sources	34
Outcome Data Sources	35
Measures	38
Independent and Control Variables	38
Programmatic Characteristics	43
Outcome Variables	58
Data Analysis	63
Chapter 3: Experimental and Control Group Comparisons _____	67
Intermediate Outcomes: Experimental versus comparison groups – Pride In Delinquency Pretest, Posttests	67
Intermediate Outcomes: Experimental versus comparison groups – Colorado Offender Attitude Survey posttests	68
Long-term Outcomes: Returns to Prison	71
Long-term Outcomes: Parolee Experimental and Control Group Comparisons – Rearrests/Revocations	72
Long-term Outcomes: Parolee Experimental and Control Group Comparisons-- Technical Violations	73
Long-term Outcomes: Parolee Experimental and Control Group Comparison – Employment.....	74
Summary	75
Chapter 4: Responsivity: Which Offenders Benefited Most from the Cognitive Skills Program? _____	76
Intermediate Outcomes	76
Long-term Outcomes	83
Summary	90

Chapter 5: The Characteristics of Effective Programs: “What Works?” _____	91
Effects of Program Completion: Returns to Prison and Rearrests.....	93
Effects of Program Characteristics on Program Completion	95
Effects of Class Size: Returns to Prison and Rearrests.....	97
Effects of Participant Performance as Assessed by Coaches: Returns to Prison and Rearrests	98
Effects of Programmatic Characteristics as Assessed by Participants: Returns to Prison and Rearrests	102
Effects of Programmatic Factors on Test Results for Pre-release Women.....	104
Summary	106
 Chapter 6: Discussion and Recommendations _____	 108
Addressing the Research Questions of the Georgia Cognitive Skills Experiment (Phase II)	109
Recommendations	118
Conclusions	123
 References _____	 124
 Appendix _____	 124

List of Tables

Table 1: Distribution of Phase II study participants across districts and study groups	15
Table 2: Comparison of the experimental and control group parolees across background social and demographic characteristics (collected at prison intake).	18
Table 3: Comparison of the experimental and control group parolees across criminal history characteristics (collected at prison intake).	20
Table 4: Comparison of the experimental and control group pre-release study members across sex and background social and demographic characteristics (collected at prison intake).	23
Table 5: Comparison of the experimental and control group pre-release study members across sex and criminal history characteristics (collected at prison intake).	28
Table 6: Observer evaluation scales, results of factor analysis and item analysis	46
Table 7: Participant and Observer evaluation scale and item scores for parole and pre- release samples	48
Table 8 : Participant evaluation scales, results of factor analysis and item analysis	57
Table 9: Pride in Delinquency (PID) and Colorado Offender Attitude Survey scores for parole and pre-release samples.....	59

Table 10: Colorado Offender Attitude Survey scale alpha reliabilities and desired direction for change	60
Table 11: Results of pretest, posttest comparisons (Pride In Delinquency Scale)	68
Table 12: Group differences on the Colorado Offender Attitude Survey	69
Table 13: PID pretest, posttest comparisons by offender characteristics, parolees	77
Table 14: PID pretest, posttest comparisons by offender characteristics, pre-release samples	79
Table 15: Probability of dropping out of the program (experimentals only): logistic regression	83
Table 16: Return to prison by experimental condition, time, and interactions: race, age, and personality type X experimental condition	88
Table 17: Summary of Differential Effects on Program Outcomes	90
Table 18: Individual-level program completion and class-level rates of program completion by program characteristics: summary of several logistic and linear regression models	96
Table 19: Graduate felony arrest/revocation and return to prison by Observer evaluation items, time, and control variables: summary of several logistic regression models	101
Table 20: Graduate felony arrest/revocation and returns to prison by Participant evaluation items	103
Table 21: Class felony arrest/revocation and returns to prison rates by Participant evaluation items	104
Table 22: Programmatic effects of test (PID and COAS) results, pre-release women	105

List of Figures

Figure 1: Data collection plan, Georgia Cognitive Skills Experiment	33
Figure 2. Summary of Interpersonal Maturity Levels	40
Figure 3: Summary of key personality constructs for the collapsed Jesness Inventory types	41
Figure 4. Session evaluation measures	45
Figure 5: Parolee return to prison rates by experimental condition and time	71
Figure 6: Parolee felony arrest/revocation rates by experimental condition and time	72

Figure 7: Parolee technical violations by experimental condition and time	73
Figure 8: Parolees employed full- or part-time by experimental condition and time	74
Figure 9: Parolee return to prison rates by experimental condition, time, control variables, and race: interaction – race X experimental condition	84
Figure 10: Parolee return to prison rates by experimental condition, time, control variables, and age: interaction age 28 to 32 X experimental condition	85
Figure 11a: Parolee return to prison rates by experimental condition, time, control variables, and personality types: interaction – aggressives X experimental condition	86
Figure 11b: Parolee return to prison rates by experimental condition, time, control variables, and personality types: interaction – dependents X experimental condition	87
Figure 11c: Parolee return to prison rates by experimental condition, time, control variables, and personality types: interaction – situationals X experimental condition	87
Figure 11d: Parolee return to prison rates by experimental condition, time, control variables, and personality types: interaction – neurotics X experimental condition	87
Figure 12: Parolee technical violations at 10 to 12 month follow up by experimental condition, control variables, and risk: interaction – risk X experimental condition	89
Figure 13: Parolee return to prison rates by program completion, time, and control variables	94
Figure 14: Parolee felony arrest/revocation rates by program completion, time, and control variables.....	94
Figure 15: Individual and class-level rates of felony arrest/revocation by class size	98
Figure 16: Graduate felony arrest/revocation rates by use of skills, time and control variables	99
Figure 17: Graduate felony arrest/revocation rates by class structure, time and control variables	100

List of Appendices

Appendix A: Session Evaluation Form	129
Appendix B: Observer Evaluation Form	131
Appendix C: Participant Evaluation Form	135
Appendix D: Colorado Offender Attitude Survey	142
Appendix E: Pride in Delinquency	151
Appendix F: Summary of key personality constructs for the original Jesness Inventory personality types	153
Appendix G: Felony arrest/revocation rates and returns to prison by experimental condition and time	154
Appendix H: Parolee technical violations at 10 to 12 month follow up by experimental condition, control variables, and risk: interaction – risk X experimental condition	155
Appendix I: Parolee felony arrest/revocation and return to prison by program completion and time	156
Appendix J: Felony arrest/revocation by class size: graduate (logistic regression) and class (linear regression) levels of analyses	157
Appendix K: Felony arrest/revocation by graduates' use of skills	158
Appendix L: Felony arrest/revocation by class structure: graduate (logistic regression) and class (linear regression) levels of analyses	159

Acknowledgements

This report concludes five years of work for the Georgia Board of Pardons and Paroles. The project extended over two phases of research and has now produced four reports. Phase I was funded by the Georgia Board of Pardons and Paroles (GBPP), through Byrne monies and Phase II was funded by the National Institute of Justice (Grant #98-CE-VX-0013). Over the course of this research it was my extreme good fortune to enjoy a constructive and enjoyable working relationship with officials from the Georgia Board of Pardons and Paroles and with my staff at the Center for Criminal Justice Research at the University of Cincinnati.

Phase II was funded as a “partnership” between the University of Cincinnati and GBPP, and it clearly lived up to those expectations. We were extremely well supported by GBPP officials and technical personnel. I am especially grateful to Beth Oxford who was the Division Director, Community-Based Services, for her wisdom in setting clear organizational expectations regarding the importance of policy-relevant research. James Bralley, Division Director, Field Operations, and John Prevost, the Director of the GBPP Office of Criminal Justice Research, arranged several workshops where project findings were discussed with GBPP officials, parole officers, and cognitive coaches. Researchers seldom have such wonderful opportunities to work with practitioners on the implications of research findings for ongoing practice.

The GBPP data management systems were outstanding. Let me offer a special thanks to “OTIS” and “FLOID” and to John Prevost and Tim Carr for their assistance in downloading accurate background and recidivism data. The remaining data were collected by hand. We wish to thank the many cognitive coaches and parole officers for their assistance in administering assessments and surveys. Renita Seabrook from the Board of Pardons and Paroles and John Prevost coordinated the data collection efforts and resolved numerous glitches. George Braught, Program Evaluation and Research Coordinator, provided valuable reviews of data collection forms and research reports, and Shalandra Robertson, Program Manager, supervised the cognitive

coaches throughout the State.

I am grateful to Lisa Spruance at the Center for Criminal Justice Research, who served as the Project Director for this study. Her talent directed the administrative, technical and creative aspects of this project. In this position, she supervised data processing and data collection efforts, prepared data, conducted much of the analyses, and was a co-author of our reports. Additional thanks are due to our statistical consultant, Dr. Neal Ritchey and two key research associates, Dr. Shelley Johnson Listwan and Jennifer Pealer. Finally, John Schwartz, Research Associate at the Center for Criminal Research, handled most of the contractual, bureaucratic, and financial matters regarding this project. He is the main reason why the rest of us had so much time to work on the project itself.

Executive Summary

Introduction

This document reports the outcome findings for Phase II of the Georgia Cognitive Skills Experiment. It is the fourth and final report of this project. Both Phase I and Phase II of the Georgia Cognitive Skills Experiment utilized a randomized controlled experimental design to compare outcomes for offenders who participated in the Cognitive Skills Program to those who did not. The central question of the Phase I study concerned the overall effectiveness of the Cognitive Skills Program (Van Voorhis, Spruance, Johnson, Ritchey, & Seabrook, 2001); Phase II expanded on this inquiry to determine whether: (a) the program was more effective for some types of offenders than for others, and (b) specific program characteristics could improve offender outcomes.

The following research questions were addressed in the present study:

1. What was the overall effectiveness of the Georgia Cognitive Skills Experiment? How did recidivism rates and test scores for offenders participating in the Cognitive Skills Program compare to those who were assigned to the comparison group?
2. Did the program benefit some types of offenders more than others? Did some types of program participants actually do worse than those who did not participate?
3. What sorts of program and facilitator characteristics differentiated the most effective cognitive programs from the least effective? Do such factors as group size, program completion, and facilitators' skills impact offender recidivism?

The Georgia Cognitive Skills Experiment was implemented by the Georgia Board of Pardons and Paroles (GBPP) as part of a larger commitment to evidence-based practice (MacKenzie, 2000). In addition to implementing a well-tested program, Reasoning and Rehabilitation (Ross & Fabiano, 1985), both phases of the Georgia Cognitive Skills Experiment successfully utilized a randomized controlled experiment, which is considered to be the most rigorous research design available to science. Finally, with the Phase II study, GBPP joined a small handful of correctional agencies to ask important questions regarding differential effects and programmatic influences on offender outcomes. Such research has been advocated for decades (Andrews, Bonta, & Hoge, 1990; Gendreau & Ross, 1987; MacKenzie, 1989; Palmer, 1980, 1994; Van Voorhis, 1987, 1994), but very few studies have actually sought such information (e.g., see Andrews & Keissling, 1980; Palmer, 1974, 2002; Robinson, 1995; Warren, 1983).

The Georgia Cognitive Skills Program (R&R) seeks to modify offenders' impulsive, egocentric, illogical and rigid thinking patterns. In doing so, it shares assumptions common to the larger field of cognitive-behavioral psychology, currently considered to be the dominant therapeutic paradigm in mental health (Dobson & Khatri, 2000). Specific objectives of the R&R program include improving offenders' interpersonal problem-solving, consequential thinking, means-end reasoning, social perspective-taking, critical and abstract reasoning, and creative thinking (Ross, Fabiano, & Ross 1989). The program is structured into 35 lessons that cover seven key components: (a) problem-solving, (b) creative thinking, (c) social skills, management of emotions, (d) negotiation skills, (e) values enhancement, and (f) critical reasoning. The class sessions build on each other in such a way that new skills are presented along with opportunities to practice previously introduced skills. Activities include role-playing, thinking games, homework assignments, and group discussions. The facilitators' handbook (Ross, Fabiano, and Ross 1989), in addition to intensive training of facilitators (coaches), attempted to foster consistency across program sites.

Methodology

Research Design.

From July 1998 to April 2000, 28 sites across Georgia participated in Phase II, including 25 parole

districts and three pre-release units housed in correctional institutions. In addition, 98 parolees attended the program in sites that did not participate in the experimental design. Those participants not included in the random assignment design are excluded from the analyses comparing experimental to control participants¹.

Sample

The sample included 1,155 parolees and 192 pre-release, prison inmates who were randomly assigned to either the treatment group or the control group. For parolees, this resulted in 574 experimental participants and 581 comparison group participants. Pre-release participants included 104 experimentals and 88 comparisons. Selection criteria for the study pool sought high-risk offenders, those with IQ scores (Culture Fair) (Cattell & Cattell, 1973) above 80, no history of sexual offending, and an absence of substance abuse problems so severe as to interfere with the offender's ability to attend or comprehend the program.

The parole study participants were predominantly African American. Their average age was 31.9. All participants were male. At the time of their arrest, 54.0% of the participants were employed on a full-time bases. At prison intake, 41.1% were rated as middle class; 63.9% were married. The majority of the parolees (66.5%) had less than a high school education. Notwithstanding GBPP directives to exclude parolees with low IQ scores, 6.2% of the sample scored at 80 or below on the Culture Fair Test (Cattell & Cattell, 1973). Even so, 70.9% of the sample read at the fifth grade reading level or higher. Jesness Inventory results found 61.4% of the parolees diagnosed at I-levels 2 or 3, indicating that cognitive and interpersonal maturity was atypically low even for a criminal population. JI personality diagnoses observed: (a) 31.7% aggressive (antisocial values, internalized criminal lifestyle), (b) 12.7% neurotic (high anxiety), (c) 29.7% dependent (easily led, immature), and (d) 25.9% situational (prosocial value system; crime is the result of adversity, poor coping skills, and perhaps substance abuse). Most participants (78.8%) scored at medium to high risk on a risk scale developed for this study. All but two participants had a prior felony on record and 56.9% of the sample had served at least one prior prison term. Prior convictions for at least one sex offense characterized 32 participants (3.4%), even though screening criteria attempted to exclude such individuals.

There were no significant differences between the comparison group and the experimental group on background or criminal history characteristics. Therefore, random assignment appears to have been strictly adhered to.

Approximately half (51.0%) of the prerelease sample were women. The majority were non-white (59.2%), single (56.1%), unemployed at the time of their arrest (66.4%) and living at a minimum standard of living or below (66.0%), with histories of substance abuse (65.3%). Most (63.3%) had less than a high school education. IQ scores below 80 were reported for 23 offenders (24.0%), and a similar proportion obtained reading scores below the fifth grade. The average age for this sample was 34.1 years. I-level diagnosis found higher levels of Interpersonal Maturity than observed for the male parole and pre-release samples. The women's sample also received more neurotic classifications and less aggressive or situational ones.

The average age of the pre-release males was 31.9 and most were non-white (64.9%). In contrast to the women's sample, the majority of male pre-release participants (53.9%) were employed full-time at their arrest. However, only 31 (34.4%) had a high school education, and 55 (60.5%) lived at the minimum standard of living or below. Only 5 pre-release, male offenders (5.4%) had IQ scores below 80, and the majority (89.1%) read at the fifth grade reading level or above. I-Level classifications were similar to those for the parole sample; however, pre-release males were observed to have more aggressive (43.7%) and neurotic (23.9%) classifications and less situational (12.7%) and dependent ones (19.7%).

Male and female pre-release, experimental and comparison groups evidenced similar demographic characteristics. This was not the case with respect to their criminal histories, however. Experimental

¹ The non-randomly assigned participants who completed the program are, however, included in the analyses of programmatic characteristics in chapter 5.

participants were more likely to have serious criminal histories than the comparison group members. For the women's sample, these differences were more pronounced than for the men's sample. For example, 38 (74.5%) members of the experiment group were classified as medium to high risk as compared to only 24 (58.5%) of the comparison group. The experimental group evidenced significantly more prior felonies (mean = 3.1) than the comparison group (mean = 2.2). Male experimental offenders also were more likely than comparisons to have: (a) classifications as medium/high risk, (b) prior incarcerations, and (c) prior violent convictions. None of these differences, were statistically significant. These group differences suggest that the experimental design may have been compromised for both the male and female pre-release samples. Another possibility is that the proportions and means characterizing the pre-release sample are not stable due to the small sample sizes. Subsequent data analyses control for these group differences.

Measures

Multiple information sources addressed the research aims of this present study. The evaluation required intake measures pertaining to offender demographic and social characteristics as well as criminal histories. Most of these were downloaded from OTIS, the State's offender-based tracking system. In addition the Jesness Inventory (Jesness, 1992) furnished personality assessments, and IQ scores were provided by the Culture Fair Test (Cattell & Cattell, 1973).

Programmatic measures included the coaches' accounts of group size, program completion, participation levels, enjoyment, levels of understanding, use of skills, and group climate. At the close of class, participants completed a participant evaluation form. Factor analysis of the items contained on this form yielded three scales, measuring: (a) adherence to social learning methods, (b) relevance, and (c) group climate. Finally classes were observed on two occasions. The items on the observers form formed four measures: (a) positive class atmosphere, (b) organization and adherence to the curriculum, (c) structure (adherence to rules), and (d) skill in dealing with difficult participants. Many of the program measures showed limited variability. That is most coaches, participants, and observes found most classes to be excellent.

Intermediate outcome measures consisted of pretest, posttest changes on the Pride in Delinquency Scale (Simourd, 1987), and posttest results on the Colorado Offender Attitude Survey (COAS)(Johnson & Hunter, 1982). These measures proved to be more valid for women offenders than for either the parolees or pre-release males.

Longterm outcome measures obtained from OTIS indicated whether parolees had been returned to prison during the follow up period. In addition, rearrest/revocation, employment, and technical violations were obtained from the GBPP FLOID database.

Results

Research findings are summarized according to the research questions for this study.

Question 1:

What was the overall effectiveness of the Georgia Cognitive Skills Experiment? How did recidivism rates for parolees participating in the Cognitive Skills Program compare to those who were assigned to the comparison group? How did the groups perform on the intermediate measures of outcome?

For parolees, event history analysis detected a 3.3% difference between experimental and comparison group returns to prison by 30 months following the program. Differences in arrest/revocation rates between the two groups at the end of 12 months was 2.5%. Neither difference was significant ($p \leq .05$). These modest effects are consistent with those found for Phase I of the project, and they remain insignificant even when the Phase I and Phase II samples are combined. Results are shown in Figures 1

and 2, below.

Figure 1: Predicted parolee return to prison rates by experimental condition and time.

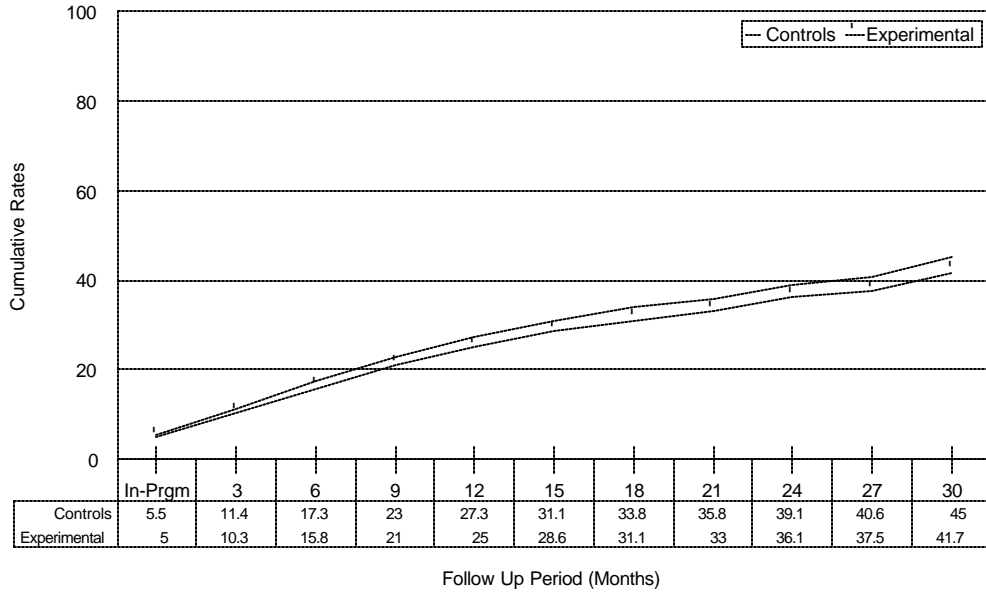
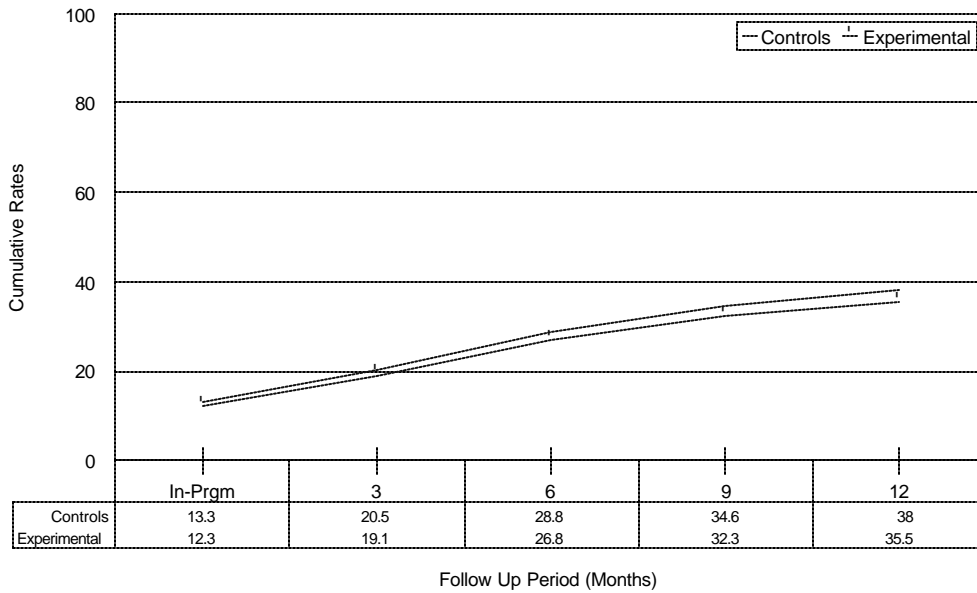


Figure 2: Predicted parolee felony arrest/revocation rates by experimental condition and time.

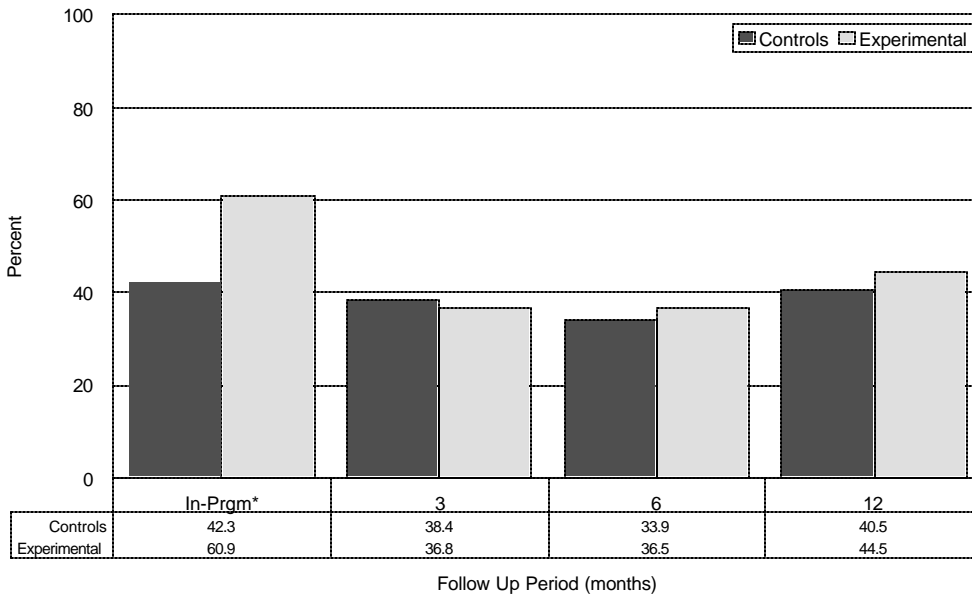


These results, along with those for Phase I, were in line with the two other large-scale R&R initiatives, one in Canada and another in Colorado. The Canadian study achieved a treatment effect of 5%, among parolees (Robinson, 1995), and the Colorado study found an effect of 3% (Johnson & Hunter, 1992). This trend has already come to the attention of meta-analysis researchers; two recently published studies report that large cognitive programs for offenders, of which R&R is only one model, found

treatment effects between 3-5% (see Wilson et al., 2000; Lipsey et al., 2001). The concern, voiced by many, is that the larger programs may have a more difficult time addressing and maintaining treatment integrity. In other words, they may be more likely to depart from appropriate screening and curriculum protocols than smaller, more contained, programs.

Figure 3 reports a significantly higher technical violation rate for experimental versus control parolees at the “during program” time period (chi square = 31.5, $p \leq .001$). Sixty-one percent of the experimental group received at least one technical violation while the cognitive skills program was in session, compared to 42% of the control group. During the post-program time periods similar proportions of experimental and control groups received technical violations.

Figure 3: Parolee technical violations by experimental condition and time.



Through each of the four time periods examined, rates of employment were high (greater than 70%) for both the experimental and control groups. Participation in the Cognitive Skill Program did not impact parolees’ employment rates.

Using MANOVA, ANOVA, and ANCOVA statistical tests, we also examined the intermediate outcomes (test results taken at the end of the class) for the pre-release and parole samples. As noted above, we had reservations about the validity of pretest and posttest measures for parolees and pre-release males, so they do not receive unwarranted discussion at this point. COAS results for pre-release females, however, were significantly better for e’s than c’s on four of the ten scales studied. PID pretest, posttest changes were not significantly different between the two groups, however, the null findings were primarily attributable to an interaction between risk and treatment assignment (see below). We do not at this time, have the recidivism data needed to assess the longterm outcomes for the pre-release samples..

Question 2:

Did the program benefit some types of offenders more than others? Did some types of program participants actually do worse than those who did not participate?

The parolees who achieved the most impressive treatment gains were: (a) white, and (b) between the ages of 28 and 32. Moreover, neurotic offenders, as classified by the Jesness Inventory (Jesness, 1996) who participated in the Cognitive Skills Program, had considerably higher recidivism rates than the comparison group members who did not participate.

Responsivity also factored into program attrition. Parolees who dropped out of the program were significantly more likely to be: (a) middle class, (b) neurotic, (c) medium to high risk, and (d) between 23 and 27 years of age.

Responsivity characteristics of women pre-release offenders followed the established screening recommendations for R&R: (a) Largest treatment effects were achieved by offenders with IQ scores above 85; (b) low risk e's achieved worse posttest scores than low risk c's. In addition, a race effect was detected for women pre-release offenders as well as parolees; white women offenders had much better posttest results than nonwhite offenders.

Question 3:

What sorts of program and facilitator characteristics differentiated the most effective cognitive programs from the least effective? Do such factors as group size, program completion, attendance, and facilitators' skills impact program completion, test results, and offender recidivism?

The following program characteristics differentiated successful programs from unsuccessful ones.

Adherence to the program design. We are not convinced that the Georgia programs had difficulties with "program drift", which is the central concern raised with regard to correctional treatment integrity (e.g., Palmer, 1978, 1992; Van Voorhis et al, 1995). Participants told us that, in most instances, they had opportunities to: (a) observe a skill demonstrated, (b) to practice the skill, and (c) to receive feedback regarding how well the skill was performed. They told us, in other words, that the classes employed the clinical dimensions dictated by their underlying cognitive-behavioral theory. Observers gave high ratings on program qualities pertaining to adherence to the curriculum. *When their ratings dropped even modestly, however, the reduced scores were significantly associated with lower completion rates.*

Relationship qualities and relevance of the material. Both participants and observers rated coaches high in qualities such as empathy and sensitivity to participants. Such findings indicated that the coaches demonstrated high relationship qualities. Finally, offender ratings indicated that the material was highly relevant to their lives. We did not detect significant relationships between these findings and offender outcomes, however.

Classroom Structure and Control. Difficulties clustered around the coaches' ability to keep good discipline and structure within the classroom. Classes that did not maintain and enforce rules regarding tardiness, absences, and unruly classroom behavior were considerably less successful than those that did. Classroom control was an area where evaluations showed the lowest score (Spruance et al., 2001) and one that affected offender outcomes. The study also speaks to the notion of "professional detachment," a quality which allows correctional workers to be empathic and approachable, but discourages tendencies to be manipulated by offenders or pulled into offender thinking (thereby modeling criminogenic thinking). Leaders with such skills were significantly more likely to have more successful classes.

Class Completion. Related to the issue of classroom control, some offenders appeared to have been dismissed from class too readily. We are not speaking, here of those who committed a new offense during the program, but rather those who were dismissed to pursue employment opportunities or presented other, non criminal reasons to excuse their participation. Presumably program staff had some discretion over whether these individuals could be dismissed from the class. The offenders dismissed by coaches for non-criminal reasons showed recidivism rates that paralleled those for the comparison group. The recidivism rate for the program graduates, however, was quite low ---24.8% for rearrests and 28.8% for returns to prison. These findings recommend for mandatory rather than voluntary participation. Results are shown in Figures 4 and 5, below. A quasi experimental design, controlling for group differences was required in order to conduct these analyses.

Figure 4: Predicted parolee return to prison rates by program completion, time, and control variables.

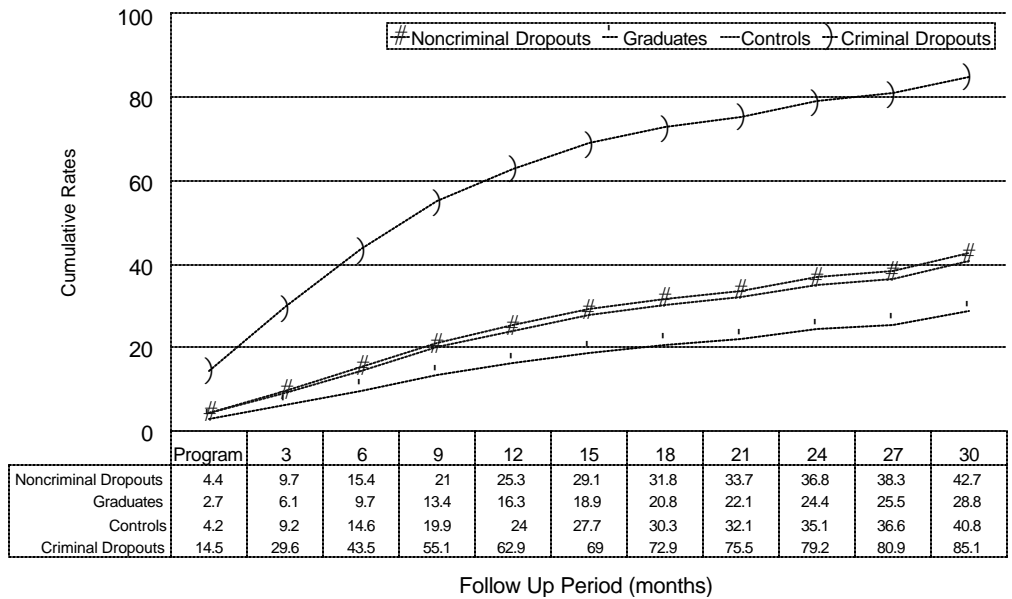
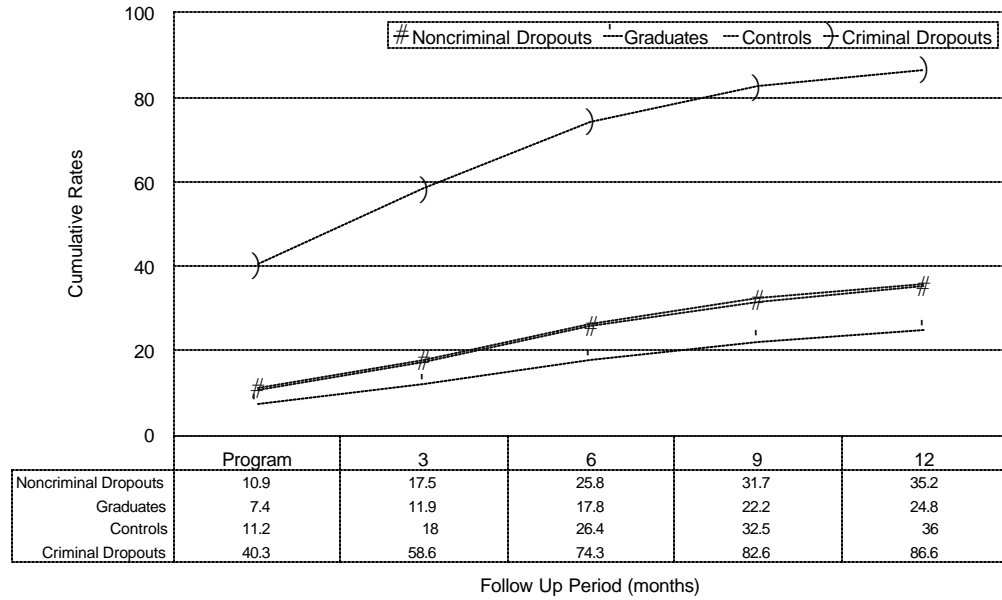
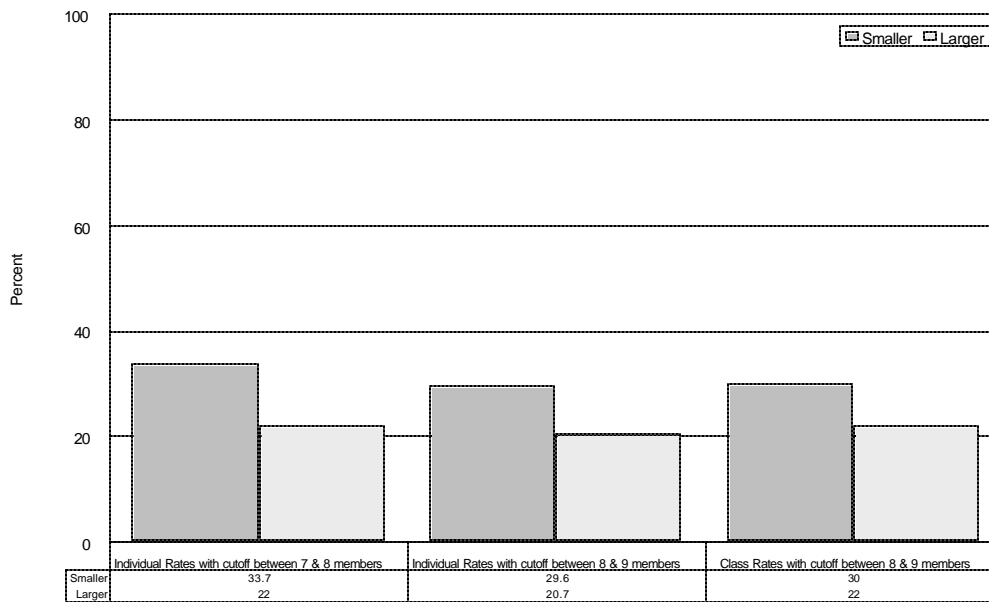


Figure 5: Predicted parolee felony arrest/revocation rates by program completion, time, and control variables.



Class size: The literature on group dynamics puts forward the well established assertion that the ideal size for groups targeting behavioral change should be between eight and 12. The manual for the Cognitive Skills Program recommends between four and eight participants. However, our study found that groups smaller than eight did not achieve the treatment gains that groups of eight or larger did. Figure 6 shows that the optimal cut point was between 7 and 8.

Figure 6: Individual- and class-level rates of felony arrest/revocation by class size.



These results should not be interpreted to mean that overly large groups are optimal. The “large” groups in this study averaged 10 members. Even though the Georgia class sizes ranged from 5 to 21, only 3 groups contained 15 or more members. Therefore, we are not asserting that groups above 15 work.

Practice and offender participation. The final series of significant program characteristics clustered around the extent to which participants were engaged in the program. Participants who were

characterized by coaches as unable to demonstrate previously taught skills, and who, themselves, reported that the skills were too difficult were more likely to recidivate than those who could demonstrate and understand the skills, even when we controlled for IQ and reading scores. Classes that were characterized by participants as having more rather than less discussion were also more likely to be successful. It didn't matter who made this observation, coaches, observers, or participants. Classes with more participation were more likely to reduce recidivism.

Classes high in participation, enjoyment, level of understanding, and use of previously taught skills also experienced less attrition. This occurred whether participation was rated by the coach or the observer.

Perhaps this concern also relates to the participant's ability to practice skills in an active, and engaged manner. We can assert, however, that the effects of participation were independent of IQ and reading level, because our analysis was controlled for these measures of ability. Therefore, this pattern of findings begins to implicate the amount of time devoted to practice and processing of skills.

The final program characteristic differentiating between success and failure concerned how well coaches who team-taught worked together. Successful programs had coaches who in the estimation of outside observers, cooperated with each other and demonstrated similarly high levels of involvement in classroom activities. At the same time, enthusiastic coaches achieved less class attrition.

Test outcomes for women offenders were impacted by coaches' adherence to the curriculum, the relevance of the class, and the group climate. As with parolees, these classes benefited from adherence to and enforcement of rules balanced with a generally favorable class atmosphere.

Recommendations

Given these findings, the following recommendations were put forward for consideration by personnel of the Georgia Board of Pardons and Paroles. Please note that they are given important discussion in the concluding chapter of the larger report.

1. Plan for important differences among offenders.

"Responsivity," the consideration of different learning styles among offenders, is a well-touted but mostly undeveloped notion in correctional treatment. Even so, the discovery of an individual attribute that suggests that a certain type of individual (e.g., low risk) is less appropriate for a program than another (e.g., medium/high risk) has two program implications. First, the program could use such findings to screen nonamenable individuals from program participation. It appeared that neurotic offenders in the program performed considerably worse than non-neurotic offenders in the comparison group. This is the first study to look at or detect this, however. Second, officials could identify ways to make the program more appropriate to such individuals. Such an approach would be appropriate to concerns regarding race and age, because the non-amenable did not do *worse* than their counterparts in the comparison group. They just didn't benefit as much as whites or parolees between the ages of 28 and 32.

2. The findings continue to recommend against voluntary participation or dismissal for employment or other non-criminal reasons.

When an offender ceases to participate because he or she has been arrested or revoked, there may be nothing that coaches or parole officers can do to keep the offender in the class. However, dismissals for employment, unruly classroom behavior, childcare, transportation, and other non-criminal reasons didn't work. As with the Phase I study, such individuals had higher recidivism rates than those who graduated. In fact, findings for the present study find this group has a recidivism rate comparable to the comparison group.

3. *Maintain class sizes at no less than 8 participants.*

Small classes did not achieve optimal outcomes among parolees. However, classes containing 8 to 12 members may be small enough to provide opportunities for interaction, practice, and full participation. They provide more role models and a greater diversity of role models. At the same time, small classes provide participation opportunities only at the cost of participants feeling constantly “on stage”. Moreover, if small groups contain reluctant participants or poor role models there is no fall-back to other class members. Our findings also underscore the fact that coaches are not the only vehicle for change.

4. *Participation and practice to the point of skill mastery appears to be an important component of groups activity.*

The importance of practice, feedback, and reinforcement for mastery of skills is deeply embedded in cognitive behavioral and social learning theory (Bandura, 1977; Meichenbaum & Jaremko, 1982), the theoretical models underlying the Georgia Cognitive Skills approach. Our findings further advocate for group activity, discussion, role plays, and practice dealing with real-life situations. Groups characterized by more facilitator talk and less active involvement from group members are less likely to be successful as those groups which promote active learning. Individuals experiencing difficulties with the skills or those who are reluctant to become engaged in group work should receive additional attention rather than be allowed to sit on the sidelines. These results not only encourage the provision of group activities but the monitoring of participants for their level of understanding.

5. *Identify ways to help coaches who are having difficulties with classroom control.*

Classroom control was the dimension most likely to be negatively rated by outside observers. Even though observers ratings were not overly negative, a modest level of difficulty decreased program effectiveness. These findings recommend for: (a) ongoing attention to classroom control in GBPP trainings, (b) monitoring of groups by personnel with higher levels of expertise and experience, and (c) in-service assistance to those coaches who are having difficulties. Moreover, classroom control should place a special emphases on expectations, rules, and enforcement of rules.

6. *Ongoing attention to quality assurance and prevention of “program drift.” should involve additional training of program observers or clinical supervisors.*

Monitoring or clinical supervision should be an ongoing programmatic task, one that occurs even when program evaluations are not in process. GBPP would be wise, in other words, to continue the observation process. However, ongoing training of supervisors and observers is essential. Such training should deal with the precision and characteristics of skills required of coaches and classes, as well as their feelings toward evaluating colleagues.

7. *Pretest, posttest, participant evaluations, and observer evaluations will require more attention to the integrity of test conditions.*

We have learned that two different assessments would be especially useful for pretest and posttest purposes. These are Criminal Sentiments Scale (Simourd, 1997) and the Psychological Inventory of Criminal Thinking Styles (Walters, 2002). Both appear to have performed better in other studies than our tests in the present research. We would be even more critical of the tests used in the present study, had they not worked exceptionally well in *one* of the program sites. The fact that PID, the COAS, and the Participant Evaluation forms proved to be extremely meaningful among women pre-release offenders unavoidably poses some questions regarding the integrity of these data in other settings. Tests must be administered under carefully controlled conditions. When they are not, they are likely to be invalid.

Chapter 1

Introduction

This document reports the outcome findings for Phase II of the Georgia Cognitive Skills Experiment. It is the fourth and final report of this project. Both Phase I and Phase II of the Georgia Cognitive Skills Experiment utilized a randomized controlled experimental design to compare recidivism rates of offenders who participated in the Cognitive Skills Program to those who did not. The central question of the Phase I study concerned the overall effectiveness of the Cognitive Skills Program (Van Voorhis, Spruance, Johnson, Ritchey, & Seabrook, 2001); Phase II expanded on this inquiry to determine whether: (a) the program was more effective for some types of offenders than for others, and (b) specific program characteristics could improve offender outcomes. In other words, Phase II continued to ask the question of overall effectiveness, but a focus on differential effectiveness and programmatic determinants of effectiveness was undertaken to assist the State's efforts to improve or fine-tune its cognitive programs. Findings suggesting, for example, that successful offenders differed from unsuccessful offenders on the bases of age, IQ, personality, race, gender, criminal history, educational status, or substance abuse history could influence future criteria for screening parolees into this program or attempts to make the sessions more responsive to these groups. At the same time, the discovery of certain programmatic features, such as group size or facilitator styles, could inform efforts to train staff and monitor sessions.

The Georgia Cognitive Skills Experiment was implemented by the Georgia Board of Pardons and Paroles (GBPP) as part of a larger commitment to evidence-based practice (MacKenzie, 2000). As a first step in this process, the Agency implemented a program that had impressive empirical support. The Georgia Cognitive Skills model is a replication of Reasoning and Rehabilitation (Ross & Fabiano, 1985), a Canadian program, found through several earlier evaluation studies to be effective in reducing offender recidivism (see Research and Statistics Branch, Correctional Services of Canada, 1991; Porporino, Fabiano, & Robinson, 1991; Robinson, 1995; Robinson, 1995). In addition to implementing a well-tested program, both phases of the Georgia Cognitive Skills Experiment successfully utilized a randomized controlled experiment, which is considered to be the most rigorous research design available to science. Finally, with the Phase II study, GBPP joined a small handful of correctional agencies to ask important questions

regarding differential effects and programmatic influences on offender outcomes. Such research has been advocated for decades (Andrews, Bonta, & Hoge, 1990; Gendreau & Ross, 1987; MacKenzie, 1989; Palmer, 1980, 1994; Van Voorhis, 1987, 1994), but very few studies have actually sought such information (e.g., see Andrews & Keissling, 1980; Palmer, 1974, 2002; Robinson, 1995; Warren, 1983).

The following research questions were addressed in the present study:

4. What was the overall effectiveness of the Georgia Cognitive Skills Experiment? How did recidivism rates for parolees participating in the Cognitive Skills Program compare to those who were assigned to the comparison group?
5. Did the program benefit some types of offenders more than others? Did some types of program participants actually do worse than those who did not participate?
6. What sorts of program and facilitator characteristics differentiated the most effective cognitive programs from the least effective? Do such factors as group size, program completion, and facilitators' skills impact offender recidivism?

Background

Our previous project reports (Van Voorhis, Murphy, & Johnson, 1999; Van Voorhis, Spruance, Johnson, Ritchey, & Seabrook, 2001; Spruance, Van Voorhis, Johnson Listwan, Pealer, & Seabrook, 2001) furnished extensive reviews of the evaluation literature regarding cognitive skills programs. These are not repeated in this report. Since the commencement of the Georgia Cognitive Skills Experiment, however, two meta-analyses specific to cognitive behavioral programs (Lipsey, Chapman & Landenberger, 2001; Wilson, Allen & MacKenzie, 2000) reported an important trend. According to both, treatment effects (difference in recidivism rates between the experiment and comparison groups) for small programs were far more substantial than those achieved by large-scale initiatives, e.g., those implemented on a statewide or agency-wide basis. Both studies observed that the poorer performance of the larger programs appeared to implicate their treatment integrity. That is, smaller programs serving less than 75 participants may have demonstrated stronger adherence to the program protocols than those programs administered to larger populations. For example, the largest study conducted to date, an evaluation of Reasoning and Rehabilitation (Robinson, 1995), involved over 2000 offenders and achieved a treatment effect of 5%. In addition to program size,

Lipsey et al., (2001) observed that the smaller-scale, more effective programs were also “demonstration projects” set up by the researchers for “the limited time required to mount an evaluation of their effects” (pg. 154).

When we limit these trends to Reasoning and Rehabilitation, the prototype for the Georgia Cognitive Skills Program, we see that the largest treatment effect was observed for the first study of R&R (Ross, Fabiano, & Ewles, 1988). In this study, 75 high-risk probationers were randomly assigned to (a) regular probation, (b) regular probation augmented by a life skills program, and (c) regular probation augmented by R&R. Recidivism rates were: (a) 69% for the probation only group, (b) 47% for the probation plus life skills group, and (c) 18% for the probation plus R&R group. The recidivism rate for the probation group augmented by R&R was 51 percent lower than the recidivism rate for the group receiving only probation. As noted above, by the time R&R was implemented across Canada to over 2000 parolees, the treatment effect was a modest 5 percent (Robinson, 1995). Another large-scale implementation of R&R involved substance abusing probationers in Colorado. Participants were randomly assigned to: (a) intensive supervision probation and no treatment, (b) intensive supervision augmented by R&R, and (c) regular probation. Adding R&R to intensive probation reduced recidivism rates by only 3%.

Both phases of the Georgia Cognitive Skills Experiment involved large samples. Phase I consisted of 468 parolees, and Phase II, the present study, included 1,190 parolees and 192 pre-release, prison inmates. Results for Phase I were in keeping with those observed for other large-scale interventions. By the end of a 30-month follow-up period, returns to prison were 3.9 percent lower for the experimental group than the comparison group (Van Voorhis et al., 2001).² With this additional study, the 3-5% treatment effect appeared to be stable across large programs (Wilson et al., 2000). However, authors remind policy makers that even small effects (3-5%) are likely to have important benefits, especially when they pertain to the large numbers of offenders participating over the full life of such interventions (see Gendreau & Ross, 1987 & Lipsey et al, 2001).

Even so, the haunting gap between the large effects of the small programs and the small effects of the large programs prompts the obvious question: How can this gap be reduced? If the authors of the meta

² Differences between the participants who completed the group and the comparison group, however, were 16.6% (Van Voorhis et al., 2001).

analyses were correct in implicating treatment integrity, suggesting that the larger programs cannot address issues of quality control as effectively as a smaller programs, it makes sense that the effectiveness of the larger programs would be diminished.

Two programmatic oversights could have eroded the treatment effects of the larger programs. First, the programs may not have screened for appropriate participants. Staff may have neglected the established screening criteria, or they could have been unaware of additional individual attributes that might have affected offenders' abilities to participate. Second, the program may not have been administered according to the directives of the program designers. For example, violations of guidelines regarding class size, classroom control, and client reinforcement, are often indicative of "program drift" into some other model that was not envisioned by its designers (Palmer, 1992; Van Voorhis, Cullen, & Applegate, 1995). These two issues are discussed below.

Appropriate Clients

A recently-coined term, "responsivity" (Andrews, Bonta, & Hoge, 1990), has directed policy makers and practitioners to the search for programs that work best for offender populations. The literature has referenced two types of responsivity---general responsivity and specific responsivity. General responsivity maintains that some programs are amenable to offender populations in general, and some are not. In fact, the identification of programs amenable to a wide array of offenders may have been the most important contribution of recent meta-analyses of offender programs. With tremendous consistency across such studies, social learning and cognitive behavioral programs were noted to be far more effective than psychoanalysis, person-centered therapies, and punishing smarter approaches to crime control (e.g., Andrews, Zinger, Hoge, Bonta, Gendreau & Cullen, 1990; Lipsey, 1995).

Although one might assume that a program noted to be generally responsive would, by definition, work with *all* offenders, studies have noted that most cognitive-behavioral and social learning approaches should be matched to medium to high risk offenders (Andrews & Friesen, 1987; Andrews & Kiessling, 1980; Research and Statistics Branch, Correctional Services of Canada, 1991), and to offenders whose IQ scores surpass 80 (Ross, Fabiano, & Ross, 1989). There have been exceptions. The large Canadian study

(Robinson, 1995) and Phase I of the Georgia Cognitive Skills Experiment (Van Voorhis et al., 2001) did not observe the expected “risk effect”.³ There, low risk offenders achieved larger treatment effects than high risk offenders (Robinson, 1995). Van Voorhis et. al, (2001) found R&R to be equally effective for both medium to high risk offenders and low risk offenders. Interestingly, both were studies of parole populations.

Age was also found to differentiate successful from unsuccessful participants. For example, Robinson (1995) found R&R to be least effective among offenders younger than 25 or older than 39; similarly Johnson & Hunter (1992) reported the program to be least effective among offenders less than 30 years of age.

The admonition against treating offenders as if they were all alike—a “one-size-fits all” mentality – has existed since the 1950s. Then the term “differential treatment” recommended that offenders be classified according to key personality and psychological attributes and then matched to programs and case managers, accordingly (Palmer, 1974; Warren, 1971). A small number of early studies found that program outcomes varied by such psychological attributes as personality and cognitive complexity (Andrews & Kiessling, 1980; Heide, 1983; Palmer, 1974, 2002; Van Voorhis, 1985; Warren, 1983). The term “differential treatment” has more recently been called “specific responsivity” (Andrews & Bonta, 1998) and has coincided with a “principle of responsivity” which asserts that “styles and modes of service must be matched to the learning styles and abilities of offenders” (Bonta, 1995; Kennedy, 2000). The concern here is for how characteristics such as motivation, personality, intelligence, age, gender, and ethnicity might affect an offender’s ability to succeed in even those programs which he or she needs.

Notwithstanding the popularity of the term, researchers were not quick to study specific responsivity (Andrews et al., 1990; Van Voorhis, 1987). However, a substantial body of research emanated from the California Youth Authority, during the 1960s, when researchers found that evaluation outcomes differed according to cognitive maturity and offender personality types (Empey & Erickson, 1972; Grant,

³ The “risk effect” refers to a frequently occurring research finding whereby much greater treatment effects are observed for medium to high risk offenders than for low risk offenders. Risk effects may also involve situations where low risk experiment participants do worse than low risk comparisons.

1965; Jesness, 1971; Palmer, 1974; Warren, 1983).⁴ Other studies noted interactions between specific interventions and empathy (Andrews & Kiessling, 1980), anxiety (Andrews, 1980, Sarason & Ganzer, 1973, Heide, 1982), Interpersonal Maturity Level (Heide, 1982), Conceptual Level (Reitsma-Street & Leschied, 1988), and Moral Development (Van Voorhis, 1985).

Simply put, when evaluation research entertains the notion of specific responsivity, researchers recognize that treatment effects can be masked by interventions that are effective among some offenders and ineffective among others. The gains achieved by offenders amenable to the intervention are then cancelled out by the poor results achieved by those who are not amenable. The result is an unimpressive treatment effect when in fact the program *did* work with some offenders. The implications of such research suggests that programs either (a) develop new policies to screen out those who are not amenable, or (b) amend program procedures and components to help accommodate those who are having difficulty with the program.

Previous research has not addressed sufficiently the effects of race and gender. There are few reports of whether interventions were more successful for African Americans than for whites, or for men than for women. Seldom have studies partitioned research populations according to race, even when the data were available to do so. At the same time, most research on correctional effectiveness has been limited to male offenders (Chesney-Lind, 2000; Morash, Bynum, & Koons, 1998). One meta-analysis of a small number of women's programs suggested that cognitive and social learning programs were effective for women (Dowden & Andrews, 1999), yet Lipsey et al. (2002) noted that almost all of the published controlled studies of cognitive behavioral programs sampled male offenders. At the same time, another group of scholars have questioned the appropriateness of cognitive behavioral programs for women offenders (Bloom & Owen, 2002; Covington, 2000). Whatever one's opinion might be, the question simply has not been studied adequately.

⁴ Interpersonal Maturity (I-level) (Jesness, 1971, 1996; Warren et al., 1966; Warren, 1983) is an assessment/classification system used to classify offenders according to personality and cognitive maturity.

Treatment Integrity

In order for programs to work they must be administered according to the specifics of their design. Facilitators must demonstrate skills they were trained to utilize, the program must be given to the clients it was intended to serve, and clients must attend and complete the program. The potential seriousness of program integrity problems has been underscored by most early evaluations of correctional treatment programs. Many of these studies examined outcomes such as recidivism, but never considered program quality. This situation became problematic for two reasons.

First, results of such evaluations were extremely misleading. Absent a measure of program quality, outcome studies created the impression that “nothing worked” (e.g., Martinson, 1974) when, in fact, nothing happened (Palmer, 1992; Van Voorhis, Cullen, & Applegate, 1995). Many such programs might have worked if properly delivered. The importance of a high quality service delivery was demonstrated recently by the Violent Juvenile Offender Project. This program existed in four sites but only showed success (reduced recidivism) in two. The factor that differentiated success from failure, in that study, was the quality of the programs’ delivery (Fagan, 1990).

A second problem associated with ignoring program processes and quality is the lost opportunity to learn about the conditions under which program models work best. Such findings contribute tremendously to the planning, implementation, facilitation, and administrative tasks associated with delivering programs. Unfortunately, comprehensive studies designed to identify optimal program conditions, and the types of offenders most amenable to specific program models have been surprisingly rare in corrections. They include Andrews and Kiessling’s (1980) study of the CaVIC experiment, as well as a series of evaluations of California Youth Authority programs during the 1960s and 1970s (e.g. Empey & Erickson, 1972; Jesness, 1971, Palmer, 2002) and to a lesser extent, the Canadian research on the cognitive skills programs (Robinson, 1995). In addition, several meta-analysis and reviews of the evaluation literature (see Andrews et al., 1990; Gendreau & Ross, 1987; Lipsey, 1989; Palmer, 1992) identified a number of general program/agency attributes associated with success. Although not specific to cognitive programs, programmatic characteristics differentiating successful from unsuccessful programming were as follows:

- Longer interventions. (Lipsey, 1989; 1995).
- Community rather than institutional programming (Andrews et al., 1990; Lipsey, 1989).
- Behavioral-oriented, skill-oriented, and multi-modal treatment (Andrews et al., 1990, Lipsey, 1989; Palmer, 1992).
- Service targeted to high risk clients (Andrews, Bonta, & Hoge, 1990; Lipsey, 1989; Andrews et al., 1990).
- Service targeted to criminogenic needs (Andrews, Bonta, & Hoge, 1990, Andrews et al., 1990).
- Services matched to offender learning and personality styles (Palmer, 1992, Andrews et al., 1990 Van Voorhis, 1987).
- Services not solely geared to a goal of deterrence (e.g., “Scared Straight”)(Andrews et al., 1990; Lipsey, 1989; Gendreau & Ross, 1987).
- Services to systems such as families which impact the lives of offenders (Lipsey, 1989; Andrews et al., 1990; Palmer, 1992; Van Voorhis, 1987).

The lessons of the meta-analyses and treatment reviews went far toward identifying the dimensions of program integrity. In fact, many of the criteria for program success have been incorporated into a program assessment tool, the Correctional Program Assessment Inventory (CPAI) (Gendreau & Andrews, 1994). As with the meta-analyses, the CPAI looks to the environment of a program (e.g., community support, leadership qualities and qualifications) as well as to within-program qualities involving facilitators’ use of appropriate reinforcement and punishment strategies and appropriate screening and classification tools. Studies have found that agencies with more favorable CPAI scores tend to achieve lower recidivism rates (Fulton, Latessa, Stichman & Travis, 1997; Holsinger, 1999).

Even though the meta analyses afforded a view of the more global characteristics of successful programs, research on the program qualities associated with the success of specific types of program has not been forthcoming. Such research would have involved numerous replications of a single program model under different program conditions (e.g., parole versus probation) and with different subjects (e.g., males versus females). The dearth of experimental replications of existing program models, according to Palmer (1992), seriously hindered chances of : (a) assuring that programs were appropriate across a variety of populations and settings, and (b) executing program modifications where needed.

Where does one look for guidance in executing sound cognitive behavioral programs if research on program processes is so limited? Directives, to date, have emerged from (a) the social learning and cognitive behavioral theoretical foundations of these programs, (b) a limited array of research pertaining to theoretically relevant components of the programs, and (c) the program manuals which guide and train group facilitators. The psychological theories directed programs to utilize appropriate contingencies (reinforcement and punishment), to target antisocial thinking, and to build prosocial thinking and strong thinking skills through practice. These practices were intended to occur in classroom settings as well as outside of the classroom in the daily lives of program participants. Practice sessions were supposed to increase in difficulty once skills at one level of difficulty were mastered (see Andrews & Bonta, 1998). For offenders, studies found prosocial modeling and contingencies to be more effective than the antisocial modeling and reinforcement/punishment (Andrews, 1980) so endemic to correctional environments (Buehler, Patterson, & Furniss, 1966). The effectiveness of the cognitive behavioral and social learning approaches was found to be enhanced by high quality relationships (Andrews & Bonta, 1998) which emphasized appropriate boundaries, enthusiasm, genuine respect toward clients, honesty, openness, flexibility, and warmth. In fact, Andrews (1980) reported that prosocial modeling and instruction often did not work when client-facilitator relationships were poorly developed.

Finally, the emergence of cognitive behavioral and social learning approaches differed from many other types of interventions, because the new program models offered detailed program manuals to personnel. These became essential to programmatic efforts to assure program consistency across settings, and to defining standards of program quality (Gendreau & Ross, 1987). In addition to detailing the importance of underlying theoretical components of modeling, contingencies, practice, and classroom rapport, the manuals spoke to issues of classroom control and provided scripted lessons, and instructional tools (such as case studies, problem scenarios and visuals) (e.g., see Novaco, 1975; Goldstein & Glick, 1987; Ross, Fabiano, & Ross, 1989).

For the present study, relevant psychological theories and the program manual for R&R (Ross, Fabiano, & Ross, 1989) served as the basis for operationalizing the intended features of R&R and for measuring the facilitators' adherence to them. In Phase II, we also examined factors typical to most process

evaluations, such as group size, program setting, and completion rates.

Against this background, the present study moves beyond the Phase I evaluation, to assist the Georgia Board of Pardons and Paroles, and the research community, in the identification of the most optimal program conditions for the Georgia Cognitive Skills Program. In identifying the most and least successful participants, and the most and least successful types of classrooms, we hope to inform plans for future programs and to assist the State's efforts to administer programs of the highest quality and effectiveness.

Description of the Georgia Cognitive Skills Program

Reasoning and Rehabilitation (Ross & Fabiano, 1985) was initially developed and piloted for the Correctional Services of Canada in the mid 1980s but has since been implemented in several sites throughout the United States as well as in Spain and Great Britain. GBPP initially adopted the program in 1992. The Georgia Cognitive Skills Program (R&R) seeks to modify offenders' impulsive, egocentric, illogical and rigid thinking patterns. In doing so, it shares assumptions common to the larger field of cognitive-behavioral psychology, currently considered to be the dominant therapeutic paradigm in mental health (Dobson & Khatri, 2000). Cognitive psychology assumes that dysfunctional thought (cognitive) processes can lead to maladaptive behavior. Therefore, teaching offenders to recognize and change criminogenic thinking (e.g., Yochelson & Samenow, 1976) and to develop more mature thinking skills (Ross & Fabiano, 1985) should reduce their propensity to crime. Specific objectives of the R&R program include improving offenders' interpersonal problem-solving, consequential thinking, means-end reasoning, social perspective-taking, critical and abstract reasoning, and creative thinking (Ross, Fabiano, & Ross 1989). Reasoning and Rehabilitation focuses more upon the processes involved in thinking as opposed to the content of thinking (or *what* offenders actually think).

The program is structured into 35 lessons that cover seven key components: (a) problem-solving, (b) creative thinking, (c) social skills, management of emotions, (d) negotiation skills, (e) values enhancement, and (f) critical reasoning. Each component is broken into sub-skills. For example, to improve problem-solving skills, participants engage in activities that target specific skills such as gathering

information, conceptualizing dilemmas, alternative thinking, and assertive communication. The class sessions build on each other in such a way that new skills are presented along with opportunities to practice previously introduced skills. Activities include role-playing, thinking games, homework assignments, and group discussions.

The facilitators' handbook (Ross, Fabiano, & Ross 1989), in addition to intensive training of facilitators (coaches), attempted to foster consistency across program sites. The manual provides detailed lesson plans and techniques for running the program sessions. Included in these lessons are suggestions on how coaches may verbalize the material to the program participants, possible comments and questions, and scheduled activities for learning new skills and practicing previously introduced skills. Furthermore, the program coaches are instructed to reinforce participants' efforts and accomplishments. Class interaction is encouraged. The curriculum is designed to cultivate an informal yet structured group atmosphere, with the participants engaging in debate-like dialogues.

Phase II of the Georgia Cognitive Skills Experiment took place from July, 1998 to April, 2000 in 28 sites throughout the State of Georgia; 25 sites were parole districts and three were correctional institutions where the program was administered to pre-release inmates. Subsequent sections of this report provide further description of the Cognitive Skills Program. For additional accounts, readers are referred to the Process Evaluation for Phase II (Van Voorhis, Spuance, Johnson-Listwan, Pealer, 2001).

Organization of the Report

The remainder of the report is organized sequentially to address the research questions posed on page 3. Following an accounting of the study methodology in Chapter 2, Chapter 3 presents research outcomes for experimental versus control group participants. The chapter details group comparisons on changes in pretest and posttest scores, returns to prison, rearrests, technical violations, and employment. Chapter 4 addresses the issue of specific responsivity, or whether certain types of offenders were more successful than others. Chapter 5 focuses on the importance of program characteristics. We conclude with a summary and discussion of findings and program recommendations. We turn now to a discussion of the methodology of the study.

Chapter 2

Methodology

Research Design

From July 1998 to April 2000, 28 sites across Georgia participated in Phase II, including 25 parole districts and three pre-release units housed in correctional institutions. In addition, 98 parolees attended the program in sites that did not participate in the experimental design. Those participants not included in the random assignment design are excluded from the analyses of experimental versus control subjects⁵.

Phase II of the Georgia Cognitive Skills Experiment employed an experimental design whereby 1190 offenders were randomly assigned to either the treatment ($n = 609$) or control group ($n = 581$).⁶ The experimental group was reduced to 574, because 35 members (5.7%) never began the cognitive skills program. The omission of these members did not change the composition of the experimental group. A regression of a dummy variable (experimental members who began the program vs. those who did not) on key demographic and criminal history variables found that group membership was unrelated to all of the background measures tested (e.g., prior violent convictions, prior incarcerations, educational attainment, marital status, age, race, sex, employment history, SES, IQ, risk level, and personality type).

Parole and corrections officers at locations participating in the experimental design compiled lists of eligible participants based on the following criteria: (a) high-risk for re-offending, (b) IQ scores (Culture Fair) (Cattell & Cattell, 1973) of at least 80, (c) no history of sexual offending, and (d) absence of substance abuse problems so severe as to interfere with the offender's ability to attend or comprehend the program. "High risk of re-offending" was not always informed by an actuarial measure of risk because GBPP did not complete risk assessments for offenders with less serious offenses or sentences of less than two years. The study participants also were to have at least 16 months remaining on their parole term in order to ensure their having adequate time to complete the four-month program and to be monitored for an

⁵ The non-randomly assigned participants who completed the program are, however, included in the analyses of programmatic characteristics in chapter 5.

⁶ The experimental group has 28 more members than the comparison group for two reasons. First, ten district lists submitted to the parole board for random assignment had one more person in the experimental group than the comparison group; three lists had an extra person in the comparison group. Second, six study groups had larger experimental than comparison groups in an effort to create suitable class sizes. For instance, a parole district might have 8 of 12 offenders randomly assigned to the experimental group in order to obtain a large enough class and to allow for the likelihood of individuals dropping out of the program.

additional 12 months thereafter. GBPP staff sent the lists of eligible offenders to the evaluation unit in Atlanta for random assignment to either the experimental or the comparison group.

Each of the 28 sites operated at least one Cognitive Skills course. Nine sites conducted a total of two courses, and an additional six ran three groups. One location, Whitworth Detention Center, held four groups. Hence, the Phase II study comprised a total of 52 groups of randomly assigned offenders. Table 1 presents the number of Cognitive Skills courses operated at each site and the number of experimental and control group members for each group. The table also indicates the sites that ran groups that were not part of the experimental design.

Table 1: Distribution of Phase II study participants across districts and study groups.

Location	Group	Experimental		Control		Total		Class with no Comparison Group	
		N	%	N	%	N	%	N	%
<u>Parole Districts:</u>									
Marietta	Group 1	20	52.6	18	47.4	38	100.0		
	Group 2	11	35.5	20	64.5	31	100.0		
	Group 3	25	48.1	27	51.9	52	100.0		
Clarkesville	Group 1	8	50.0	8	50.0	16	100.0		
Dahlonega	Group 1	12	52.2	11	47.8	23	100.0		
	NCG#	----	----	----	----	----	----	11	100.0
Milledgeville	Group 1	8	44.4	10	55.6	18	100.0		
	Group 2	10	50.0	10	50.0	20	100.0		
Thomas	Group 1	9	47.4	10	52.6	19	100.0		
Savannah	Group 1	12	52.2	11	47.8	23	100.0		
Jesup	Group 1	7	58.3	5	41.7	12	100.0		
Waycross	Group 1	10	50.0	10	50.0	20	100.0		
	Group 2	7	46.7	8	53.3	15	100.0		
Fitzgerald	NCG	----	----	----	----	----	----	9	100.0
Columbus	Group 1	15	48.4	16	51.6	31	100.0		
	NCG	----	----	----	----	----	----	7	100.0
Jonesboro	Group 1	9	40.9	13	59.1	22	100.0		
	Group 2	9	50.0	9	50.0	18	100.0		
Lawrenceville	Group 1	9	50.0	9	50.0	18	100.0		
	Group 2	7	50.0	7	50.0	14	100.0		
Augusta	Group 1	7	36.8	12	63.2	19	100.0		
	Group 2	12	50.0	12	50.0	24	100.0		
	Group 3	9	45.0	11	55.0	20	100.0		
Brunswick	Group 1	10	50.0	10	50.0	20	100.0		
	Group 2	9	52.9	8	41.1	17	100.0		
	Group 3	10	50.0	10	50.0	20	100.0		
Cairo	Group 1	13	50.0	13	50.0	26	100.0		
North Fulton	Group 1	9	42.9	12	57.1	21	100.0		
	Group 2	16	50.0	16	50.0	32	100.0		
	Group 3	18	41.9	25	58.1	43	100.0		

NCG indicates classes with no comparison groups.

Table 1 continued.

Location	Group	Experimental		Control		Total		Class with no Comparison Group	
		N	%	N	%	N	%	N	%
LaGrange	Group 1	12	60.0	8	40.0	20	100.0		
	NCG	----	----	----	----	----	----	9	100.0
	NCG	----	----	----	----	----	----	8	100.0
Griffin	Group 1	12	50.0	12	50.0	24	100.0		
LaFayette	Group 1	6	50.0	6	50.0	12	100.0		
	Group 2	7	53.8	6	46.2	13	100.0		
Jefferson	NCG	----	----	----	----	----	----	11	100.0
Jefferson	Group 1	11	47.8	12	52.2	23	100.0		
South Dekalb	Group 1	15	50.0	15	50.0	30	100.0		
	Group 2	17	43.6	22	56.4	39	100.0		
Lyons	Group 1	9	50.0	9	50.0	18	100.0		
	NCG	----	----	----	----	----	----	12	100.0
	NCG	----	----	----	----	----	----	10	100.0
Statesboro	Group 1	7	46.7	8	53.3	15	100.0		
	Group 2	9	50.0	9	50.0	18	100.0		
	Group 3	7	46.7	8	53.3	15	100.0		
South Metro	Group 1	9	50.0	9	50.0	18	100.0		
	Group 2	12	46.2	14	53.8	26	100.0		
South Richmond	Group 1	9	42.9	12	57.1	21	100.0		
	Group 2	9	50.0	9	50.0	18	100.0		
Carrollton	NCG	----	----	----	----	----	----	8	100.0
Douglas	NCG	----	----	----	----	----	----	9	100.0
Hartwell	Group 1	9	90.0	1	10.0	10	100.0		
	Group 2	9	81.8	2	18.2	11	100.0		
<u>Institutions:</u>									
Coastal St. Prison	Group 1	12	52.2	11	47.8	23	100.0		
Whitworth Det. Cent.	Group 1	10	52.6	9	47.4	19	100.0		
	Group 2	10	83.3	2	16.7	12	100.0		
	Group 3	10	50.0	10	50.0	20	100.0		
	Group 4	10	50.0	10	50.0	20	100.0		
Metro St. Prison	Group 1	8	50.0	8	50.0	16	100.0		
	Group 2	17	53.1	15	46.9	32	100.0		
	<u>Group 3</u>	<u>27</u>	<u>54.0</u>	<u>23</u>	<u>46.0</u>	<u>50</u>	<u>100.0</u>		
Total		574	49.7	581	50.3	1155	100.0	94	100.0

Sample

Parole Participants

All analyses were conducted separately for the parole and pre-release groups. Table 2 shows demographic and social history characteristics at prison intake for the parolee sample as a whole and separated according to group (e/c) assignment. The parole participants were predominantly African American. Their average age was 31.9 with the largest (modal) age group (43.9%) between the ages of 26 and 36. All participants were male. At the time of their arrest, 54.0 percent of the participants were employed on a full-time bases. At prison intake, 41.1% were rated as middle class;⁷ 63.9% were married.

The majority of the participants (66.5%) had less than a high school education. Notwithstanding GBPP directives to exclude parolees with low IQ scores, 6.2 percents of the sample scored at 80 or below on the Culture Fair (Cattell & Cattell, 1973) assessment. Even so, 70.9% of the sample read at the fifth grade reading level or higher. Jesness Inventory results found 61.4% of the parolees were diagnosed at I-levels 2 or 3, indicating that cognitive and interpersonal maturity was atypically low even for a criminal population. JI personality diagnoses observed: (a) 31.7% aggressive (antisocial values, internalized criminal lifestyle), (b) 12.7 percent neurotic (high anxiety), (c) 29.7% dependent (easily led, immature), and (d) 25.9% situational (prosocial value system; crime is the result of adversity, poor coping skills, and perhaps substance abuse).

⁷ No established index of social class was used for this purpose. Prison intake personnel, however, were guided by clear criteria for social class ratings. These were as follows: a) welfare (receiving some form of public assistance at the time of incarceration, regardless of other income), b) occasionally employed (occasionally employed), c) minimum standard (annual income meets the government's Minimum Standard of Living for a family of 4, d) middle class (making more than the Minimum Standard of Living and having some resources, such as property, savings or investments).

Table 2: Comparison of the experimental and control group parolees across background social and demographic characteristics⁸ (collected at prison intake).

Characteristic		Experimental		Control		Total	
		N	%	N	%	N	%
Age:	18-25	129	28.1	123	25.5	252	26.8
	26-36	195	42.5	219	45.3	414	43.9
	36-50	120	26.1	123	25.5	243	25.8
	50+	<u>15</u>	<u>3.3</u>	<u>18</u>	<u>3.7</u>	<u>33</u>	<u>3.5</u>
	Total	459	100.0	483	100.0	942	100.0
Average Age:		31.9		31.9		31.9	
Race:	White	141	30.7	140	29.0	281	29.8
	<u>Nonwhite</u>	<u>318</u>	<u>69.3</u>	<u>343</u>	<u>71.0</u>	<u>661</u>	<u>70.2</u>
	Total	459	100.0	483	100.0	942	100.0
Sex:	Male	470	100.0	493	100.0	963	100.0
	<u>Female</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>
	Total	470	100.0	493	100.0	963	100.0
Employment status:							
	Employed full time	234	51.3	270	56.6	504	54.0
	Employed part time	51	11.2	48	10.1	99	10.6
	Unemployed < 6 mo.	65	14.3	52	10.9	117	12.5
	Unemployed for 6+ mo.	69	15.1	73	15.3	142	15.2
	Never worked but capable	14	3.1	11	2.3	25	2.7
	Student	3	0.7	1	0.2	4	0.4
	Incapable of work	10	2.2	9	1.9	19	2.0
	<u>Other</u>	<u>10</u>	<u>2.2</u>	<u>13</u>	<u>2.7</u>	<u>23</u>	<u>2.5</u>
	Total	456	100.1	477	100.0	933	99.9
Social class:							
	Welfare	36	8.0	34	7.3	70	7.6
	Occasionally employed	58	12.9	43	9.2	101	11.0
	Minimum standard	171	38.0	191	40.9	362	39.5
	Middle class	183	40.7	194	41.5	377	41.1
	<u>Other</u>	<u>2</u>	<u>0.4</u>	<u>5</u>	<u>1.1</u>	<u>7</u>	<u>0.8</u>
	Total	450	100.0	467	100.0	917	100.0
Education:							
	Less than high school	38	8.3	55	11.4	93	9.9
	Some high school	264	57.9	267	55.4	531	56.6
	High school	104	22.8	118	24.5	222	23.7
	At least some technical school	15	3.3	12	2.5	27	2.9
	At least some college	34	7.5	30	6.2	64	6.8
	<u>Other</u>	<u>1</u>	<u>0.2</u>	<u>0</u>	<u>0.0</u>	<u>1</u>	<u>0.1</u>
	Total	456	100.0	482	100.0	938	100.0
Marital status:							
	Single (never married)	276	60.7	322	66.9	598	63.9
	Married	48	10.5	52	10.8	100	10.7
	Separated	27	5.9	20	4.2	47	5.0
	Divorced (not remarried)	46	10.1	31	6.4	77	8.2
	Widowed	3	0.7	3	0.6	6	0.6
	Common law marriage	54	11.9	53	11.0	107	11.4
	<u>Other</u>	<u>1</u>	<u>0.2</u>	<u>0</u>	<u>0.0</u>	<u>1</u>	<u>0.1</u>
	Total	455	100.0	481	99.9	936	99.9

⁸ None of the differences between the experimental and control groups are statistically significant.

Table 2 continued.

Characteristic	Experimental		Control		Total	
	N	%	N	%	N	%
Substance abuse:						
History of substance abuse	250	54.5	267	55.3	517	54.9
<u>No history of substance abuse</u>	<u>209</u>	<u>45.5</u>	<u>216</u>	<u>44.7</u>	<u>425</u>	<u>45.1</u>
Total	459	100.0	483	100.0	942	100.0
IQ:						
50 to 80	28	6.3	29	6.1	57	6.2
81 to 99	153	34.2	177	37.2	330	35.7
100 to 115	198	44.2	194	40.8	392	42.4
<u>116 to 139</u>	<u>69</u>	<u>15.4</u>	<u>76</u>	<u>16.0</u>	<u>145</u>	<u>15.7</u>
Total	448	100.1	476	100.1	924	100.0
Average IQ:	101.8		101.7		101.8	
Reading level:						
Below 5 th grade	126	28.3	142	29.8	268	29.1
<u>5th grade or above</u>	<u>320</u>	<u>71.7</u>	<u>334</u>	<u>70.2</u>	<u>654</u>	<u>70.9</u>
Total	446	100.0	476	100.0	922	100.0
I-Level:						
I-Levels 2 and 3	252	61.9	257	60.9	509	61.4
<u>I-Level 4</u>	<u>155</u>	<u>38.1</u>	<u>165</u>	<u>39.1</u>	<u>320</u>	<u>38.6</u>
Total	407	100.0	422	100.0	829	100.0
Collapsed Jesness Inventory type:						
Aggressives	134	32.9	129	30.6	263	31.7
Neurotics	53	13.0	52	12.3	105	12.7
Dependents	118	29.0	128	30.3	246	29.7
<u>Situationals</u>	<u>102</u>	<u>25.1</u>	<u>113</u>	<u>26.8</u>	<u>215</u>	<u>25.9</u>
Total	407	100.0	422	100.0	829	100.0

Table 3 shows the criminal history characteristics of the parolees. Most participants (78.8%) scored at medium to high risk on a risk scale developed for this study (see page 42, below). All but two participants had a prior felony on record and 56.9% of the sample had served at least one prior prison term. In keeping with their serious prior histories, prior convictions included: (a) violent offenses (44.8%), (b) drug sales (45.5%), and (c) drug possession (64.5%). Prior convictions for at least one sex offense characterized 32 participants (3.4%), even though screening criteria attempted to exclude such individuals.

Table 3: Comparison of the experimental and control group parolees across criminal history characteristics⁹ (collected at prison intake).

Characteristic		Experimental		Control		Total	
		N	%	N	%	N	%
Risk:	Low	104	22.8	94	19.7	198	21.2
	<u>Medium/High</u>	<u>352</u>	<u>77.2</u>	<u>383</u>	<u>80.3</u>	<u>735</u>	<u>78.8</u>
	Total	456	100.0	477	100.0	933	100.0
Number of prior incarcerations:							
	None	198	43.1	193	40.0	391	41.5
	One	128	27.9	137	28.4	265	28.1
	Two	65	14.2	61	12.6	126	13.4
	Three	28	6.1	43	8.9	71	7.5
	Four	22	4.8	24	5.0	46	4.9
	<u>Five or more</u>	<u>18</u>	<u>3.9</u>	<u>25</u>	<u>5.2</u>	<u>43</u>	<u>4.6</u>
	Total	459	100.0	483	100.1	942	100.0
	Mean	1.8		2.0		1.9	
	Median	1.0		1.0		1.0	
Number of prior felony convictions:							
	None	1	0.2	1	0.2	2	0.2
	One	117	25.5	117	24.2	234	24.8
	Two	126	27.5	129	26.7	255	27.1
	Three	79	17.2	101	20.9	180	19.1
	Four	55	12.0	56	11.6	111	11.8
	<u>Five or more</u>	<u>81</u>	<u>17.6</u>	<u>79</u>	<u>16.4</u>	<u>160</u>	<u>17.0</u>
	Total	459	100.0	483	100.0	942	100.0
	Mean	2.9		2.9		2.9	
	Median	2.0		2.0		2.0	
Number of prior violent convictions:							
	None	260	56.6	269	55.7	529	56.2
	One	72	15.7	71	14.7	143	15.2
	Two	42	9.2	56	11.6	98	10.4
	Three	28	6.1	29	6.0	57	6.1
	Four	19	4.1	16	3.3	35	3.7
	<u>Five or more</u>	<u>38</u>	<u>8.3</u>	<u>42</u>	<u>8.7</u>	<u>80</u>	<u>8.5</u>
	Total	459	100.0	483	100.0	942	100.1
	Mean	1.3		1.3		1.3	

⁹ None of the differences between the experimental and control groups are statistically significant.

Table 3 continued.

Characteristic	Experimental		Control		Total	
	N	%	N	%	N	%
Number of prior property convictions:						
None	137	29.8	128	26.5	265	28.1
One	69	15.0	72	14.9	141	15.0
Two	57	12.4	40	8.3	97	10.3
Three	28	6.1	52	10.8	80	8.5
Four	20	4.4	31	6.4	51	5.4
Five	21	4.6	24	5.0	45	4.8
Six	20	4.4	22	4.6	42	4.5
Seven	15	3.3	17	3.5	32	3.4
<u>Eight or more</u>	<u>92</u>	<u>20.0</u>	<u>97</u>	<u>20.1</u>	<u>189</u>	<u>20.1</u>
Total	459	100.0	483	100.1	942	100.1
	Mean	4.4	4.7		4.5	
	Median	2.0	3.0		2.0	
Number of prior drug sale convictions:						
None	252	54.9	271	56.1	523	55.5
One	111	24.2	109	22.6	220	23.4
<u>Two or more</u>	<u>96</u>	<u>20.9</u>	<u>103</u>	<u>21.3</u>	<u>199</u>	<u>21.1</u>
Total	459	100.0	483	100.0	942	100.0
	Mean	0.8	0.8		0.8	
	Median	0.0	0.0		0.0	
Number of prior drug possession convictions:						
None	166	36.2	165	34.2	331	35.1
One	74	16.1	90	18.6	164	17.4
Two	64	13.9	58	12.0	122	13.0
Three	40	8.7	55	11.4	95	10.1
Four	35	7.6	38	7.9	73	7.7
<u>Five or more</u>	<u>80</u>	<u>17.4</u>	<u>77</u>	<u>15.9</u>	<u>157</u>	<u>16.7</u>
Total	459	99.9	483	100.0	942	100.0
	Mean	2.4	2.3		2.3	
	Median	1.0	1.0		1.0	
Number of prior sex crime convictions:						
None	447	97.4	463	95.9	910	96.6
One	10	2.2	10	2.1	20	2.1
<u>Two or more</u>	<u>2</u>	<u>0.4</u>	<u>10</u>	<u>2.1</u>	<u>12</u>	<u>1.3</u>
Total	459	100.0	483	100.1	942	100.0
	Mean	0.0	0.1		0.0	
	Median	0.0	0.0		0.0	

Tables 2 and 3, report the experimental and comparison groups to be nearly identical on social, demographic and criminal history factors. As with Phase I of this study, GBPP personnel appeared to have adhered consistently to the directives of random assignment.

Pre-release Participants

Social and demographic characteristics for the institutionalized (pre -release) participants are shown in Table 4. This sample consisted of 192 offenders, approximately half of whom (51.0%) were women. Background characteristics differ for males and females; therefore we discuss the groups separately.

As shown in Table 4, the majority of the women in the pre-release groups were non-white (59.2%), single (56.1%), unemployed at the time of their arrest (66.4%) and living at a minimum standard of living or below (66.0%), with histories of substance abuse (65.3%). Most (63.3%) had less than a high school education. IQ scores of less than 80 were reported for 23 offenders (24.0%), and a similar proportion obtained reading scores below the fifth grade. The average age for this sample was 34.1 years. I-level diagnosis found higher levels of interpersonal maturity than observed for the male parole and pre-release samples. The women's sample also received more neurotic classifications and less aggressive or situational ones.

The average age of the pre-release males was 31.9 and most were non-white (64.9%). In contrast to the women's sample, the majority of male pre-release participants (53.9%) were employed full-time at their arrest. However, only 31 (34.4%) had a high school education, and 55 (60.5%) lived at the minimum standard of living or below. Only 5 pre -release, male offenders (5.4%) had IQ scores below 80, and the majority (89.1%) read at the fifth grade reading level or above. I-Level classifications were similar to those for the parole sample; however, pre-release males were observed to have more aggressive (43.7%) and neurotic (23.9%) classifications and less situational (12.7%) and dependent ones (19.7%).

Table 4: Comparison of the experimental and control group pre-release study members across sex and background social and demographic characteristics¹⁰ (collected at prison intake).

Characteristic		Experimental		Control		Total	
		N	%	N	%	N	%
Women Pre-Release Groups							
Age:	18-25	10	19.2	9	19.6	19	19.4
	26-36	22	42.3	20	43.5	42	42.9
	36-50	16	30.8	16	34.8	32	32.7
	<u>50+</u>	<u>4</u>	<u>7.7</u>	<u>1</u>	<u>2.2</u>	<u>5</u>	<u>5.1</u>
	Total	52	100.0	46	100.1	98	100.1
Average Age:		34.4		33.9		34.1	
Race:	White	21	40.4	19	41.3	40	40.8
	<u>Nonwhite</u>	<u>31</u>	<u>59.6</u>	<u>27</u>	<u>58.7</u>	<u>58</u>	<u>59.2</u>
	Total	52	100.0	46	100.0	98	100.0
Sex:	Male	0	0.0	0	0.0	0	0.0
	<u>Female</u>	<u>52</u>	<u>100.0</u>	<u>46</u>	<u>100.0</u>	<u>98</u>	<u>100.0</u>
	Total	52	100.0	46	100.0	98	100.0
Employment status:							
	Employed full time	14	27.5	13	31.7	27	29.3
	Employed part time	3	5.9	1	2.4	4	4.3
	Unemployed < 6 mo.	7	13.7	7	17.1	14	15.2
	Unemployed for 6+ mo.	14	27.5	11	26.8	25	27.2
	Never worked but capable	5	9.8	4	9.8	9	9.8
	Student	0	0.0	0	0.0	0	0.0
	Incapable of work	4	7.8	1	2.4	5	5.4
	<u>Other</u>	<u>4</u>	<u>7.8</u>	<u>4</u>	<u>9.8</u>	<u>8</u>	<u>8.7</u>
	Total	51	100.0	41	100.0	92	99.9
Social class:							
	Welfare	9	17.3	10	22.2	19	19.6
	Occasionally employed	0	0.0	0	0.0	0	0.0
	Minimum standard	27	51.9	18	40.0	45	46.4
	Middle class	15	28.8	17	37.8	32	33.0
	<u>Other</u>	<u>1</u>	<u>1.9</u>	<u>0</u>	<u>0.0</u>	<u>1</u>	<u>1.0</u>
	Total	52	99.9	45	100.0	97	100.0
Education:							
	Less than high school	6	11.5	8	17.4	14	14.3
	Some high school	28	53.8	20	43.5	48	49.0
	High school	10	19.2	9	19.6	19	19.4
	At least some technical school	3	5.8	2	4.3	5	5.1
	At least some college	5	9.6	7	15.2	12	12.2
	<u>Other</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>
	Total	52	99.9	46	100.0	98	100.0

¹⁰ None of the differences between the experimental and control groups are statistically significant.

Table 4 continued.

Characteristic	Experimental		Control		Total	
	N	%	N	%	N	%
Women Pre-Release Groups continued						
Marital status:						
Single (never married)	31	59.6	24	52.2	55	56.1
Married	6	11.5	8	17.4	14	14.3
Separated	3	5.8	4	8.7	7	7.1
Divorced (not remarried)	6	11.5	7	15.2	13	13.3
Widowed	3	5.8	3	6.5	6	6.1
Common law marriage	3	5.8	0	0.0	3	3.1
<u>Other</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>
Total	52	100.0	46	100.0	98	100.0
Substance abuse:						
History of substance abuse	36	69.2	28	60.9	64	65.3
<u>No history of substance abuse</u>	<u>16</u>	<u>30.8</u>	<u>18</u>	<u>39.1</u>	<u>34</u>	<u>34.7</u>
Total	52	100.0	46	100.0	98	100.0
IQ:						
50 to 80	14	26.9	9	20.5	23	24.0
81 to 99	25	48.1	19	43.2	44	45.8
100 to 115	8	15.4	10	22.7	19	18.8
<u>116 to 139</u>	<u>5</u>	<u>9.6</u>	<u>6</u>	<u>13.6</u>	<u>11</u>	<u>11.5</u>
Total	52	100.0	44	100.0	97	100.1
Average IQ:						
	89.0		93.1		90.8	
Reading level:						
Below 5 th grade	12	23.1	11	24.4	23	23.7
<u>5th grade or above</u>	<u>40</u>	<u>76.9</u>	<u>34</u>	<u>75.6</u>	<u>74</u>	<u>76.3</u>
Total	52	100.0	45	100.0	97	100.0
I-Level:						
I-Level s 2 and 3	21	44.7	25	62.5	46	52.9
<u>I-Level 4</u>	<u>26</u>	<u>55.3</u>	<u>15</u>	<u>37.5</u>	<u>41</u>	<u>47.1</u>
Total	47	100.0	40	100.0	87	100.0
Collapsed Jesness Inventory type:						
Aggressives	9	19.1	11	27.5	20	23.0
Neurotics	22	46.8	9	22.5	31	35.6
Dependents	12	25.5	14	35.0	26	29.9
<u>Situationals</u>	<u>4</u>	<u>8.5</u>	<u>6</u>	<u>15.0</u>	<u>10</u>	<u>11.5</u>
Total	47	99.9	40	100.0	87	100.0

Table 4 continued.

Characteristic		Experimental		Control		Total	
		N	%	N	%	N	%
Men Pre-Release Groups							
Age:	18-25	13	25.0	10	23.8	23	24.5
	26-36	20	38.5	18	42.9	38	40.4
	36-50	17	32.7	14	33.3	31	33.0
	<u>50+</u>	<u>2</u>	<u>3.8</u>	<u>0</u>	<u>0.0</u>	<u>2</u>	<u>2.1</u>
	Total	52	100.0	42	100.0	94	100.0
Average Age:		32.0		31.8		31.9	
Race:	White	20	38.5	13	31.0	33	35.1
	<u>Nonwhite</u>	<u>32</u>	<u>61.5</u>	<u>29</u>	<u>69.0</u>	<u>61</u>	<u>64.9</u>
	Total	52	100.0	42	100.0	94	100.0
Sex:	Male	52	100.0	42	100.0	94	100.0
	<u>Female</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>
	Total	52	100.0	42	100.0	94	100.0
Employment status:							
	Employed full time	25	52.1	23	56.1	48	53.9
	Employed part time	1	2.1	3	7.3	4	4.5
	Unemployed < 6 mo.	13	27.1	5	12.2	18	20.2
	Unemployed for 6+ mo.	6	12.5	4	9.8	10	11.2
	Never worked but capable	1	2.1	3	7.3	4	4.5
	Student	0	0.0	0	0.0	0	0.0
	Incapable of work	0	0.0	1	2.4	1	1.1
	<u>Other</u>	<u>2</u>	<u>4.2</u>	<u>2</u>	<u>4.9</u>	<u>4</u>	<u>4.5</u>
	Total	48	100.1	41	100.0	89	99.9
Social class:							
	Welfare	3	5.9	2	5.0	5	5.5
	Occasionally employed	3	5.9	6	15.0	9	9.9
	Minimum standard	23	45.1	18	45.0	41	45.1
	Middle class	22	43.1	14	35.0	36	39.6
	<u>Other</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>
	Total	51	100.0	40	100.0	91	100.1
Education:							
	Less than high school	4	7.8	7	16.7	11	11.8
	Some high school	29	56.9	21	50.0	50	53.8
	High school	11	21.6	12	28.6	23	24.7
	At least some technical school	0	0.0	0	0.0	0	0.0
	At least some college	7	13.7	2	4.8	9	9.7
	<u>Other</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>
	Total	51	100.0	42	100.1	93	100.0

Table 4 continued.

Characteristic	Experimental		Control		Total	
	N	%	N	%	N	%
Men Pre-Release Groups continued						
Marital status:						
Single (never married)	31	59.6	27	64.3	58	61.7
Married	4	7.7	5	11.9	9	9.6
Separated	2	3.8	3	7.1	5	5.3
Divorced (not remarried)	11	21.2	5	11.9	16	17.0
Widowed	0	0.0	1	2.4	1	1.1
Common law marriage	4	7.7	1	2.4	5	5.3
<u>Other</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>	<u>0</u>	<u>0.0</u>
Total	52	100.0	42	100.0	94	100.0
Substance abuse:						
History of substance abuse	37	71.2	25	59.5	62	66.0
<u>No history of substance abuse</u>	<u>15</u>	<u>28.8</u>	<u>17</u>	<u>40.5</u>	<u>32</u>	<u>34.0</u>
Total	52	100.0	42	100.0	94	100.0
IQ:						
50 to 80	3	5.8	2	5.0	5	5.4
81 to 99	12	23.1	15	37.5	27	29.3
100 to 115	34	65.4	20	50.0	54	58.7
<u>116 to 139</u>	<u>3</u>	<u>5.8</u>	<u>3</u>	<u>7.5</u>	<u>6</u>	<u>6.5</u>
Total	52	100.1	40	100.0	92	99.9
Average IQ:	103.0		100.0		102.4	
Reading level:						
Below 5 th grade	6	11.5	4	10.0	10	10.9
<u>5th grade or above</u>	<u>46</u>	<u>88.5</u>	<u>36</u>	<u>90.0</u>	<u>82</u>	<u>89.1</u>
Total	52	100.0	40	100.0	92	100.0
I-Level:*						
I-Levels 2 and 3	25	54.3	20	80.0	45	63.4
<u>I-Level 4</u>	<u>21</u>	<u>45.7</u>	<u>5</u>	<u>20.0</u>	<u>26</u>	<u>36.6</u>
Total	46	100.0	25	100.0	71	100.0
Collapsed Jesness Inventory type:*						
Aggressives	17	37.0	14	56.0	31	43.7
Neurotics	15	32.6	2	8.0	17	23.9
Dependents	8	17.4	6	24.0	14	19.7
<u>Situationals</u>	<u>6</u>	<u>13.0</u>	<u>3</u>	<u>12.0</u>	<u>9</u>	<u>12.7</u>
Total	46	100.0	25	100.0	71	100.0

Male and female experimental groups and comparison groups evidenced similar demographic characteristics. As can be seen in Table 5, however, this was not the case with respect to their criminal histories. Experimental participants were more likely to have serious criminal histories than the comparison group members. For the women's sample, these differences were more pronounced than for the men's sample. For example, 38 (74.5 percent) members of the experiment group were classified as medium to high risk as compared to only 24 (58.5%) of the comparison group. The experimental group evidenced significantly more prior felonies (mean = 3.1) than the comparison group (mean = 2.2). Male experimental offenders also were more likely than comparisons to have: (a) classifications as medium/high risk, (b) prior incarcerations, and (c) prior violent convictions. None of these differences, however, were statistically significant.

These group differences suggest that the experimental design may have been compromised for both the male and female pre-release samples. Another possibility, however, is that the proportions and means characterizing the pre-release sample are not stable due to the small sample size. If that is the case, the addition of more cases would likely make the experimental and control groups more similar. Although our data analyses controlled for offense history characteristics, it was not possible to utilize many control variables without overly compromising the power of the statistical tests.

Table 5: Comparison of the experimental and control group pre-release study members across sex and criminal history characteristics (collected at prison intake).

Characteristic		Experimental		Control		Total	
		N	%	N	%	N	%
Women Pre-Release Groups							
Risk	Low	13	25.5	17	41.5	30	32.6
	<u>Medium/High</u>	<u>38</u>	<u>74.5</u>	<u>24</u>	<u>58.5</u>	<u>62</u>	<u>67.4</u>
	Total	51	100.0	41	100.0	92	100.0
Number of prior incarcerations:							
	None	23	44.2	30	65.2	53	54.1
	One	9	17.3	7	15.2	16	16.3
	Two	10	19.2	5	10.9	15	15.3
	Three	6	11.5	3	6.5	9	9.2
	Four	2	3.8	1	2.2	3	3.1
	<u>Five or more</u>	<u>2</u>	<u>3.8</u>	<u>0</u>	<u>0.0</u>	<u>2</u>	<u>2.0</u>
	Total	52	99.8	46	100.0	98	100.0
	Mean	1.7		1.3		1.5	
	Median	1.0		1.0		1.0	
Number of prior felony convictions:							
	None	0	0.0	1	2.2	1	1.0
	One	14	26.9	21	45.7	35	35.7
	Two	11	21.2	6	13.0	17	17.3
	Three	12	23.1	9	19.6	21	21.4
	Four	5	9.6	3	6.5	8	8.2
	<u>Five or more</u>	<u>10</u>	<u>19.2</u>	<u>6</u>	<u>13.0</u>	<u>16</u>	<u>16.3</u>
	Total	52	100.0	46	100.0	98	99.9
	Mean*	3.1		2.2		2.7	
	Median	3.0		2.0		2.0	
Number of prior violent convictions:							
	None	24	46.2	27	58.7	51	52.0
	One	8	15.4	7	15.2	15	15.3
	Two	7	13.5	6	13.0	13	13.3
	Three	3	5.8	3	6.5	6	6.1
	Four	2	3.8	0	0.0	2	2.0
	<u>Five or more</u>	<u>8</u>	<u>15.4</u>	<u>3</u>	<u>6.5</u>	<u>11</u>	<u>11.2</u>
	Total	52	100.1	46	99.9	98	99.9
	Mean	1.8		1.0		1.4	
	Median	1.0		0.0		0.0	

* Significance of X^2 or $t < .05$

Table 5 continued.

Characteristic	Experimental		Control		Total	
	N	%	N	%	N	%
Women Pre-Release Groups continued						
Number of prior property convictions:						
None	9	17.3	22	47.8	31	31.6
One	3	5.8	5	10.9	8	8.2
Two	8	15.4	2	4.3	10	10.2
Three	3	5.8	0	0.0	3	3.1
Four	3	5.8	3	6.5	6	6.1
Five	4	7.7	1	2.2	5	5.1
Six	1	1.9	4	8.7	5	5.1
Seven	1	1.9	2	4.3	3	3.1
<u>Eight or more</u>	<u>20</u>	<u>38.5</u>	<u>7</u>	<u>15.2</u>	<u>27</u>	<u>27.6</u>
Total	52	100.1	46	99.9	98	100.1
	Mean*	7.0		3.6		5.4
	Median	4.5		1.0		2.5
Number of prior drug sale convictions:						
None	41	78.8	31	67.4	72	73.5
One	9	17.3	9	19.6	18	18.4
<u>Two or more</u>	<u>2</u>	<u>3.8</u>	<u>6</u>	<u>13.0</u>	<u>8</u>	<u>8.2</u>
Total	52	99.9	46	100.0	98	100.1
	Mean	0.3		0.5		0.4
	Median	0.0		0.0		0.0
Number of prior drug possession convictions:						
None	27	51.9	30	65.2	57	58.2
One	6	11.5	5	10.9	11	11.2
Two	7	13.5	4	8.7	11	11.2
Three	3	5.8	2	4.3	5	5.1
Four	5	9.6	3	6.5	8	8.2
<u>Five or more</u>	<u>4</u>	<u>7.7</u>	<u>2</u>	<u>4.3</u>	<u>6</u>	<u>6.1</u>
Total	52	100.0	46	99.9	98	100.0
	Mean	1.6		0.9		1.3
	Median	0.0		0.0		0.0
Number of prior sex crime convictions:						
None	49	94.2	42	91.3	91	92.9
One	0	0.0	2	4.3	2	2.0
<u>Two or more</u>	<u>3</u>	<u>5.8</u>	<u>2</u>	<u>4.3</u>	<u>5</u>	<u>5.1</u>
Total	52	100.0	46	99.9	98	100.0
	Mean	0.3		0.2		0.2
	Median	0.0		0.0		0.0

Table 5 continued.

Characteristic		Experimental		Control		Total	
		N	%	N	%	N	%
Men Pre-Release Groups							
Risk	Low	4	8.3	8	19.5	12	13.5
	<u>Medium/High</u>	<u>44</u>	<u>91.7</u>	<u>33</u>	<u>80.5</u>	<u>77</u>	<u>86.5</u>
	Total	48	100.0	41	100.0	89	100.0
Number of prior incarcerations:							
	None	16	30.8	16	38.1	32	34.0
	One	13	25.0	12	28.6	25	26.6
	Two	7	13.5	6	14.3	13	13.8
	Three	6	11.5	5	11.9	11	11.7
	Four	7	13.5	1	2.4	8	8.5
	<u>Five or more</u>	<u>3</u>	<u>5.8</u>	<u>2</u>	<u>4.8</u>	<u>5</u>	<u>5.3</u>
	Total	52	100.1	42	100.1	94	99.9
	Mean	1.7		1.6		1.6	
	Median	1.5		1.0		1.0	
Number of prior felony convictions:							
	None	0	0.0	0	0.0	0	0.0
	One	11	21.2	7	16.7	18	19.1
	Two	10	19.2	9	21.4	19	20.2
	Three	13	25.0	12	28.6	25	26.6
	Four	8	15.4	10	23.8	18	19.1
	<u>Five or more</u>	<u>10</u>	<u>19.2</u>	<u>4</u>	<u>9.5</u>	<u>14</u>	<u>14.9</u>
	Total	52	100.0	42	100.0	94	99.9
	Mean*	3.1		3.0		3.0	
	Median	3.0		3.0		3.0	
Number of prior violent convictions:							
	None	32	61.5	23	54.8	55	58.5
	One	11	21.2	11	26.2	22	23.4
	Two	6	11.5	3	7.1	9	9.6
	Three	1	1.9	3	7.1	4	4.3
	Four	1	1.9	2	4.8	3	3.2
	<u>Five or more</u>	<u>1</u>	<u>1.9</u>	<u>0</u>	<u>0.0</u>	<u>1</u>	<u>1.1</u>
	Total	52	99.9	42	100.0	94	100.1
	Mean	0.8		0.8		0.8	
	Median	0.0		0.0		0.0	

Table 5 continued.

Characteristic	Experimental		Control		Total	
	N	%	N	%	N	%
Men Pre-Release Groups continued						
Number of prior property convictions:						
None	11	21.2	11	26.2	22	23.4
One	6	11.5	6	14.3	12	12.8
Two	8	15.4	4	9.5	12	12.8
Three	3	5.8	3	7.1	6	6.4
Four	3	5.8	6	14.3	9	9.6
Five	0	0.0	1	2.4	1	1.1
Six	3	5.8	4	9.5	7	7.4
Seven	1	1.9	2	4.8	3	3.2
<u>Eight or more</u>	<u>17</u>	<u>32.7</u>	<u>5</u>	<u>11.9</u>	<u>22</u>	<u>23.4</u>
Total	52	100.1	42	100.0	94	100.1
	Mean	6.3		3.4		5.0
	Median	3.0		2.5		3.0
Number of prior drug sale convictions:						
None	30	57.7	22	52.4	52	55.3
One	13	25.0	10	23.8	23	24.5
<u>Two or more</u>	<u>9</u>	<u>17.3</u>	<u>10</u>	<u>23.8</u>	<u>19</u>	<u>20.2</u>
Total	52	100.0	42	100.0	94	100.0
	Mean	0.6		1.0		0.8
	Median	0.0		0.0		0.0
Number of prior drug possession convictions:						
None	19	36.5	14	33.3	33	35.1
One	10	19.2	9	21.4	19	20.2
Two	5	9.6	7	16.7	12	12.8
Three	9	17.3	4	9.5	13	13.8
Four	5	9.6	2	4.8	7	7.4
<u>Five or more</u>	<u>4</u>	<u>7.7</u>	<u>6</u>	<u>14.3</u>	<u>10</u>	<u>10.6</u>
Total	52	99.9	42	100.0	94	99.9
	Mean	1.9		1.9		1.9
	Median	1.0		1.0		1.0
Number of prior sex crime convictions:						
None	50	96.2	41	97.6	91	96.8
One	2	3.8	0	0.0	2	2.1
<u>Two or more</u>	<u>0</u>	<u>0.0</u>	<u>1</u>	<u>2.4</u>	<u>1</u>	<u>1.1</u>
Total	52	100.0	42	100.0	94	100.0
	Mean	0.0		0.1		0.1
	Median	0.0		0.0		0.0

Data Collection

Multiple information sources addressed the research issues for the present study; these are shown in Figure 1. For study participants, the evaluation required data pertaining to: (a) offender demographic and social characteristics as well as criminal histories, (b) scores on risk assessments, a personality inventory, an IQ test, and pretests and posttests, and (c) follow-up data describing returns to prison, rearrests, employment and technical violations. Programmatic data included: (a) measures of group size, program completion rates, and attendance, (b) facilitators' accounts of participation levels, enthusiasm, and level of understanding, (c) participants' evaluations, and (d) observers/supervisors' evaluations. Strategies for obtaining these data, and their availability are discussed below.

Figure 1: Data collection plan, Georgia Cognitive Skills Experiment.

Data Element	Data Collection Administration	Time of Administration
<u>Intake Data</u>		
<ul style="list-style-type: none"> a. Demographic & offense history. b. Risk Assessment c. Intelligence Score (test & score) d. Intake Assessments: <ul style="list-style-type: none"> --Pride in Delinquency Scale (crim. attitudes) --Jesness Inventory (cognitive maturity & personality). 	<ul style="list-style-type: none"> a. Extracted from Georgia OTIS. b. Extracted from Georgia OTIS. c. Extracted from Georgia OTIS. d. Paper & Pencil Test, completed by the participant. 	<ul style="list-style-type: none"> a. Within one month of prison intake. b. Within one month of prison intake. c. Within one month of prison intake. d. Assessment session first week of class.
<u>Program Data</u>		
<ul style="list-style-type: none"> a. Observer's Evaluation (adherence R&R) b. Participant's Evaluation. (adherence to R&R, attendance, contracts with parole officer). c. Coach's Evaluation of Participant (attendance, level of participation and understanding). 	<ul style="list-style-type: none"> a. Observer behavioral checklist, completed by experienced observer. b. Paper & pencil evaluation, completed by participant. c. Session evaluation form, completed by the coach. 	<ul style="list-style-type: none"> a. Twice for each R&R group, session 17 & 30. b. End of program. c. Five times for each R&R group, same sessions for each group.
<u>Post-Program Assessments</u>		
<ul style="list-style-type: none"> a. Colorado Attitude Survey and Pride in Delinquency Scale. 	<ul style="list-style-type: none"> a. Paper & pencil test, completed by participant. 	<ul style="list-style-type: none"> a. Final class sessions.
<u>Post-Program Follow-up</u>		
<ul style="list-style-type: none"> a. Parolee Evaluation Form: collateral program participation, employment, new arrests, technical violations. b. Returns to prison. 	<ul style="list-style-type: none"> a. Monitoring form completed by supervising parole officer, and from FLOYD for later participants. b. Extracted from OTIS 	<ul style="list-style-type: none"> a. At the end of the 3rd, 6th, 9th and 12th month following participant's completion of R&R. b. During program and at 3 month intervals following the conclusion of program- 3-30 months following end of program.

Intake Data Sources

Participant Background Data. The State of Georgia maintains extensive data on its offenders through the Georgia Offender Tracking Information System (OTIS). Data obtained during prison intake and diagnostic interviews, and subsequently entered into OTIS, provided background measures of social and demographic characteristics. OTIS also furnished data pertaining to the criminal histories of offenders, (e.g., prior incarcerations, prior felonies and prior violent offenses) along with reading (WRAT) and IQ (Culture Fair) scores. Complete sets of social and demographic data derived from OTIS were available for 92.1% of the parole sample (n = 887) and 89.9% (n = 172) of the pre-release sample. Criminal history data were available for 97.8% (n = 942) and 100.0% (n = 192) of the parole and prerelease samples, respectively.

Psychological Classifications and Pretest Scores. Psychological classifications of offenders into three levels of Interpersonal Maturity and four personality types (Warren et al., 1966) (described in the measurement section below) were provided by the Jesness Inventory (Jesness, 1992). GBPP staff administered the 155-item true/false, paper and pencil inventories to the experimental and control groups at the beginning of the Cognitive Skills Program. Inventory results were available for 86.1% (n = 961) of the parole and 82.3% (n = 158) of the pre-release participants.

Facilitators also administered the Pride In Delinquency Scale as a pre-test of offender values. These were administered to both experimental and comparison group subjects. The PID was a 10-item paper and pencil assessment (Simourd, 1997). Tests were available for 89.6% (n = 863) of the parole sample and 84.4% (n = 162) of the pre-release group.

Program Data Sources

Descriptions of what took place during the Georgia Cognitive Skills Program derive from session evaluation forms, observer evaluation forms, and participant evaluation forms. Of course, these forms were obtained for experimental but not comparison group members.

Session Evaluation Data. At the end of each class, coaches completed a “session evaluation form” (see Appendix A) for all class members. Evaluation staff obtained these forms from five sessions (the same

sessions for each group) which spanned the full cognitive skills course. The five sessions attended to the following issues: (a) problem-solving, (b) creative thinking, (c) social skills, (d) values enhancement, and (e) critical reasoning. Although each of the 52 program groups provided the five session evaluation forms, session evaluations of at least two of the five sessions were available for only 485 (84.5%) of the program participants—82.6% (n = 388) of the parole sample and 93.3% (N=97) of the prerelease sample. Data were unavailable due to some participants dropping out of the program or being absent. Indicators of group size also were obtained from the session evaluation rosters.

Observer Evaluation Data. At two points during the delivery of each program (sessions 17 and 30), observers assessed classroom effectiveness and completed the observer evaluation form (see Appendix B). These observers consisted of 11 cognitive coaches with the experience of having run at least three Cognitive Skills Programs and attended in-service training related to the observation process. Cognitive coach skills under examination included class structure, sensitivity, organization, delivery, participation, and feedback to participants. Of the 44 parole groups, observers evaluated 38 (84.6%) of the seventeenth session and 37 (84.1%) of the thirtieth session. Similar figures for the pre-release group were 8 (100.0%) of the seventeenth session and 8 (100.0%) of the thirtieth session.

Participant Evaluation Data. Upon completion of the program, participants provided their own assessments of the program via a participants' evaluation form (see Appendix C). Items on the survey tapped participants' perceptions of the coach's teaching methods, relevance of the class, and classroom climate. These were available for: (a) 270 (57.4%) of the parolees assigned to the experimental group, and 251 (89.6%) of those who completed, and (b) 76 (73.1%) of the experimental pre-release offenders and 73 (96.1%) of those who completed.

Outcome Data Sources

Intermediate (upon completion of the program) and long-term follow-up measures were available through two posttests, and automated data pertaining to new offenses, technical violations, and employment.

Offender post-tests. The Colorado Offender Attitude Survey (COAS) (Johnson & Hunter, 1992)

(Appendix D) and the Pride in Delinquency Scale (PID) (Simourd, 1997) (Appendix E) were administered to experimental and comparison group participants upon the experimental groups' completion of the Cognitive Skills Program. Both groups also completed the PID at the beginning of the class. These were paper and pencil assessments. PID pretests were available for 863 (89.6 %) of the parole group and 162 (84.4%) of the pre-release group. Posttests were available for 583 (60.5%) of the parole group and 125 (65.1%) of pre-release group. The Colorado Offender Attitude Survey was available for 598 (62.1%) of the parole sample and 127 (66.1%) of the pre-release sample.

Recidivism Recidivism data included two indicators of criminal behavior: (a) returns to prison and (b) felony arrests/revocations. The GBPP provided this information through data files extracted from OTIS and from follow-up forms completed by parole supervisors and a program coordinator at the GBPP. No recidivism data were obtained for the pre-release participants at the time this report was prepared, because they had not been released for a sufficient period of time.

The GBPP tracked parolee returns to prison through OTIS. These data were available for all but four parolee study participants for times ranging from the beginning of their participation in the study until the conclusion of our data collection period (n = 959 or 99.6% of the sample). Because the GBPP conducted the Cognitive Skills classes at different times throughout the study period, and evaluation staff obtained readmission to prison data at one point in time (October 2001), the length of time covered on the readmission to prison data varies; parolees in groups who began their classes earlier in the study have lengthier follow up time frames on this outcome variable than those who entered into the study more recently. For this reason, data on returns to prison were available for all parolee cognitive classes and their control group counterparts up to 12 months after the program end dates, then fewer parolees are represented in the remaining follow up periods. Data for subsequent time periods (15, 18, 21, 24, 27, 30, and 33 months) characterize smaller proportions of the sample, so that return to prison data were collected on only 122 parolees or 12.2 percent of the sample at the 33 month follow up period.

Felony arrest/revocation data were collected through follow-up forms completed by parole supervisors and an additional GBPP data base, FLOID (Field Log of Interaction Data). From the beginning of Phase II (June, 1998) until May 2001, detailed follow up forms were compiled by parole supervisors

through referencing the FLOID database. Beginning in May, 2001, these data were downloaded directly into an SPSS readable file for researchers at the University of Cincinnati. The felony arrest/revocation data for experimental and comparison group participants originally were organized into four time periods: at the end dates of the program classes and at three, six, and twelve months after the program completion date. If parolees were discharged during one of the follow-up frames, their felony arrest data were censored for the subsequent follow-up periods. Only 18 parolees (1.9% of the sample) are missing felony arrest/revocation follow-up data for reasons other than attrition through discharge. Data unavailable due to discharge are as follows: (a) during program, 0.5%; (b) 3 month, 1.8%; (c) 6 month, 2.8%, and (d) 12 month, 4.8%. These are cumulative percentages. Five cases (2 control group members and 3 experimental group members) were discharged so early in the study (i.e., during the first half of the time that the class was being held) that they are excluded from all analyses pertaining to arrest/revocation.

Technical violation and employment outcome data. In addition to providing information on arrests and revocations, the follow up forms completed by GBPP staff and information downloaded from FLOID cover the following aspects of performance while on parole: (a) compliance with conditions of parole, (b) employment status, (c) residence,¹¹ (d) technical violations, (e) hearings, and (f) participation in other treatment programs. On such indicators as employment and technical violations, the follow-up measures describe only those parolees who remained on supervision during the time period in question and do not include those who had been discharged or returned to prison prior to the follow up period under examination. Cumulative data unavailable for technical violations and employment due to discharge or revocation are as follows: (a) during program, 1.4%; (b) 3 month, 9.2%; (c) 6 month, 15.6%; and (d) 12 month, 22.5%. There are several instances where parolees were in jail throughout a follow-up time period, and hence were not “at risk” of technical violations nor able to be employed in the community. These individuals were not included in the analyses for that particular time period, but may be represented in later follow up periods when they are once again in the community. The proportion of parolees excluded from analyses due to being in jail varies from 1.6% at the 12 month period to 3.4% at the 6 month period. Additional data are missing due to an inability to obtain the information. Across the four follow up time

¹¹ Data pertaining to residential stability were incomplete and, therefore, are not included in our outcome analyses.

frames, 4.5% to 5.9% of employment data and 1.9% to 2.2% of technical violation data could not be collected.

Measures

Independent and Control Variables.

Independent measures included the focal group assignment variable (Es versus Cs), as well as individual characteristics which were employed as control variables or as covariates, depending upon the analysis. In addition, program measures served as independent variables for tests of their impact on the outcomes of the program participants.

Social and Demographic Measures. These included intake measures pertaining to participants' age, race, sex, marital status, social class, educational attainment, employment status, and substance abuse history.

Age was represented in the analyses as a set of dummy variables, because we did not assume a linear relationship between age and the various indicators of parole success. The age dummy variables indicate whether parolees were 18-22, 23-27, 28-32, 33-37, and 38 years old and older at the time that the Cognitive Skills program began. As noted earlier, two previous studies of R&R found that age was a factor in the success of participants. In one study, those between the ages of 25 and 37 (Robinson 1995) had lower post-program recidivism rates than those in other age groups. Another study found that those older than 38 (Johnson & Hunter, 1992) had lower post-program recidivism rates.

Race was dichotomized according to "white" and "nonwhite" (white = 1). With the exception of three participants the nonwhite offenders were all African Americans; two study members were Native American and one was Hispanic.

Analysis of the differential effects of sex was conducted for the pre-release sample which included both males and females. All of the parolees, however, were male.

The employment and education data measured parolees' self-reported characteristics at prison intake. Subjects reported their employment status as one of the following: employed full-time, part-time, unemployed for less than six months, unemployed for six months or more, never had a job but were capable of working, student, incapable of work and "other". For purposes of multivariate analysis this variable was collapsed into four dummy variables: employed full-time, employed part-time, unemployed less than six months, and "other". Occasionally, the measure was dichotomized (employed vs. unemployed). Study members provided their educational attainment by noting the highest degree/certificate

of educational attainment or the highest grade completed if they had not finished high school. The variable was collapsed into a dummy variable indicating whether the participants had completed high school/earned a GED.

Socioeconomic status was ascertained by prison diagnostic counselors who categorized inmates as “welfare,” “occasionally employed,” “minimum standard,” or “middle class,” according to criteria specified above. This measure was collapsed into a dummy variable indicating whether parolees were middle class or not (middle class=1). Marital status at prison admission was coded according to seven categories: (a) single (never married), (b) married, (c) separated, (d) divorced (not remarried), (e) widowed, (f) common law marriage, and (g) other. Categories were collapsed to indicate married or unmarried (married=1).

Educational and Psychological Assessments. Intelligence scores. obtained through the Culture Fair Intelligence Test (Cattell & Cattell, 1973), involved a short assessment, used primarily as a screening tool, which did not require reading skills or familiarity with ethnocentric referents. In the analyses, IQ remained an interval level variable when employed as a control variable. However, when we assessed the impact of IQ on program effectiveness (see Chapter 4), we explored two cut point for dichotomizing the scale—one at 80 and another at 85). A cut-point of 80 coincided with the screening guidelines recommended in the program manual (Ross, Fabiano, & Ross 1989). However, 85 provided better variability on the item (i.e., more participants in the low IQ group).

The Wide-Range Achievement Test (WRAT) (Reid, 1986; Reynolds, 1986) measured offenders’ reading levels. These scores coincided with school grades, with the number after the decimal point representing the number of month in the grade. For example, a study participant with a reading score of 9.2 would read at a level typical of a person who is two months into his or her 9th grade in school. In subsequent analysis, reading level is dichotomized at the 5th grade reading level, the proficiency level suggested for cognitive programming (Ross, Fabiano, & Ross 1989). Support for the construct validity of intelligence and reading scores is suggested by a significant correlation between the two ($r = .40, p \leq .001$).

Personality and cognitive maturity measures were obtained through the Jesness Inventory (Jesness, 1996). Earlier research on correctional effectiveness found that specific programs, even good

programs, worked better with some offenders than others. The earliest of these studies noted specific personality and maturity determinants.

Study participants were portrayed according to two psychological dimensions of the JI system: (a) Interpersonal Maturity (I-level) and (b) personality subtypes. Building on the theoretically-derived I-level system developed by Sullivan, Grant, and Grant (1957), and later Warren (Warren et al., 1966), the JI placed offenders into one of three levels of Interpersonal Maturity, I-level 2 through I-level 4. I-level assesses how individuals perceive the world, particularly in terms of emotions and motivations (Warren, 1983). Knowledge of a person’s I-level also provides insight regarding the maturity of one’s world-view; movement to a higher level indicates increases in the complexity of thoughts and degree of interpersonal maturity. Figure 2 describes the characteristics associated with each I-level classification.

Figure 2. Summary of Interpersonal Maturity Levels (Harris, 1988; Jesness, 1988; Warren, 1966).

Construct	Definition	Relationship to Criminal Behavior
I-level 2	The primary concern of I-level 2s is whether or not their needs are being met. Other people serve merely as sources of gratification or are seen as barriers to their satisfaction. They are unable to understand or predict the behavior of others.	Criminal behavior may result from poor impulse control. Also, external pressures may lead to fear or anger. An inability to cope with the negative affect then results in delinquency.
I-level 3	The primary concern of I-level 3s focuses on sources of power. They are aware that their behaviors have consequences for others, yet they lack empathy. Their view of others is limited to stereotyped roles. They attempt to manipulate the environment to their favor, and typically engage in only short-term planning.	Criminal behavior result from attempts to gain peer approval, gratification of material needs, or an attempt to gain control in a situation via a “bad guy” role.
I-level 4	The primary concern of I-level 4s is developing their own individuality. They have internalized values and standards, and are rigid in their application of those standards when judging themselves and others. They feel guilt when they do not measure up to these standards, or feel conflict when admiring others’ behavior.	Criminal behavior may result in response to internal discord or situational tension that leads to acting out. These conflicts may yield internalized criminal standards and value systems.

In addition, the authors of the I-level classification system (see Warren et al., 1966) empirically identified 13 personality subtypes among juvenile offenders; the Jesness Inventory furnished scales measuring nine of these subtypes. Research among adults found that the 9 subtypes could be further collapsed into four personality types: aggressive, neurotics, dependents, and situationals (Van Voorhis 1994). Figure 3 describes the collapsed types. Appendix F presents the nine subtypes that make up the four collapsed personality types.

Figure 3: Summary of key personality constructs for the collapsed Jesness Inventory types (Van Voorhis, 1994).

Construct	Definition	Relationship to Criminal Behavior
Aggressives (Aa, Cfc, Mp) ¹	Aggressives tend to be manipulative. Their behavior is unpredictable and negative. Aggressives feel alienated and hostile, and have antisocial values. They are not likely to possess prosocial values.	Aggressives have a high probability of criminal behavior.
Neurotics (Na and Nx)	Neurotics tend to be anxious and insecure. They are often cynical, hostile, and act inappropriately when anxious.	Neurotics have a high probability of criminal behavior when they are unable to cope with anxiety.
Dependents (Ap and Cfm)	Dependents tend to engage in conforming behavior. They often follow others, including criminals. This type is less clearly defined among adults than among juveniles.	Dependents are less likely to engage in criminal behavior than other types. When criminal behavior does occur it may be the result of the influence of others and their limited cognitive functioning.
Situationals (Se and Ci)	Situationals tend to view convention positively. They often maintain good relationships and be prosocial, but can be rigid.	Situationals are less likely to participate in criminal behavior, and such behavior is situational if it does occur.

¹ The abbreviations indicate which of the original nine personality types are included in the collapsed types. See Appendix F.

The four personality types were each represented as dummy variables in the analyses. The I-level measure was collapsed into a dummy variable with I-level four equaling “1” and “0” depicting levels 2 and 3.

Staff at the University of Cincinnati scored the inventories using The Jesness Inventory Computer Program Windows™ Version (1996) software. For the software to score the inventories, each item required a valid response. Of the 1021 JI inventories received, 146 (14.3%) posed scoring problems due to missing or double responses on items. In an effort to make use of these problematic inventories, scoring rules were created at the University of Cincinnati where staff would assign a response to missing items under specific conditions. First, inventories would remain unscored, and therefore not used, if four or more of the 155 items were not sufficiently marked. Second, the dimensions of the personality types addressed by the specific items were taken into consideration. If multiple problematic items tapped into a single dimension, the inventory remained unscored. Finally, for those inventories not excluded based on the above criteria, responses to the problematic items were selected that favored the more prosocial of the two possible options, true or false. The inventories could be scored using the software once each item was assigned a response. One hundred and eight (75.5%) of the 143 problematic Jesness inventories could be fixed through this procedure. Overall, a total of 980 (84.8%) Jesness Inventories generated personality and I-level types.

Risk scores. Previous research on cognitive interventions found them more effective for medium to high risk offenders than for low risk offenders (Andrews & Bonta 1998; Correctional Services of Canada 1991; Johnson & Hunter, 1992), unless they were parolees (Robinson 1995; Van Voorhis et al., 2001). A measure of risk of reoffending allowed for further examination of whether outcomes were more favorable for higher compared to lower risk study participants.

Because risk scores were unavailable for 29.4% of all cases, we created a risk score from individual variables available through OTIS. Our intention was to recreate as nearly as possible the Salient Factor Risk Score (SFS) (Hoffman 1994), a valid and widely used risk assessment instrument for parolees. In doing so, it was necessary to modify three of the seven variables contained in the Salient Factor Score. Whereas the SFS called for inclusion of juvenile offenses, prior convictions and prior incarceration, information available to us pertained solely to the adult record. In addition, we were not able to obtain a

measure of whether the offender had a recent commitment free period of three years before the most recent offense. Finally, our measure of alcohol or drug addiction did not precisely mirror the corresponding dimension on the SFS. The SFS called for heroin/opiate dependence, and such a variable is not routinely collected by most correctional agencies. As an alternative measure, offenders were considered at higher risk if their diagnostic interviews indicated histories of addiction to alcohol or drugs.

The new risk score was a summation of the individual items available through OTIS. Due to data missing on the individual variables used to create this score, it was necessary to impute values for 20 cases.¹² Twenty-one cases remained without a risk score because no data were available on any of the items used to create the risk score. Scores varied from one to ten with low scores indicating higher risk. For our analyses, the interval level score was collapsed into a dummy variable. Medium to high risk study members (i.e., those with risk scores from one through seven) received a value of “1” on the dummy variable.

Although our risk measure is a modified estimate of the SFS, it does tap a number of factors known to be static predictors of reoffending (Van Voorhis & Brown 1996). In fact, the new measure showed both construct and predictive validity. It was significantly related to the Georgia OTIS score among those offenders who had a risk score on record ($\gamma = .81, p < .001$ for Phase II; $\gamma = .71, p < .001$ for Phase I) and to offender recidivism over the course of the follow-up period for this research ($r = -.35, p < .001$ for returns to prison and $r = -.34, p < .001$ for rearrests/revocations).

Programmatic Characteristics

Programmatic constructs included whether participants assigned to the experimental group received the full dosage of treatment and participated in appropriately sized groups. We also examined participants' level of understanding, enjoyment, extent of participation, and application of skills. In addition, coaches were rated by participants and an outside observer according to: (a) classroom structure and control, (b) interaction style with participants, (c) organization and professional demeanor, and (d) adherence to theoretical (social learning) dimensions of the program. Finally, participants indicated their

¹² In doing so, the interval level risk score was regressed on the other demographic and criminal history variables describing the case. The resulting regression equation was used to calculate values to replace the missing risk scores. 44

level of comfort with the group and evaluated the relevance of R&R to their daily lives. Measures reflecting these dimensions pertained solely to the experimental participants¹³.

Measures of program completion and class size. Offenders' completion of the Georgia Cognitive Skills Program was noted in Phase I to be an important determinant of program effectiveness (Van Voorhis et al., 2001). The measure of program completion was drawn from the session evaluation form for the 30th class session. These forms categorized each offender as: (a) never beginning the program, (b) beginning but not completing, or (c) remaining in the program by the 30th session, thereby completing or graduating from the program). The analysis depicted three modes of program completion: (a) graduated from the program, (b) did not complete the program for reasons related to criminal behavior, and (c) did not complete the program for non-criminal reasons. The latter describes offenders who were dismissed for reasons such as employment, health, childcare, or other non-criminal situations. Measures pertaining to class size were also obtained from the session evaluation forms. Group size is the average of the number of people who began the class and the number of people still in the class at session 30. These averages varied from 5 to 23 (mean = 9.3; median = 8.0). Each experimental participant received a score pertaining to the average size of the group in which he or she participated.

Coaches' determination of participants' enjoyment, understanding, application of skills, and level of participation, in addition to their characterization of the classroom climate consisted of five point scales ranging from negative (e.g., not attentive) to positive (e.g. attentive, volunteers) ratings (see Figure 4). A dichotomous variable characterized the participants' use of previously taught skills. The participation, enjoyment, understanding, and group atmosphere ratings were averaged across the five sessions. Hence, these measures represented mean scores for the five time periods, with greater scores indicating more favorable performance. The group climate average pertained to the group as a whole and therefore was the same for all participants in the group. Indicators of the participant's use of skills, as presented in Figure 4, were not averaged, but were reported as three separate measures for sessions 3, 17, and 30¹⁴.

¹³ The group of study members who were not part of the experimental design were also assessed on the various programmatic characteristics.

¹⁴ In later analyses, participant's use of skills is measured as an average of four reports of skill usage.

Figure 4. Session evaluation measures.

Dimension	Value¹	Description
Level of participation	1	Not attentive
	2	Attentiveness wavered
	3	Generally attentive, quiet and reluctant to answer
	4	Attentive, volunteers
	5	Attentive, volunteers
Level of enjoyment	1	Expressed strong dislike and was disruptive
	2	Expressed some dislike
	3	Were noncommittal
	4	Showed some enjoyment
	5	Expressed enjoyment and joined in willingly
Level of understanding	1	Is confused, resistant and quite
	2	Questions the concepts and applications
	3	Seems to understand but cannot apply
	4	Understands, cannot offer own examples but agrees with examples
	5	Understands and can offer examples of the skills
Group atmosphere	1	Divisive, quiet
	2	Some disagreement, not listening
	3	Listening, little active participation
	4	Agreeable, some participation
	5	Cohesive, participating
Use of previous skills	0	No, did not use previously taught skills
	1	Yes, used previously taught skills

¹ On the session evaluation form, more favorable responses were assigned lower values. The values and corresponding descriptions presented here reflect ratings after being reverse-coded.

Observer evaluation measures. Observers critiqued two program sessions (sessions 17 and 30) according to three-point scales (0 = needs improvement, 1 = satisfactory, 2 = very satisfactory) on 45 items. Scores for each item were averaged across the two session evaluation forms. Factor analysis (principle component extraction with varimax rotation) provided the basis for creating the scales by analyzing the interrelationships among items to reveal underlying themes (or factors) within the data (Hair, 1992). As shown in Table 6, each scale achieved eigen values greater than 4, and all factor loadings surpassed .60. Alpha reliabilities for each scale were greater than .85. Although additional factors with sufficiently high eigen values (>1.0) were produced by this analysis, they did not lend themselves to clear interpretation.

Table 6: Observer evaluation scales, results of factor analysis and item analysis^a.

Item #	Scale Name and Items	Factor Loadings
Facilitators Skills in Creating a Positive Classroom Atmosphere		
24	Objectivity/open-minded stance toward participants.	.67
25	Flexibility: ability to respond to where the group is.	.77
32	Attention to participants questions.	.78
33	Use of positive enforcement	.62
35	Ability to keep participants safe from ridicule from others.	.70
36	Skill in identifying participants feelings, beliefs, thoughts or attitudes.	.70
37	Ability to empathize with participants.	.80
38	Ability to confront participants without demeaning them.	.73
39	Respect for participants.	.76
40	Professional detachment (e.g., from personal issues, from taking comments personally, etc.).	.78
41	Boundaries with participants.	.66
	Eigenvalue	24.3
	Explained variance	54.00%
	Chronbach's alpha	.97
Organization and Professional Demeanor		
11	Adherence to scheduled starting and ending time.	.71
12	Adherence to lesson plan for the day.	.74
13	Adequacy of preparation for the lesson.	.73
19	Organization of the lesson.	.62
22	Summary of the lesson at the end.	.62
42	Level of cooperation between coaches.	.76
43	Involvement of both coaches	.75
	Eigenvalue	3.87
	Explained variance	8.61%
	Chronbach's alpha	.87
Classroom Structure		
3	Evidence of rules and sanctions for missing group meetings.	.84
4	Evidence of enforcement of rules and sanctions for missing group meetings.	.83
5	Evidence of rules and sanctions for tardiness.	.89
6	Evidence of enforcement of rules and sanctions for tardiness.	.86
7	Evidence of rules and sanctions for poor classroom behavior.	.87
8	Evidence of enforcement of rules and sanctions for poor classroom behavior.	.80
	Eigenvalue	2.51
	Explained variance	5.58%
	Chronbach's alpha	.95
Skill in Dealing with Difficult Participants		
1	Coaches control of the classroom.	.86
9	Ability to deal with difficult participants (e.g., hostile, angry, non-participating, disrespectful).	.64
10	Ability to deal with participants who dominate group discussions.	.63
18	Timely completion of days lessons.	.65
44	Engagement of most rather than some participants.	.64
	Eigenvalue	2.07
	Explained variance	4.60%
	Chronbach's alpha	.94

^aPrincipal component analysis, varimax.

The first scale described the coaches' skill in interacting with participants, according to such favorable characteristics as flexibility, boundaries, sufficient detachment, listening and reinforcement qualities, and understanding of participants. A second factor measured coaches' adherence to the prescribed structure for each class, including adherence to schedules and lesson content, preparation, and level of cooperation among co-facilitators. The third scale depicted classroom control issues such as appropriate use and enforcement of rules and sanctions. Finally, a separate scale describes the coaches' ability to deal with difficult participants. Scales represent averages of all items contained in the scales and range from 0 to 2; high scores indicate more favorable evaluations than low scores.

For descriptive purposes, analyses of the impact of group characteristics also examined the influence of specific items contained on the observer evaluation form. As discussed in Chapter 5, however, this required certain precautions regarding the increased likelihood of finding significance among multiple tests. Measures of central tendencies for the scales and items, shown in Table 7, indicated limited variability; the overwhelming majority of observations resulted in extremely favorable scores.

Table 7: Participant and Observer evaluation scale and item scores for parole and pre-release samples.

Scale/item	Parole			Pre-release		
	Mean	s.d	Range	Mean	s.d.	Range
Observers' Evaluation Form						
<u>Observation 1 Scales</u>						
Facilitators' skills in creating a positive classroom atmosphere	1.9	.3	1.0 – 2.0	2.0	.1	1.8 – 2.0
Organization and professional demeanor	1.8	.3	.9 – 2.0	1.9	.1	1.8 – 2.0
Classroom structure	1.7	.5	.0 – 2.0	1.6	.7	.0 – 2.0
Skill in dealing with difficult participants	1.8	.4	.2 – 2.0	1.9	.1	1.7 – 2.0
<u>Observation 1 Items</u>						
1 Coaches' control over participants.	1.8	.5	.0 – 2.0	2.0	.0	2.0 – 2.0
2 Coaches' ability to manage and enforce the class rules.	1.8	.5	.0 – 2.0	2.0	.0	2.0 – 2.0
3 Evidence of rules and sanctions for missing group meetings.	1.7	.5	.0 – 2.0	1.6	.7	.0 – 2.0
4 Evidence of enforcement of rules and sanctions for missing group meetings.	1.7	.5	.0 – 2.0	1.9	.4	1.0 – 2.0
5 Evidence of rules and sanctions for tardiness.	1.7	.6	.0 – 2.0	1.6	.7	.0 – 2.0
6 Evidence of enforcement of rules and sanctions for tardiness.	1.6	.6	.0 – 2.0	1.6	.7	.0 – 2.0
7 Evidence of rules and sanctions regarding classroom behavior.	1.7	.5	.0 – 2.0	1.6	.7	.0 – 2.0
8 Evidence of enforcement of rules and sanctions regarding classroom behavior.	1.7	.5	.0 – 2.0	1.9	.4	1.0 – 2.0
9 Ability to deal with difficult participants.	1.8	.5	.0 – 2.0	2.0	.0	2.0 – 2.0
10 Ability to deal with participants who dominate group discussion.	1.7	.6	.0 – 2.0	2.0	.0	2.0 – 2.0
11 Adherence to scheduled starting and ending time.	1.9	.3	1.0 – 2.0	1.9	.4	1.0 – 2.0
12 Adherence to lesson plan for the day.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
13 Adequacy of preparation for the lesson.	1.8	.4	1.0 – 2.0	1.8	.5	1.0 – 2.0
14 Clarity of presentations.	1.8	.4	.0 – 2.0	2.0	.0	2.0 – 2.0
15 Use of examples, illustrations, definitions, etc.	1.8	.5	.0 – 2.0	1.9	.4	1.0 – 2.0
16 Relevance of the examples, illustrations, definitions, etc.	1.9	.4	.0 – 2.0	1.9	.4	1.0 – 2.0
17 Use of a variety of instructional strategies.	1.7	.5	1.0 – 2.0	1.5	.5	1.0 – 2.0
18 Timely completion of the day's lesson.	1.8	.4	1.0 – 2.0	1.9	.4	1.0 – 2.0
19 Organization of the lesson.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
20 Review of points from the previous lesson.	1.7	.6	.0 – 2.0	1.8	.5	1.0 – 2.0
21 Pace, not too slow or too fast.	1.8	.5	.0 – 2.0	1.8	.5	1.0 – 2.0
22 Summary of the lesson at the end.	1.8	.5	.0 – 2.0	2.0	.0	2.0 – 2.0
23 Coaches' skill in assisting participants to self-discovery.	1.8	.4	1.0 – 2.0	1.8	.5	1.0 – 2.0

Table 7 continued.

Scale/item	Parole			Pre-release			
	Mean	s.d	Range	Mean	s.d.	Range	
Observers' Evaluation Form							
24	Objectivity/open-minded stance toward participants.	1.8	.4	1.0 – 2.0	1.8	.5	1.0 – 2.0
25	Flexibility: ability to respond to where the group is.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
26	Enthusiasm.	1.9	.3	1.0 – 2.0	1.5	.5	1.0 – 2.0
27	Attention to cognitive processes instead of content.	1.8	.4	1.0 – 2.0	1.8	.5	1.0 – 2.0
28	Ability to recognize when participants do not understand lesson concepts.	1.7	.6	.0 – 2.0	2.0	.0	2.0 – 2.0
29	Coaches' feedback to participants on their performance.	1.7	.5	.0 – 2.0	2.0	.0	2.0 – 2.0
30	Balance between participants' and coaches' discussion.	1.7	.6	.0 – 2.0	2.0	.0	2.0 – 2.0
31	Attention to participants' need for opportunities to practice.	1.6	.7	.0 – 2.0	2.0	.0	2.0 – 2.0
32	Attention to participants' questions.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
33	Use of positive reinforcement.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
34	Participants' level of understanding by the end of the lesson.	1.7	.6	.0 – 2.0	1.9	.4	1.0 – 2.0
35	Ability to keep participants safe from ridicule from other participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
36	Skill in identifying participants feelings, beliefs, thoughts, or attitudes.	1.8	.4	1.0 – 2.0	1.9	.4	1.0 – 2.0
37	Ability to empathize with participants.	1.9	.3	1.0 – 2.0	1.9	.4	1.0 – 2.0
38	Ability to confront participants without demeaning them.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
39	Respect for participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
40	Professional detachment.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
41	Boundaries with participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
42	Level of cooperation between the coaches.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
43	Involvement of both coaches.	1.7	.5	1.0 – 2.0	2.0	.0	2.0 – 2.0
44	Engagement of most rather than some participants.	1.9	.4	.0 – 2.0	1.9	.4	1.0 – 2.0
45	Level of group participation.	1.8	.4	1.0 – 2.0	1.9	.4	1.0 – 2.0

Table 7 continue.

Scale/item	Parole			Pre-release		
	Mean	s.d	Range	Mean	s.d.	Range
Observers' Evaluation Form						
<u>Observation 2 Scales</u>						
Facilitators' skills in creating a positive classroom atmosphere	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
Organization and professional demeanor	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
Classroom structure	1.8	.5	.0 – 2.0	2.0	.0	2.0 – 2.0
Skill in dealing with difficult participants	1.9	.4	.3 – 2.0	2.0	.0	2.0 – 2.0
<u>Observation 2 Items</u>						
1 Coaches' control over participants.	1.9	.4	.0 – 2.0	2.0	.0	2.0 – 2.0
2 Coaches' ability to manage and enforce the class rules.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
3 Evidence of rules and sanctions for missing group meetings.	1.9	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
4 Evidence of enforcement of rules and sanctions for missing group meetings.	1.9	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
5 Evidence of rules and sanctions for tardiness.	1.7	.5	.0 – 2.0	2.0	.0	2.0 – 2.0
6 Evidence of enforcement of rules and sanctions for tardiness.	1.7	.6	.0 – 2.0	2.0	.0	2.0 – 2.0
7 Evidence of rules and sanctions regarding classroom behavior.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
8 Evidence of enforcement of rules and sanctions regarding classroom behavior.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
9 Ability to deal with difficult participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
10 Ability to deal with participants who dominate group discussion.	1.9	.4	.0 – 2.0	2.0	.0	2.0 – 2.0
11 Adherence to scheduled starting and ending time.	1.9	.4	.0 – 2.0	2.0	.0	2.0 – 2.0
12 Adherence to lesson plan for the day.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
13 Adequacy of preparation for the lesson.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
14 Clarity of presentations.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
15 Use of examples, illustrations, definitions, etc.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
16 Relevance of the examples, illustrations, definitions, etc.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
17 Use of a variety of instructional strategies.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
18 Timely completion of the day's lesson.	1.9	.4	.0 – 2.0	2.0	.0	2.0 – 2.0
19 Organization of the lesson.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
20 Review of points from the previous lesson.	2.0	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
21 Pace, not too slow or too fast.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
22 Summary of the lesson at the end.	1.9	.4	.0 – 2.0	2.0	.0	2.0 – 2.0
23 Coaches' skill in assisting participants to self-discovery.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
24 Objectivity/open-minded stance toward participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0

Table 7 continued.

Scale/item	Parole			Pre-release			
	Mean	s.d	Range	Mean	s.d.	Range	
Observers' Evaluation Form							
25	Flexibility: ability to respond to where the group is.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
26	Enthusiasm.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
27	Attention to cognitive processes instead of content.	1.8	.5	.0 – 2.0	2.0	.0	2.0 – 2.0
28	Ability to recognize when participants do not understand lesson concepts.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
29	Coaches' feedback to participants on their performance.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
30	Balance between participants' and coaches' discussion.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
31	Attention to participants' need for opportunities to practice.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
32	Attention to participants' questions.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
33	Use of positive reinforcement.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
34	Participants' level of understanding by the end of the lesson.	1.8	.4	.0 – 2.0	2.0	.0	2.0 – 2.0
35	Ability to keep participants safe from ridicule from other participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
36	Skill in identifying participants feelings, beliefs, thoughts, or attitudes.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
37	Ability to empathize with participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
38	Ability to confront participants without demeaning them.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
39	Respect for participants.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
40	Professional detachment.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
41	Boundaries with participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
42	Level of cooperation between the coaches.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
43	Involvement of both coaches.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
44	Engagement of most rather than some participants.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
45	Level of group participation.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0

Table 7 continued.

Scale/item	Parole			Pre-release		
	Mean	s.d	Range	Mean	s.d.	Range
Observers' Evaluation Form						
<u>Average Observation Scales</u>						
Facilitators' skills in creating a positive classroom atmosphere	1.9	.3	1.0 – 2.0	2.0	.0	1.9 – 2.0
Organization and professional demeanor	1.9	.3	.9 – 2.0	2.0	.1	1.9 – 2.0
Classroom structure	1.9	.4	1.0 – 2.0	1.8	.4	1.0 – 2.0
Skill in dealing with difficult participants	1.8	.3	.6 – 2.0	2.0	.1	1.8 – 2.0
<u>Average Observation Items</u>						
1 Coaches' control over participants.	1.9	.4	.5 – 2.0	2.0	.0	2.0 – 2.0
2 Coaches' ability to manage and enforce the class rules.	1.8	.4	.5 – 2.0	2.0	.0	2.0 – 2.0
3 Evidence of rules and sanctions for missing group meetings.	1.8	.4	1.0 – 2.0	1.8	.4	1.0 – 2.0
4 Evidence of enforcement of rules and sanctions for missing group meetings.	1.8	.5	.0 – 2.0	1.9	.2	1.5 – 2.0
5 Evidence of rules and sanctions for tardiness.	1.7	.4	1.0 – 2.0	1.8	.4	1.0 – 2.0
6 Evidence of enforcement of rules and sanctions for tardiness.	1.6	.5	.0 – 2.0	1.8	.4	1.0 – 2.0
7 Evidence of rules and sanctions regarding classroom behavior.	1.7	.4	1.0 – 2.0	1.8	.4	1.0 – 2.0
8 Evidence of enforcement of rules and sanctions regarding classroom behavior.	1.8	.4	1.0 – 2.0	1.9	.2	1.5 – 2.0
9 Ability to deal with difficult participants.	1.8	.4	.5 – 2.0	2.0	.0	2.0 – 2.0
10 Ability to deal with participants who dominate group discussion.	1.8	.4	.5 – 2.0	2.0	.0	2.0 – 2.0
11 Adherence to scheduled starting and ending time.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
12 Adherence to lesson plan for the day.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
13 Adequacy of preparation for the lesson.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
14 Clarity of presentations.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
15 Use of examples, illustrations, definitions, etc.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
16 Relevance of the examples, illustrations, definitions, etc.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
17 Use of a variety of instructional strategies.	1.8	.4	1.0 – 2.0	1.8	.3	1.5 – 2.0
18 Timely completion of the day's lesson.	1.8	.4	.5 – 2.0	1.9	.2	1.5 – 2.0
19 Organization of the lesson.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
20 Review of points from the previous lesson.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
21 Pace, not too slow or too fast.	1.8	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
22 Summary of the lesson at the end.	1.8	.4	.0 – 2.0	1.8	.3	1.0 – 2.0
23 Coaches' skill in assisting participants to self-discovery.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
24 Objectivity/open-minded stance toward participants.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0

Table 7 continued.

Scale/item	Parole			Pre-release			
	Mean	s.d	Range	Mean	s.d.	Range	
Observers' Evaluation Form							
25	Flexibility: ability to respond to where the group is.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
26	Enthusiasm.	1.8	.3	1.0 – 2.0	1.8	.3	1.5 – 2.0
27	Attention to cognitive processes instead of content.	1.8	.4	1.0 – 2.0	1.9	.2	1.5 – 2.0
28	Ability to recognize when participants do not understand lesson concepts.	1.8	.4	.5 – 2.0	2.0	.0	2.0 – 2.0
29	Coaches' feedback to participants on their performance.	1.8	.4	.5 – 2.0	2.0	.0	2.0 – 2.0
30	Balance between participants' and coaches' discussion.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
31	Attention to participants' need for opportunities to practice.	1.8	.4	.5 – 2.0	2.0	.0	2.0 – 2.0
32	Attention to participants' questions.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
33	Use of positive reinforcement.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
34	Participants' level of understanding by the end of the lesson.	1.8	.4	.5 – 2.0	1.9	.2	1.5 – 2.0
35	Ability to keep participants safe from ridicule from other participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
36	Skill in identifying participants feelings, beliefs, thoughts, or attitudes.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
37	Ability to empathize with participants.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0
38	Ability to confront participants without demeaning them.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
39	Respect for participants.	1.9	.2	1.0 – 2.0	2.0	.0	2.0 – 2.0
40	Professional detachment.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
41	Boundaries with participants.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
42	Level of cooperation between the coaches.	1.9	.3	1.0 – 2.0	2.0	.0	2.0 – 2.0
43	Involvement of both coaches.	1.8	.4	1.0 – 2.0	2.0	.0	2.0 – 2.0
44	Engagement of most rather than some participants.	1.9	.3	.5 – 2.0	1.9	.2	1.5 – 2.0
45	Level of group participation.	1.9	.3	1.0 – 2.0	1.9	.2	1.5 – 2.0

Table 7 continued.

Scale/item	Parole			Pre-release		
	Mean	s.d	Range	Mean	s.d.	Range
Participants' Evaluation Form15						
Adherence to social learning methods.	4.3	.4	2.7 – 5.0	4.3	.5	2.6 – 5.0
Relevance of skills to participant.	4.2	.5	1 – 5	4.3	.6	2.6 – 5.0
Group climate.	3.9	.6	2 – 5	3.8	.6	2 – 5
1 At first I hated going to class.	2.4	1.1	1 – 5	2.8	1.3	1 – 5
2 My thoughts and feelings seem clearer to me now than they were before I participated in this class.	4.1	.8	1 – 5	4.2	.8	2 – 5
3 In the end, I enjoyed going to class.	4.2	.9	1 – 5	4.3	.8	2 – 5
4 By using the skills I have learned, I will know how to get out of a bad situation.	4.1	.8	1 – 5	4.2	.8	2 – 5
5 The skills and examples seemed pretty realistic.	4.1	.7	1 – 5	4.2	.8	2 – 5
6 The coaches were easy to talk to.	4.4	.7	1 – 5	4.3	.7	2 – 5
7 The skills were too easy.	3.2	1.0	1 – 5	3.3	.9	1 – 5
8 The skills were too difficult.	3.7	.7	1 – 5	3.8	.7	1 – 5
9 At least one of the coaches seemed really bored with this group.	4.0	.9	1 – 5	4.1	.9	1 – 5
10 The exercises were helpful.	4.1	.7	1 – 5	4.3	.7	2 – 5
11 The coaches seemed enthusiastic about teaching the class.	4.2	.8	1 – 5	4.2	.9	1 – 5
12 I attended these classes regularly.	4.2	.8	1 – 5	4.4	.7	2 – 5
13 I don't feel I needed the skills.	3.8	.9	1 – 5	4.0	.9	1 – 5
14 When I had trouble understanding parts of a skill, the coaches did a pretty good job of giving us examples and showing us what they meant.	4.4	.7	1 – 5	4.5	.6	3 – 5
15 Sometimes I had trouble understanding the skills.	3.0	1.1	1 – 5	3.1	1.2	1 – 5
16 The skills were explained too quickly.	4.0	.6	1 – 5	4.1	.6	1 – 5
17 The group members seemed to have trouble understanding the skills.	3.8	.7	1 – 5	3.9	.7	2 – 5
18 The coaches let me know from time to time whether I was understanding the skills.	4.0	.6	2 – 5	3.9	.8	1 – 5
19 The coaches gave me suggestions for how to improve the way I handled problems.	4.3	.6	2 – 5	4.3	.8	1 – 5
20 Sometimes we "role played" parts of the lesson.	4.4	.7	1 – 5	4.3	.7	1 – 5
21 Group members often did not cooperate well with the coach.	3.9	.8	1 – 5	3.7	1.0	1 – 5
22 Sometimes group members were teased and the coaches did not do anything about it.	4.1	.8	1 – 5	4.1	.8	1 – 5

15 Items were recoded so that for each statement higher scores reflect more positive assessments of the class.

Table 7 continued.

Scale/item	Parole			Pre-release			
	Mean	s.d	Range	Mean	s.d.	Range	
Participants' Evaluation Form16							
23	Sometimes we got some good discussions going.	4.4	.7	1 – 5	4.5	.6	2 – 5
24	Just a few people seemed to do most of the talking.	3.6	1.0	1 – 5	3.6	1.1	1 – 5
25	I felt that the coaches understood where I was coming from.	4.2	.6	1 – 5	4.2	.8	1 – 5
26	The coaches were helpful to me.	4.4	.7	1 - 5	4.4	.7	2 – 5
27	At least one of the coaches seemed to take our comments too personally.	3.9	.8	1 – 5	4.0	.9	1 – 5
28	Coaches used examples (video, pictures, or practice sessions) to help us understand the skills.	4.3	.6	1 – 5	4.3	.7	2 – 5
29	The coaches read from a manual most of the time.	2.6	1.1	1 – 5	2.9	1.1	1 – 5
30	I felt comfortable stating my own opinions in the classes.	4.2	.7	1 – 5	4.2	.8	1 – 5
31	Other group members did not seem to respect what I had to say.	4.0	.7	1 – 5	3.9	.8	1 – 5
32	I think I participated a lot in these classes.	4.2	.7	1 – 5	4.2	.7	1 – 5
33	I had several chances to practice the skills during group meetings.	4.2	.6	1 – 5	4.1	.8	1 – 5
34	I practiced the skills between meetings.	3.9	.8	1 – 5	3.9	.9	1 – 5
35	Most group members did not seem to be taking the classes seriously.	3.8	.9	1 – 5	3.7	.8	2 – 5
36	I completed my assignments to write in my anger diary most of the time.	3.3	1.1	1 – 5	3.4	1.1	1 – 5
37	I really did not take the anger diary seriously.	3.6	.9	1 – 5	3.6	.9	1 – 5
38	Not including the Cognitive Skills meetings, how many times per month did you meet with your supervising parole officer while you were in the Cognitive Skills program?						
	Once a month	45.2%					
	Twice a month	33.5%					
	Three times a month	10.3%					
	More than three times a month	11.0%					

16 Items were recoded so that for each statement higher scores reflect more positive assessments of the class.

Participant evaluation measures. Program participants provided the final set of program measures. At the end of the course, participants responded whether they 1) “strongly agreed,” 2) “agreed,” 3) had “no opinion,” 4) “disagreed,” or 5) “strongly disagreed” with 37 statements. Again factor analysis (principle component extraction and varimax rotation) (see Table 8) identified three factors among twenty-one items which were used to create the following scales: (a) coaches’ adherence to social learning methods, (b) relevance of the skills to program participants, and (c) group climate. The scale scores reflected the mean values of the items in the scales, and higher scores indicated more positive perceptions of the program. Chronbach’s alpha for each scale was greater than .75. The impact of single participant evaluation items on outcomes are analyzed for descriptive purposes. As with the observer evaluation measures results are mostly favorable (see Spruance et al., 2001). Scale/item ranges, and means are shown in Table 7.

Table 8 : Participant evaluation scales, results of factor analysis and item analysis^a.

Item	Factor Loading
Factor 1: Adherence to social learning methods	
The coaches gave me suggestions for how to improve the way I handled problems.	.63
Sometimes we “role played” parts of the lesson.	.56
Sometimes we got some good discussions going.	.56
I felt that the coaches understood where I was coming from.	.57
Coaches used examples (video, pictures, or practice sessions) to help us understand the skills.	.67
I felt comfortable stating my own opinions in the classes.	.70
I think I participated a lot in these classes.	.68
I had several chances to practice the skills during group meetings.	.67
	Eigenvalue 4.68
	Percentage of variance explained 12.64
	Internal consistency (Chronbach’s Alpha) .85
Factor 2: Relevance to of skills to participant	
My thoughts and feelings seem clearer to me now than they were before I participated in this class.	.71
In the end, I enjoyed going to class.	.62
By using the skills I have learned, I will know how to get out of a bad situation.	.75
The skills and examples seemed pretty realistic.	.66
The coaches were easy to talk to.	.57
The exercises were helpful.	.66
The coaches seemed enthusiastic about teaching the class.	.51
	Eigenvalue 3.92
	Percentage of variance explained 10.60
	Internal consistency (Chronbach’s Alpha) .82
Factor 3: Group climate	
Other group members did not seem to respect what I had to say.	.57
Group members often did not cooperate well with the coach.	.71
The group members seemed to have trouble understanding the skills.	.56
Sometimes group members were teased and the coaches did not do anything about it.	.58
Just a few people seemed to do most of the talking.	.62
Most group members did not seem to be taking the classes seriously.	.70
	Eigenvalue 3.59
	Percentage of variance explained 9.70
	Internal consistency (Chronbach’s Alpha) .78

^aPrincipal component analysis, varimax.

Outcome Variables

Intermediate outcome measures. The Pride in Delinquency Scale (PID) was intended to show change in offender thinking from the beginning to the end of the program. The PID is a 10-item, self-report, scale which asks for reactions to each of ten offenses. Respondents were asked to indicate what their level of pride versus shame would be on a ten point scale ranging from –10 (ashamed) to +10 (proud). After 100 points are added to the total of the ratings for the 10 items, scores can range from 0 to 200. The assessment was developed in Canada (Shields & Whitehall, 1991) and used initially with juvenile offenders. Subsequent studies have found the PID to be valid for adult offenders as well (Simourd, 1997). With adults, the PID has been found to be significantly correlated with Level of Service Inventory (LSI) scores (Andrews & Bonta, 1995), institutional misconducts, and prior offenses (Simourd, 1997).

Our experience with the PID raised some issues over the course of the study. Parole officers voiced concern about the instrument's face validity and noted that offenders often were amused by the test. In addition research staff at the University of Cincinnati flagged a number of assessments as potentially flawed. Post-tests were more seriously affected than pretests, and involved 68 (7.1%) of the parole cases but only 8 (4.2%) of the prerelease cases. Problems included cases where the handwriting for a pretest did not match that of the same individual's posttest and instances of response sets, e.g., where the same score had been selected for all items. Such problems often involved assessments that were received late and affected entire classes.

Tests of the validity of the PID produced disappointing results. Among parolees, neither the pretests nor the posttests were significantly related to risk or recidivism measures. Among prerelease women offenders, the PID posttest was significantly correlated with risk ($r=.38, p\leq.01$); the pretest was not. The instrument did not appear to be valid for pre-released male offenders. We had no recidivism data for this group. Ranges and standard deviations for the pretest and posttest scores are shown in Table 9 for both the parole and pre-trial samples. Alpha measures of internal consistency were unavailable for the present samples, but have been reported in other studies (Simourd, 1997).

Table 9: Pride in Delinquency (PID) and Colorado Offender Attitude Survey scores for parole and pre-release samples.

Scale/item	Parole			Pre-release			
	Mean	s.d	Range	Mean	s.d.	Range	
Pride In Delinquency							
Pretest	49.9	42.3	0 - 200	55.6	38.3	0 - 188	
Posttest	54.3	46.4	0 - 198	52.9	37.5	0 - 192	
Colorado Attitude Survey¹⁷							
1	Belief that criminal behavior/drug use is wrong (increase)	4.5	.5	2.0 – 5.0	4.3	.7	2.2 – 5.0
2	Favorable attitudes toward police (increase)	3.2	.7	1.0 – 5.0	3.2	.6	1.6 – 5.0
3	Favorable attitudes toward courts and judges (increase)	3.0	.7	1.0 – 5.0	2.9	.7	1.0 – 5.0
4	Belief that your parole officer is supportive (increase)	3.7	.6	1.2 – 5.0	3.4	.7	2.0 – 4.0
5	Belief that others regard you positively (increase)	3.9	.6	1.8 – 5.0	3.7	.6	2.2 – 5.0
6	Perceived prospects for achieving life goals (increase)	4.7	.4	3.0 – 5.0	4.6	.5	3.0 – 5.0
7	Problem solving ability (increase)	3.4	.5	2.0 – 5.0	3.4	.5	1.7 – 4.6
8	Empathy for others (increase)	3.6	.6	1.0 – 5.0	3.7	.7	1.3 – 5.0
9	Awareness and sympathy toward victims of crime (increase)	3.7	.7	1.5 – 5.0	3.8	.7	2.0 – 5.0
10	Self control (increase)	3.9	.6	1.3 – 5.0	3.9	.7	1.3 – 5.0
11	Acceptance of rationalizations for crime (decrease)	2.0	.6	1.0 – 4.0	2.0	.7	1.0 – 3.0
12	Tolerant attitudes toward drug use (decrease)	2.3	.7	1.0 – 4.8	2.3	.7	1.0 – 4.3
13	Sense of powerlessness/fatalism (decrease)	2.3	.6	1.0 – 4.3	2.3	.6	1.0 – 3.8
14	Normlessness/accepting illegitimate means (decrease)	2.1	.7	1.0 – 4.7	2.0	.7	1.0 – 3.7
15	Susceptibility of peer influences toward deviance (decrease)	2.0	.7	1.0 – 5.0	1.9	.7	1.0 – 4.3
16	General susceptibility to external influence (decrease)	2.4	.6	1.0 – 4.4	2.5	.7	1.0 – 3.8
17	Exposure to criminal friends (decrease)	1.5	.7	1.0 – 5.0	1.9	.8	1.0 – 5.0
18	Access to criminal resources (decrease)	2.2	.8	1.0 – 4.8	2.4	.9	1.0 – 4.3
	Prosocial sentiments scale (increase)	3.8	.4	2.6 – 5.0	3.8	.4	3.0 – 4.7
	Cognitive Skills Scale (increase)	3.8	.3	2.8 – 4.8	3.8	.4	2.9 – 4.7

¹⁷ Possible values vary from 1 to 5.

The Colorado Offender Attitude Survey (Johnson & Hunter, 1992) produced 18 scales and two summary scales measuring cognitive skills and criminogenic thinking. This study used the instrument as a post-test, only; no pre-tests were administered. The scales along with their measures of central tendency are reported in Table 9. The scales' measures of internal consistency appear in Table 10. Only eleven scales were found to have sufficient internal reliability (Chronbach's $\alpha > .60$). As a result, outcome analyses reported in Chapters 3 and 4 utilized only scales 1, 2, 4, 5, 6, 11, 15, 17, and the two summary scales. Even though most of the discarded scales measured cognitive skills, a combination of all the items comprising the cognitive skills scales was found to be internally consistent ($\alpha = .80$).

Table 10: Colorado Offender Attitude Survey scale alpha reliabilities and desired direction for change.

Scale Name	Alpha	Desired Direction
1. Belief that criminal behavior/drug use is wrong	.83	Increase
2. Favorable attitudes toward police	.64	Increase
3. Favorable attitudes toward courts and judges	.60	Increase
4. Belief that your parole officer is supportive	.70	Increase
5. Belief that others regard you positively	.66	Increase
6. Perceived prospects for achieving life goals	.67	Increase
7. Problem solving ability	.40	Increase
8. Empathy for others	.31	Increase
9. Awareness and sympathy toward victims of crime	.39	Increase
10. Self control	.53	Increase
11. Acceptance of rationalizations for crime	.76	Decrease
12. Tolerant attitudes toward drug use	.48	Decrease
13. Sense of powerlessness/fatalism	.46	Decrease
14. Normlessness/accepting illegitimate means	.51	Decrease
15. Susceptibility of peer influences toward deviance	.69	Decrease
16. General susceptibility to external influence	.37	Decrease
17. Exposure to criminal friends	.90	Decrease
18. Access to criminal resources	.75	Decrease
Prosocial sentiments scale	.80	Increase

Cognitive Skills Scale	.80	Increase
------------------------	-----	----------

At the time of its adoption for use in Georgia, the Colorado Offender Attitude Survey had not been well-validated. Moreover, Phase I results were quite disappointing; scale alphas were nearly identical to those shown in Table 10. Pretest, posttests comparisons reported for Phase I showed few significant differences (Van Voorhis et. al., 2001). By the time these limitations were discovered, Phase II was already underway.

Even so, Phase II validity tests for COAS scales were somewhat more promising among some of the groups studied during this research. Among the pre-release women, for example, the COAS criminal sentiments scale and the cognitive skills scale were significantly related to our dichotomous risk measure-- $r = -.20, p \leq .10$ and $r = -.25, p \leq .05$, respectively. Three of the eight additional scales also were significantly related to our dichotomous risk measure ($p \leq .05$). Two of the scales are related significantly to our full risk scale that varies from 1 to 10 ($p \leq .05$). As noted earlier, recidivism data were not available for the pre-release participants. No significant correlations between the COAS scales and risk were noted for male pre-release inmates.

Predictive validity tests for parolees found significant relationships between seven of the ten scales and either rearrests or returns to prison ($p \leq .05$). For example, both of the summary scales, criminal sentiments ($r = -.16, p \leq .01$) and cognitive skills ($r = -.16, p \leq .01$) were significantly related to returns to prison. Only three of the 10 scales were related to our dichotomous risk measure (scales 5, 6, and 17). Five scales, including both the criminal sentiments and cognitive skills scales, correlated significantly ($p \leq .05$) with parolees' risk when represented by the full risk scale that varies from 1 to 10. Among parolees, the PID and COAS scores were unrelated.

Criminal behavior outcome measures. The criminal recidivism data for both returns to prison and rearrests were coded into discrete time periods. For rearrests, the time periods included a during program period, and 4 additional time periods corresponding to 3, 6, 9, and 12 months following the date that the experimental participants completed their class. Although reports of the rearrests were not made during the 9th month, the activities of parolees during the time frames were usually associated with dates. Therefore, we were able to determine whether the arrests indicated on the 12 month forms took place within the time

frames of six-through-nine-months or ten-through-twelve-months after the program end date.¹⁸ Each follow up time period spanned three-months, with the exception of the “during program” time frame, which averaged 3.8 months.

We had hoped to differentiate misdemeanor arrests, felony arrests and revocations, but our data did not afford the opportunity to do so. This offense distinction often was not noted on the follow-up forms. Moreover, as is common to parole practice, the difference between an arrest and a revocation was often a matter of the parole agency’s discretion. Therefore, it was necessary to create a broader variable to capture more serious outcomes related to offending behavior which may or may not have resulted in a return to prison. A decision rule was created which coded the variable as “yes” if: (a) an arrest was clearly noted as a felony, (b) a revocation was noted; or (c) an arrest not noted as a felony or misdemeanor occurred along with an additional action, such as jail or revocation. Thus, the “felony arrest/revocation” dependent variable indicated whether a parolee was revoked and/or arrested for a felony during one of five periods of the study, including a during-program period and four post-program periods.

Data on readmission to prison were available during the program time frame through as much as 33 months after the program end date. The “readmission to prison” measure indicated whether a parolee was returned to prison (y/n) during one of the 12 time periods (one in-program period and 11, 3-month, follow up periods).

Technical violation and employment outcome data. The time periods covered by the technical violation and employment analyses include the following: (a) during program, (b) from program end to three months after program completion, (c) from four to six months following program completion, and (d) from seven to twelve months after program completion. Because technical violations were not reported with dates of the events, it was not possible to disaggregate the last time period into two three-month periods as we did with the rearrest data.

The technical violation outcome measure indicated whether parolees incurred at least one technical violation during each discrete follow up time period. The types of technical violations included:

¹⁸ Of the 79 felony arrests indicated on the 12 month follow up forms, 10 did not have dates associated with the arrests. In those instances, the events (arrests) were assigned randomly to either the nine or twelve month time period.

(a) electronic monitoring violations, (b) positive drug tests, (c) employment violations, (d) reporting violations, (e) absconding, (f) special condition violations, or (g) “other”.¹⁹

Employment data were collected to indicate the following variations of employment status: employed full-time, employed part-time, student full-time, not required to work, unemployed/not in school, and “other” which included having been detained or having absconded throughout the three-month follow up period. Of interest in these analyses was whether the proportion of parolees employed, including full- and part-time employment, differed by experimental condition. The measure of employment indicated whether, during a discrete time follow up time period, a parolee was employed either full- or part-time as opposed to the other variations of employment status mentioned above.

Data Analysis

Data analyses proceeded in the following manner. First, all outcome measures, including pretest posttest results, returns to prison, felony arrest/revocations, technical violations, and employment, were examined by comparing results for parolees in the experimental group to parolees in the control group. Outcome analyses for the pre-release groups involved only the pretest, posttest comparisons, because recidivism data were not available. Outcomes measured were then re-examined in conjunction with individual characteristics of the parolees (e.g., IQ, risk, race, and personality type, etc.) to determine whether such characteristics affected the impact of the Cognitive Skills Program. Finally, results for experimental group graduate members were examined for whether or not they were affected by specific programmatic factors. Of course, comparison group members were not included in this analysis, because there were no comparable programmatic measures for them.

Pretest, posttest analyses. PID analyses involved determining whether changes in pretest to posttest scores were significantly more favorable for experimental than control group participants. Multivariate analysis of variance (MANOVA) for repeated measures was used for this purpose. A second portion of the analysis, which tests for effects that are differentiated according to individual factors, utilized multivariate analysis of covariance (MANCOVA). Results of the Colorado Attitude Survey, were

¹⁹ Other, in this case, included non-compliance with additional disposition criteria or failure to comply with treatment/program requirements.

determined through analysis of variance (ANOVA) and individual effects were examined through analysis of covariance (ANCOVA). Because posttest results were not available for all members of the experimental and comparison groups, statistical tests involved controlling for factors which differentiated the comparison group from the experimental group. Even with the data attrition, there were no differences between the experimental and comparison groups among the parolees. Analysis of the pre-release women's sample, however, involved statistically controlling for race, class and risk; reading scores and number of prior incarcerations were controlled during analyses of data for pre-release males.

Returns to prison and felony arrests/revocations. The effects of the cognitive program on the likelihood of being arrested/revoked or readmitted to prison were revealed through the use of (discrete time) event history analysis. This statistical procedure estimated the probability or likelihood of an event happening (e.g., arrest/revocation) while considering the different lengths of time that the parolees were available or "at risk" for the event to occur. By using event history analyses we were able to see whether the proportion of study members who failed at the various discrete time periods differed for parolees assigned to the experimental conditions compared to the control group, and whether the rates at which parolees recidivated varied throughout the follow-up period.

As noted above, the discrete time periods were, for the most part, three-month time periods. The "during program" time period was an exception, because the length of programs varied from 1.5 to 5.8 months, with the average duration being 3.8 months. Hence the typical "during program" time frame was slightly longer than the other discrete time periods of the analysis.

Three factors contributed to variation in the length of parolees' time at risk. First, some parolees failed sooner than others in the evaluation. It would have been inaccurate to treat the length of time at risk as equivalent for parolees who failed early on parole compared to parolees who still had not recidivated at 30 months. The second scenario that lead to variation in the length of parolees' time at risk was availability of data. As mentioned above, readmission to prison data were available up to 33 months after class completion for only the groups that began their Cognitive Skills Program early in the study. By the 15 month follow-up period, groups started to fall out of the analysis. Finally, some parolees were discharged. In the analysis of arrest/revocation data, once a parolee was discharged, he was considered a success during

a specific time period, but data for subsequent time periods were no longer available.

Event history analysis accounted for these different lengths of time at risk by changing the units of analysis from persons to person-periods at risk. A case was created for each parolee at each discrete time period up to the point when he either failed or there no longer was recidivism data for him due to discharge or conclusion of his follow-up time frame. For example, a “person” unit for a parolee who did not fail on parole and for whom data extended from “during the program” through 30 months after the program was completed (a total of 11 discrete time periods), would be transformed into 11 “person-period” units. A parolee who recidivated during the six month period would be represented by only three “person-period” units: (a) during program, (b) three-months following the program, and (c) six months following the program. Once a parolee failed on parole or the span of data were exhausted, he was censored from subsequent time periods of the analysis. Survival (recidivism) curves shown in the analyses were nevertheless cumulative functions, reflecting success or failure during earlier time periods.

Once the overall effect of treatment was established via the comparisons between experimental and comparison groups, the analyses moved to examine differential treatment effects, analyzing whether some types of parolees did better than others. Individual factors included: (a) risk level, (b) inclination toward violence (as indicated by a history of convictions for violent crimes), (c) IQ, (d) reading level, (e) educational attainment, (f) marital status, (g) SES, (h) race, (i)age, (j) employment history, (k) personality types, and (l) interpersonal maturity level. Interaction terms were computed for each of the above characteristics by experimental condition (i.e., whether the parolee was in the experimental or control group). The interaction terms were then entered into the event history analyses, along with additive variables of all of the individual characteristics to partial out the effects of the individual level characteristics. Finally, interaction terms were tested through a decrement to chi square analysis for the extent to which they represented an improvement over the more simplistic additive model. When the decrement to chi square of the interaction model was less than .05, the interaction model was considered a significant improvement over the more parsimonious additive model, and indicative of instances where recidivism varied according to individual attributes.

The effects of programmatic characteristics on recidivism typically are revealed through both

event history analysis using logistic regression and linear regression. Logistic regression is used when examining the impact of program qualities on the probability of individuals recidivating. These analyses include experimental study members who graduated from the program. Linear regression is appropriate for determining how program characteristics affect the class-level rates of graduates recidivating. Programmatic predictors of rearrest and returns to prison include class size, participants' performance during the program, coach styles of interaction, and the structure of the classes. Each program characteristic is entered into the analyses separately; one model for each characteristic. Statistical models at the individual-level include controls for relevant demographic and criminal history characteristics.

Technical Violations and Employment. Initial comparisons between experimental and comparison groups determined whether or not a technical violation or employment occurred. For each follow-up period, we employed chi square analysis. Next, differential effects were examined through logistic regression and the use of interaction terms. Tests for differential effects, however, were only conducted for 12-month time period.

Chapter 3

Experimental and Control Group Comparisons

In this short chapter, we compare intermediate and long-term outcomes for comparison and experimental groups. Analyses are conducted separately for the parole and pre-release samples. In addition, test outcomes for men's and women's pre-release groups are analyzed for each group because (as discussed in Chapter 2) the two groups differ considerably from one another. Results for parolees who were not randomly assigned to experimental and comparison follow results for the randomly assigned parolees.

Intermediate Outcomes: Experimental versus comparison groups – Pride In Delinquency Pretest, Posttests

Tests of the effects of the Georgia Cognitive Skills Experiment on pretest, posttest changes in Pride In Delinquency scores are shown in Table 11. No significant differences were found between experimental and comparison parole group participants even when the sample is restricted to most valid cases (those which were not flagged by coders as appearing to be invalid). In addition, scores for parolees who were not randomly assigned were similar to those in the experimental group and showed no significant change over time.

Limited sample sizes marred tests for the pre-release groups. Table 11 suggests, however, that experimental, male, pre-release, inmates achieved more favorable posttest changes than comparison group participants, however, differences were not statistically significant. MANOVA tests of male pre-release inmates controlled for reading scores and prior incarcerations.

Differences for women, controlling for risk, class, and race, also were not significant. However, the results suggest that women in the experimental group performed less favorably upon the completion of the program than at the beginning, but again, these differences were not significant. We will explain in Chapter 4, that the increase in women's scores from pre-test to posttest is mainly attributable to low risk offenders whose posttest scores were considerably higher than their pre-test scores.

Table 11: Results of pretest, posttest comparisons (Pride In Delinquency Scale).

Scale (Desired direction)	Group	N	Pretest Means	Posttest means	Sig.	Nature of Change
Effects for Parolees						
PID Scale (decrease)	Experimental	279	50.0	53.7		
	Control	248	48.6	53.1		
PID Scale (most valid responses)	Experimental	250	48.6	51.7		
	Control	216	51.5	55.4		
Effects for Pre-release Males						
PID Scale (decrease)	Experimental	35	63.3	54.7		
	Control	18	53.1	56.8		
PID Scale (most valid responses)	Experimental	35	63.6	54.7		
	Control	18	53.1	56.8		
Effects for Pre-release Females						
PID Scale (decrease)	Experimental	31	40.2	55.2		
	Control	27	46.1	42.6		
PID Scale (most valid responses)	Experimental	31	40.3	55.2		
	Control	19	53.7	42.1		

**** $p \leq .001$; *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$

**Intermediate Outcomes: Experimental versus comparison groups —
Colorado Offender Attitude Survey posttests.**

Posttest scores for the Colorado Offender Attitude Survey are shown in Table 12. Results for the parolees find experimental participants achieved much less favorable posttest results than the members of the comparison group. Differences were significant for five of the ten comparisons. Results for the unassigned parolees are in line with those for the experimental group, but no comparisons could be made to a comparison group or to a pretest. Counter-indicated results such as these usually involve a check of raw data for evidence of systematic sources or error. In this case, however, we were unable to conduct a random

check for consistency, because COAS data were sent to UC in scantron-generated databases.

Among male pre-release inmates a significant difference between experimental and comparison group members was found on scale 17, Exposure to Criminal Friends, but not on any of the other scales. Again, members of the comparison group achieved more favorable results than those who participated in the Cognitive Skills Program.

Results for the female pre-release inmates made considerably more sense. Experimental participants achieving significantly higher mean scores on four of the ten scales: (a) Belief that Criminal Behavior is Wrong, (b) Prospects for Achieving Life Goals, (c) Belief that the Parole Officer is Supportive, and (d) the Total Cognitive Skills Scale.

Table 12: Group differences on the Colorado Offender Attitude Survey.

Scale	Experimental		Control		Sig.	Nature of Change	
	N	Mean	N	Mean			
Parolees (Parolees not in random assignment)							
1	Belief that criminal behavior is wrong (increase)	316 (70) ^a	4.5 (4.4)	283	4.6	p=.05	Cs less criminogenic
2	Favorable attitudes toward police (increase)	317 (70)	3.1 (3.0)	286	3.3	p=.05	Cs less criminogenic
4	Belief that P.O. is supportive (increase)	318 (70)	3.7 (3.8)	285	3.8		
5	Belief that others regard you positively (increase)	319 (70)	3.9 (3.8)	286	4.0	p=.001	Cs more positive
6	Prospects for achieving life goals (increase)	318 (70)	4.7 (4.8)	285	4.7		
11	Rationalizes criminal behavior (decrease)	318 (70)	2.0 (2.0)	286	1.9		
15	Susceptibility to peer influence (decrease)	318 (70)	2.0 (2.1)	286	1.9	p=.05	Cs less criminogenic
17	Exposure to criminal friends (decrease)	318 (70)	1.6 (1.8)	286	1.5		
	Total Criminal Sentiments (increase)	318 (70)	3.8 (3.7)	286	3.9	p=.001	Cs less criminogenic
	Total Cognitive Skills (increase)	318 (70)	3.8 (3.8)	286	3.9		

^a Figures in parentheses pertain to parole districts which did not participate in random assignment. They are offered for descriptive purposes only.

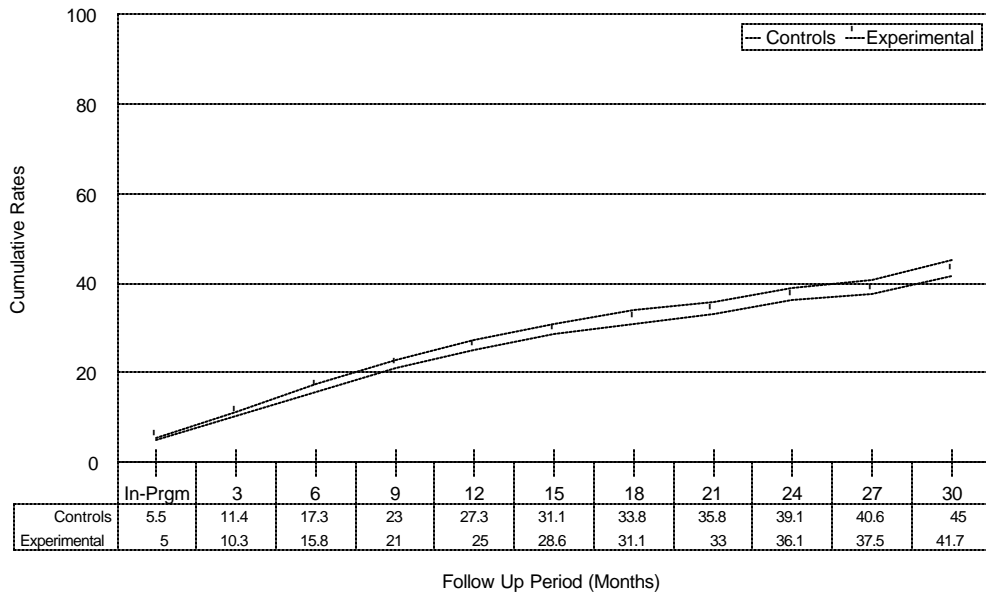
Table 12 continued.

Scale		<u>Experimental</u>		<u>Control</u>		Sig.	Nature of Change
		N	Mean	N	Mean		
Pre-release Males							
1	Belief that criminal behavior is wrong (increase)	46	4.2	19	4.3		
2	Favorable attitudes toward police (increase)	46	3.2	19	3.3		
4	Belief that P.O. is supportive (increase)	46	3.3	19	3.2		
5	Belief that others regard you positively (increase)	46	3.6	19	3.8		
6	Prospects for achieving life goals (increase)	46	4.7	19	4.9		
11	Rationalizes criminal behavior (decrease)	46	2.0	19	1.9		
15	Susceptibility to peer influence (decrease)	46	2.0	19	1.9		
17	Exposure to criminal friends (decrease)	46	2.1	19	1.7	p=.05	Es less criminogenic
Total Criminal Sentiments (increase)		46	3.7	19	3.7		
Total Cognitive Skills (increase)		46	3.8	19	3.9		
Pre-release Females							
1	Belief that criminal behavior is wrong (increase)	31	4.6	29	4.2	p=.01	Es less criminogenic
2	Favorable attitudes toward police (increase)	31	3.2	29	3.3		
4	Belief that P.O. is supportive (increase)	31	3.7	29	3.4	p=.01	Es less criminogenic
5	Belief that others regard you positively (increase)	31	3.8	29	3.7		
6	Prospects for achieving life goals (increase)	31	4.8	29	4.4	p=.01	Es less criminogenic
11	Rationalizes criminal behavior (decrease)	31	2.0	29	2.0		
15	Susceptibility to peer influence (decrease)	31	1.8	29	1.9		
17	Exposure to criminal friends (decrease)	31	1.8	29	1.9		
Total Criminal Sentiments (increase)		31	3.9	29	3.8		
Total Cognitive Skills (increase)		31	3.9	29	3.7	p=.05	Es less criminogenic

Long-term Outcomes: Returns to Prison

Results of event history analysis for parolees randomly assigned to experimental and comparison groups are shown in Figure 5. The readmission to prison analyses focused on 6,413 person-period units (experimentals = 3,162, controls = 3,251) derived from 959 parolee units. The curves shown in Figure 5 represent the predicted cumulative failure rates (inverse of a cumulative survival function) for each group by length of time (months) from the beginning of the Cognitive Skills Program. Throughout the 12 follow-up periods of the analysis, the control group members were readmitted to prison at a slightly elevated rate compared to the experimental group subjects. By end of the 30-month period, the predicted return to prison rates for experimental group and comparison group were 41.7% and 45.0%, respectively. The difference in rates for the two groups, 3.3%, was not statistically significant. These results were similar to those found for Phase I (see Van Voorhis et al., 2001). The predicted rates derive from the logistic regression equation presented in Appendix G.

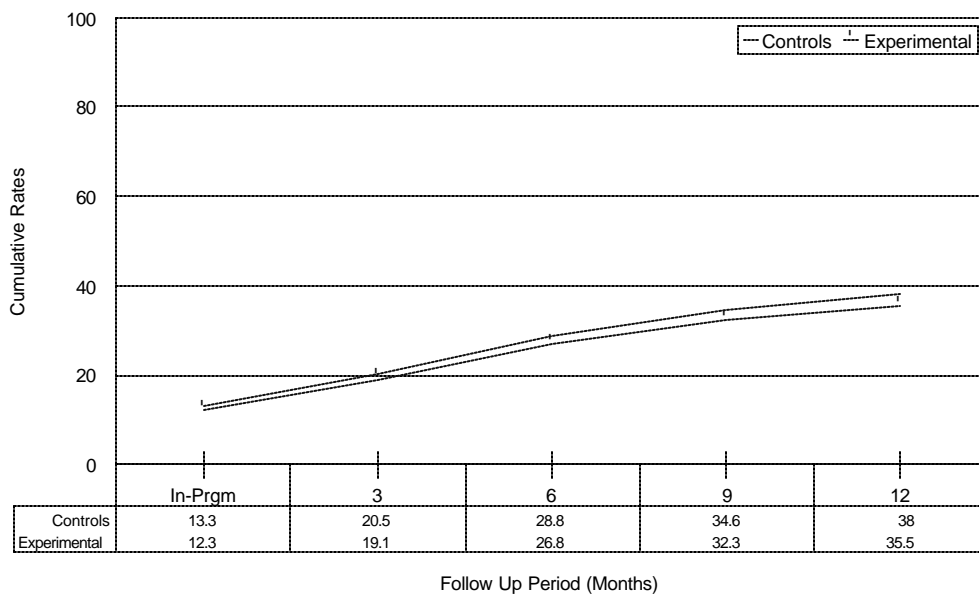
Figure 5: Predicted parolee return to prison rates by experimental condition and time.



**Long-term Outcomes: Parolee Experimental and Control Group Comparisons—
Rearrests/Revocations**

The felony arrest/revocation analyses used 3,702 person-period units (experimentals = 1,837; controls = 1,865) derived from 940 parolee units. As shown in Figure 6, by end of the twelve-month period 35.0% of parolees were arrested for a felony or revoked. Similar to the readmission model discussed above, group membership did not significantly affect rates of felony arrest/revocation. The control group parolees were arrested/revoked at slightly greater rates than the experimentals but the differences were not statistically significant. At the end of the twelve-month period 38.0% of controls were arrested/revoked compared to 35.5% of the experimental group. The logistic model is located in Appendices G.

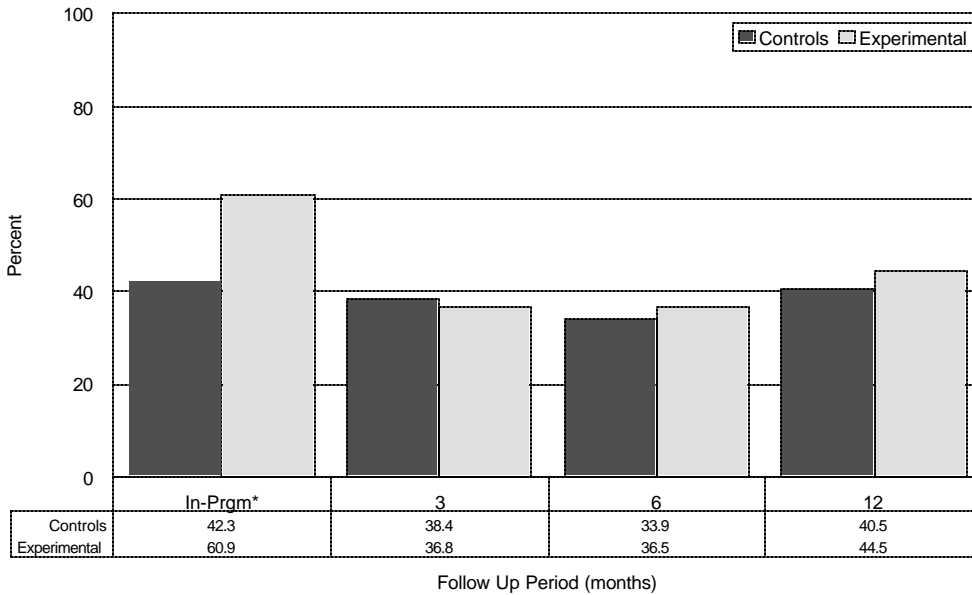
Figure 6: Predicted parolee felony arrest/revocation rates by experimental condition and time.



**Long-term Outcomes: Parolee Experimental and Control Group Comparisons—
Technical Violations**

Technical violations were examined across four time periods, during the program, three, six, and twelve month time frames. In contrast to the recidivism analyses, above, these analyses only involved those offenders who remained on parole supervision. Analyses were conducted separately for each time period and were not cumulative across time periods. Figure 7 reports a significantly higher technical violation rate for experimental versus control subjects at the “during program” time period (chi square = 31.5, $p \leq .001$). Sixty-one percent of the experimental group received at least one technical violation while the cognitive skills program was in session, compared to 42% of the control group. During the post-program time periods similar proportions of experimental and control groups received technical violations.

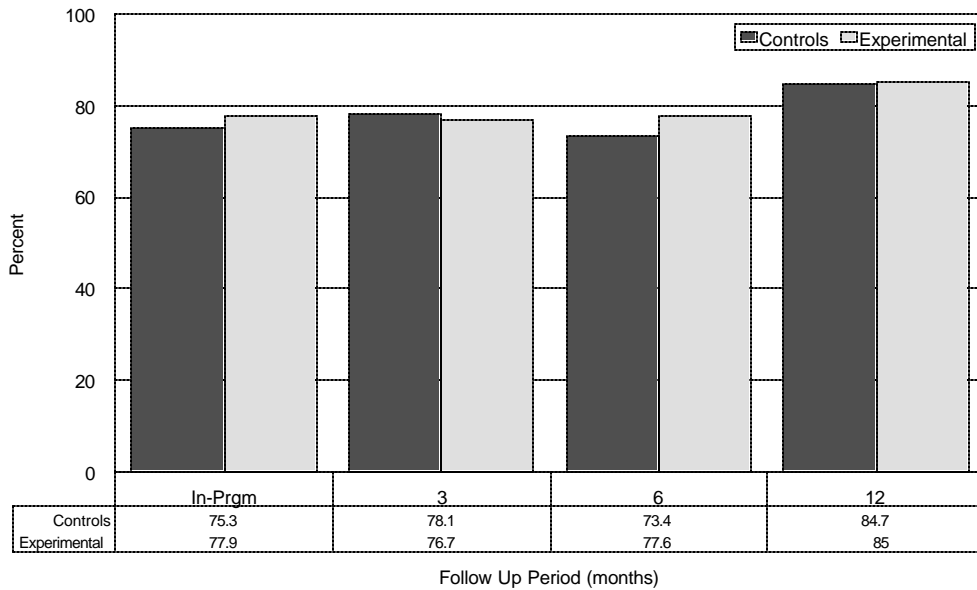
Figure 7: Parolee technical violations by experimental condition and time.



**Long-term Outcomes: Parolee Experimental and Control Group Comparisons—
Employment.**

Through each of the four time periods examined, rates of employment were high (greater than 70%) for both the experimental and control groups. Moreover, similar proportions of experimental and control group participants were employed at each time period. Chi square analyses indicated that assignment to either the experimental or comparison group did not affect whether parolees were employed during program or three-, six-, or twelve-months following program participation. Figure 8 presents the employment trends across the four time periods. Although a slightly greater proportion of control subjects were employed during the three month period (78.1% compared to 76.7% of the experimental group), experimental group employment figures were slightly better than the control group at the during program, six- and twelve- month time frames. None of these group differences are statistically significant.

Figure 8: Parolees employed full- or part-time by experimental condition and time.



Summary

In sum, a strict comparison between experimental and comparison group members found very few statistically significant differences. Only women pre-release offenders showed favorable and significant posttest results on the Colorado Attitude Survey. On several scales, parolees in the comparison group achieved *more* favorable COAS test results than those who participated in the cognitive class. Pride In Delinquency changes from pretest to posttest were not significant, regardless of the sample.

The Cognitive Skills Program also showed no significant impact on post program recidivism. Recidivism rates were only slightly lower for experimental participants than comparison group members (3.3% for returns to prison and 2.5% for rearrests/revocations). The program evidenced only negligible effects on technical violations and employment.

These results receive more discussion and interpretation in the concluding chapter of this report. Findings reported in this chapter represent our “first pass” through these data. The next two chapters, analyze the program’s impact in a more comprehensive manner. It is important to note that the study groups analyzed in this chapter combine those offenders who were most amenable to the program as well as those who were least amenable. In addition, 30 percent of the experimental group consisted of members who actually did not finish the program. The full research samples also do not differentiate between high and modest quality programs. These factors receive more attention in Chapters 4 and 5.

Chapter 4

Responsivity: Which Offenders Benefited Most from the Cognitive Skills Program?

Some types of experimental offenders achieved more favorable post-program outcomes than others; and others did worse than similar offenders in the comparison group. In this chapter, we compare specific types of experimental and comparison group participants on test, program completion, and recidivism measures. The analyses, in other words, attempted to identify offender characteristics that differentiated success from failure. The individual attributes considered for their differential impacts on offender outcomes were: (a) age, (b) race, (c) social class, (d) risk, (e) history of violence (f) marital status, (g) pre-arrest employment status, (h) education, (i) interpersonal maturity level, (j) personality, (k) reading level, (l) PID scores,²⁰ and (m) IQ.

Intermediate Outcomes

Generally, results were more meaningful for the recidivism and program completion analyses than for analyses involving pretests and posttests. Women pre-release participants were an exception to this generalization; pre-test, posttest comparisons on the PID measures and COAS posttests produced more meaningful results than those for either male parolees or pre-release offenders.

Pride In Delinquency Tests. Detection of offender-specific correlates of improvements or lack of improvements on the PID tests involved analyses for interaction effects between program participation and offender attributes. For example, to determine whether low risk offenders scored better on the PID assessments than medium to high risk offenders, we conducted a repeated measures test with risk as a covariate and risk and the experiment/comparison assignment measure as an interaction term (MANCOVA). Of importance to this analysis was whether the interaction term was found to be significant.

²⁰ PID scores were only considered in analyses of recidivism, technical violations, and employment. The analysis of PID scores dichotomized the PID score at the 25th decile, a point below which offenders were considered to have scored within the range of pro-social thinking. The test sought to determine whether offenders with prosocial thinking were harmed by the program or whether the program achieved its largest

Results for parolees are shown in Table 13. Only one interaction effect was detected, an interaction between marital status and PID change scores. Married experimental participants achieved more favorable PID improvements than married comparison group participants. Unmarried experimental and comparison group participants showed very little change from pretest to posttest. This effect remained significant when a multivariate analysis controlled for all other individual factors considered in these analyses.²¹

Table 13: PID pretest, posttest comparisons by offender characteristics, parolees.

Characteristic		N	Pretest Mean	Posttest Mean	Sig.	Nature of Change
Risk Categories	<u>High/Medium</u>					
	Experimental	253	51.94	55.53		
	Control	233	48.64	52.97		
	<u>Low Risk</u>					
	Experimental	19	34.84	35.53		
	Control	11	58.18	60.27		
Race	<u>Nonwhite</u>					
	Experimental	184	57.03	55.68		
	Control	175	51.31	53.63		
	<u>White</u>					
	Experimental	88	37.59	54.73		
	Control	69	43.38	50.23		
Age	<u>Below 26</u>					
	Experimental	97	50.13	47.74		
	Control	78	54.55	56.95		
	26-35					
	Experimental	104	50.32	51.62		
	Control	100	52.37	52.72		
	36-45					
	Experimental	48	57.52	63.81		
	Control	46	42.57	54.48		
	<u>Above 45</u>					
	Experimental	23	41.09	72.30		
	Control	20	26.20	39.20		
Social Class	<u>Middle</u>					
	Experimental	106	45.59	52.96		
	Control	98	47.45	53.70		
	<u>Other</u>					
	Experimental	166	52.38	54.89		
	Control	146	50.19	53.02		

²¹ The control variables were not intended to serve as controls for group differences, but rather to remove shared variation between marital status and other independent variables – in other words, to assess for the direct effects of marital status.

Table 13: continued.

Characteristic		N	Pretest Mean	Posttest Mean	Sig.	Nature of Change
Education	<u>Less than HS</u>					
	Experimental	179	51.64	56.63		
	Control	156	52.19	55.62		
	<u>High School Grad</u>					
	Experimental	93	49.01	49.34		
	Control	88	43.56	49.17		
Reading Scores	<u>Below Grade 5</u>					
	Experimental	67	60.49	65.49		
	Control	67	59.28	62.43		
	<u>Above Grade 5</u>					
	Experimental	205	47.56	50.42		
	Control	177	45.21	49.84		
IQ (Culture Fair)	<u>Below 85</u>					
	Experimental	28	55.57	65.49		
	Control	26	52.92	62.43		
	<u>Above 85</u>					
	Experimental	244	50.19	53.64		
	Control	218	48.61	53.22		
Married	<u>Not Married</u>					
	Experimental	208	52.44	49.50		
	Control	182	52.14	55.26		
	<u>Married</u>					
	Experimental	64	45.22	69.19	p<.05	Married Es performed worse than married Cs
	Control	62	40.06	47.53		
I-Level	<u>I2 & I3</u>					
	Experimental	170	53.76	57.31		
	Control	149	53.33	53.07		
	<u>I4</u>					
	Experimental	102	45.71	48.85		
	Control	95	42.40	53.65		
Jesness Personality Types	<u>Aggressive</u>					
	Experimental	86	56.45	61.29		
	Control	81	58.07	55.06		
	<u>Neurotic</u>					
	Experimental	41	60.73	54.49		
	Control	34	52.85	53.09		
	<u>Dependent</u>					
	Experimental	81	53.02	52.89		
	Control	69	45.43	48.22		
	<u>Situational</u>					
	Experimental	64	33.78	45.88		
	Control	60	38.97	56.87		

No significant interaction effects were identified for male pre-release offenders (see Table 14), however, the sample size and cell sizes were small, thereby limiting the statistical power of these test. In addition, PID assessments for male offenders were found to have limited validity among male parolees (see page 58, Chapter 2).

Table 14: PID pretest, posttest comparisons by offender characteristics, pre-release samples.

Characteristic		N	Pretest Mean	Posttest Mean	Sig.	Nature of Change
Differential Effects for Women Offenders						
Risk Categories	<u>High/Medium</u>					
	Experimental	24	50.04	52.83		
	Control	12	55.08	50.08		
	<u>Low Risk</u>					
	Experimental	7	6.57	63.29	p<.01	Low risk Es perform worse.
	Control	15	38.87	36.53		
Race	<u>Nonwhite</u>					
	Experimental	22	41.23	68.91		
	Control	12	53.75	44.67		
	<u>White</u>					
	Experimental	9	37.78	21.67		
	Control	15	39.93	40.87		
Social Class	<u>Middle</u>					
	Experimental	9	39.11	24.78		
	Control	14	36.50	37.50		
	<u>Other</u>					
	Experimental	22	40.68	67.64		
	Control	13	56.38	48.00		
Education	<u>Less than HS</u>					
	Experimental	22	41.05	63.45		
	Control	19	51.26	39.95		
	<u>High School Grad</u>					
	Experimental	9	38.22	35.00		
	Control	19	51.26	39.95		
Reading Scores	<u>Below Grade5</u>					
	Experimental	9	28.67	100.56		
	Control	5	37.72	69.80		
	<u>Above Grade 5</u>					
	Experimental	22	28.59	36.64		
	Control	22	48.00	36.36		

Table 14 continued.

Characteristic		N	Pretest Mean	Posttest Mean	Sig.	Nature of Change
Differential Effects for Women Offenders continued						
IQ (Culture Fair)	<u>Below 85</u>				p<.01	Es with IQ>85 improved, Es and Cs with IQ<85 performed worse.
	Experimental	13	31.92	82.77		
	Control	9	42.22	51.44		
	<u>Above 85</u>					
	Experimental	18	46.22	35.28		
	Control	18	48.00	38.11		
Married	<u>Not Married</u>					
	Experimental	28	40.61	54.68		
	Control	23	26.46	47.78		
	<u>Married</u>					
	Experimental	3	36.67	60.00		
	Control	4	44.25	12.50		
History of violence	<u>Yes, history</u>					
	Experimental	18	43.00	58.56		
	Control	16	46.19	31.56		
	<u>No history</u>					
	Experimental	13	36.38	50.54		
	Control	11	45.91	58.55		
I-Level	<u>I2 & I3</u>					
	Experimental	13	35.31	61.31		
	Control	13	40.62	39.38		
	<u>I4</u>					
	Experimental	18	43.78	50.78		
	Control	14	51.14	45.50		
Differential Effects for Male Offenders						
Risk Categories	<u>High/Medium</u>					
	Experimental	32	63.16	52.72		
	Control	14	53.79	60.50		
	<u>Low Risk</u>					
	Experimental	3	64.33	76.33		
	Control	4	50.75	43.75		
Race	<u>Nonwhite</u>					
	Experimental	19	63.32	58.05		
	Control	10	54.60	49.20		
	<u>White</u>					
	Experimental	16	63.19	50.81		
	Control	8	51.25	66.25		

Table 14 continued.

Characteristic		N	Pretest Mean	Posttest Mean	Sig.	Nature of Change
Differential Effects for Male Offenders continued						
Social Class	<u>Middle</u>					
	Experimental	16	70.56	59.00		
	Control	8	58.50	63.38		
	<u>Other</u>					
	Experimental	19	57.11	51.16		
	Control	10	48.80	51.50		
Education	<u>Less than HS</u>					
	Experimental	22	61.45	48.32		
	Control	13	50.92	57.92		
	<u>High School Grad</u>					
	Experimental	13	66.31	65.62		
	Control	5	58.80	53.80		
Reading Scores	<u>Below Grade5</u>					
	Experimental	4	62.25	50.25		
	Control	0	0.00	0.00		
	<u>Above Grade 5</u>					
	Experimental	31	63.39	55.32		
	Control	18	53.11	56.78		
IQ (Culture Fair)	<u>Below 85</u>					
	Experimental	2	57.00	42.50		
	Control	0	0.00	0.00		
	<u>Above 85</u>					
	Experimental	33	63.64	55.48		
	Control	18	53.11	56.78		
Married	<u>Not Married</u>					
	Experimental	31	62.19	55.00		
	Control	16	50.63	53.06		
	<u>Married</u>					
	Experimental	4	71.50	52.75		
	Control	2	73.00	86.50		
History of violence	<u>Yes. history</u>					
	Experimental	15	58.40	50.93		
	Control	8	43.75	54.50		
	<u>No history</u>					
	Experimental	20	66.90	57.60		
	Control	10	60.60	58.60		
I-Level	<u>I2 & I3</u>					
	Experimental	18	56.28	47.78		
	Control	14	46.64	52.21		
	<u>I4</u>					
	Experimental	17	70.65	62.12		
	Control	4	75.75	72.75		

Table 14 also shows that tests for women offenders in the prerelease sample identified two interaction effects. One suggests that medium to high risk program participants (in comparison to medium to high risk control group members) performed better than low risk groups; the other finds that women in the experimental group who also had IQ scores of less than 85 achieved worse post program gains than comparison group members with low IQ scores. For women with IQ scores greater than 85, experimental participants performed better than the comparison group. Both findings are consistent with the screening criteria recommended by program designers. 22

Colorado Attitude Surveys. These analyses considered only the summary measures of the COAS-- the criminal sentiments scale and cognitive skills scale. Tests of all COAS scales and potential interactions would have proven cumbersome and difficult to interpret.

Among parolees, no significant interactions were found. In other words, within categories pertaining to offender characteristics the same pattern as found in the previous chapter (experimental offenders achieved lower scores than comparison offenders) was found in the ANCOVA analysis. That is, offender attributes did not change the general pattern of findings for the group as a whole.

Among pre-release males, an interaction effect ($p < .05$) between treatment condition and marital status was found. Married experimental offenders scored higher on the cognitive skills scale than married comparison group participants. At the same time, unmarried comparisons were more successful than unmarried experimentals.

Results for women offenders in pre-release found interactions between program participation (or not) and race ($p < .05$), education ($p < .10$), and IQ ($p < .01$). All significant interactions occurred with the criminal sentiments scale and not the cognitive skills scale. Significantly higher program gains were observed for a) white program participants, b) women with less than a high school education (significance was only $p < .10$), and c) women with IQ scores greater than 85.

Program Completion Program completion data revealed that 145 (30.9%) of 470 experimental participants did not complete the Cognitive Skills Program. Non-completion occurred for crime-related and other reasons, such as employment, illness, family concerns, or coaches' decisions to terminate the client

22 It was not possible to examine a saturated model, with appropriate control variables, for the pre-release samples, because they were too small.

for missing too many classes or being disruptive. Because program completion may be considered either an intermediate outcome or a programmatic measure, we consider it in both this chapter and in chapter 5.

In this chapter, we conduct an attrition analysis to determine whether some types of offenders were more likely to fail to complete than others types. Our analyses were restricted to experimental participants. Logistic regression regressed the dichotomous program completion measure on age, race, social class, risk, history of violence, IQ, reading level, I-level, personality, and marital status. Results are shown in Table 15. Offenders most likely to fail to complete were: (a) middle class, (b) medium to high risk, (c) neurotic, and (d) participants between the ages of 23 and 27.

Table 15: Probability of dropping out of the program (experimentals only): logistic regression.

Independent Variables	B	Exp(B)
Age 23 to 27	0.516	1.676**
White	-0.070	0.933
Middle class	0.469	1.598**
Medium/high risk	0.577	1.781**
History of violence	-0.395	0.674*
IQ above 85	-0.026	0.974
Reads above 5 th grade level	0.270	1.310
High school graduate	-0.117	0.890
I level 4	-0.320	0.726
Neurotic	0.688	1.990**
Married	0.429	1.044
Intercept	-1.41*	
Model Chi Square	19.42**	
-2 log likelihood	498.76	

**** $p \leq .001$; *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$

Long-term Outcomes

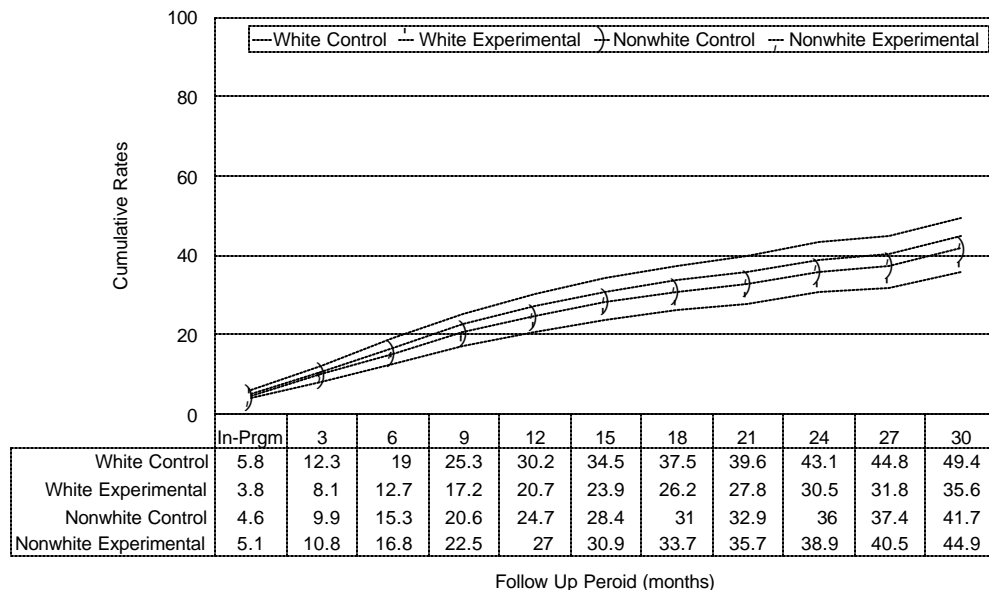
Returns to Prison and Rearrests. Tests for differential effects on returns to prison and rearrests/revocations involved entering interaction terms into the event history models, along with control variables to partial out the effects of other individual-level characteristics.²³ Three statistically significant interaction effects emerged from these analyses ($p < .05$) with respect to returns to prison: (a) race by

experimental condition, (b) age by experimental condition, and (c) personality type by experimental condition. Interaction analysis of the same independent variables on rearrests and revocations produced no significant findings.

We discuss the race, age, and personality finding for returns to prison below. The regression models for interaction analyses are located in Table 16 (p. 88).

Race by experimental condition. As shown in Figure 9, white parolees benefited from the cognitive skills treatment more than the nonwhite parolees. The white parolees assigned to the experimental group returned to prison at significantly lower rates than white parolees in the control group; by the end of the follow-up period, 49.4% of whites in the control group and only 35.6% of whites in the experimental group were returned to prison-- a treatment effect of 13.8%. Among nonwhites, all but three of whom were African American, the treatment effect was 3.2%, similar to that for the entire sample of parolees; the recidivism rate for nonwhites in the control group was 41.7 compared to a rate of 44.9% for nonwhites in the experimental group. Decrement to chi square analyses (Table 16) indicated that the race interaction model is a significant improvement over the additive model.

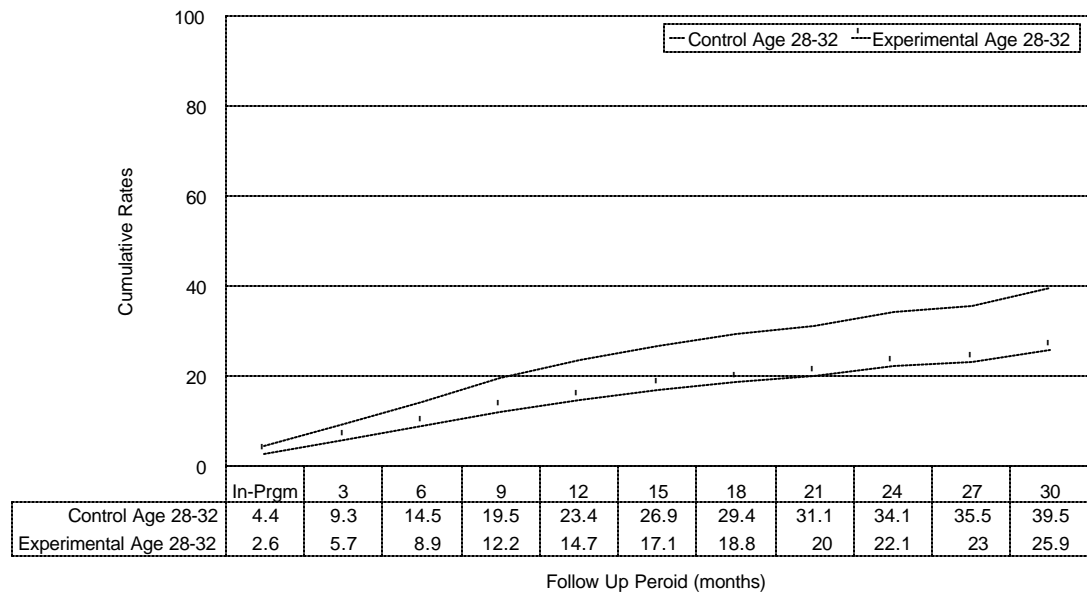
Figure 9: Predicted parolee return to prison rates by experimental condition, time, control variables, and race: interaction – race X experimental condition.



23 Control variables, in this instance, are not required to correct for group differences between experimental and comparison participants. Instead, we control for additional offender characteristics in order to remove statistically shared variation between the attribute being tested and other offender attributes.

Age by experimental condition. Returns to prison were modestly affected by participants' age. Treatment gains (differences between Es and Cs) were greatest, and slightly statistically significant ($p = .06$), for subjects between the ages of 28 and 32 (a 13.6% difference) (see Figure 10). Among offenders below 22 and above 38, the experimental participants appeared to have higher returns to prison than the comparison group, however, these differences were not statistically significant. Decrement to chi square analysis found that the interaction model is a slight improvement ($p = .07$) to the additive model. The logistic model is located in Table 16.

Figure 10: Predicted parolee return to prison rates by experimental condition, time, control variables, and age: interaction age 28 to 32 X experimental condition.



Personality by experimental condition. The personality interaction model examined if aggressives, neurotics, dependents, and situationals responded differently to the Cognitive Skills Program. Gains made through treatment for the aggressives (2.9%), dependents (7.3%), and situationals (6.4%) are fairly similar to each other and are not statistically significant for any group (see Figures 11a, 11b, and 11c). The cognitive skills program, however, significantly increased the likelihood of neurotics returning to prison; 35.3% of neurotics in the control group and 54.9% of experimental group neurotics returned to prison (see Figure 11d). The decrement to chi square analyses presented in Table 16 indicates that the personality interaction model is a borderline improvement over the additive model ($p = .09$). However, when the interaction model is re-run with a dummy variable capturing personality as “neurotic” = 1 and “other personality types” = 0, the interaction model is a significant improvement over the more additive model ($p \leq .05$).

Figure 11a: Predicted parolee return to prison rates by experimental condition, time, control variables, and personality types: interaction – aggressives X experimental condition.

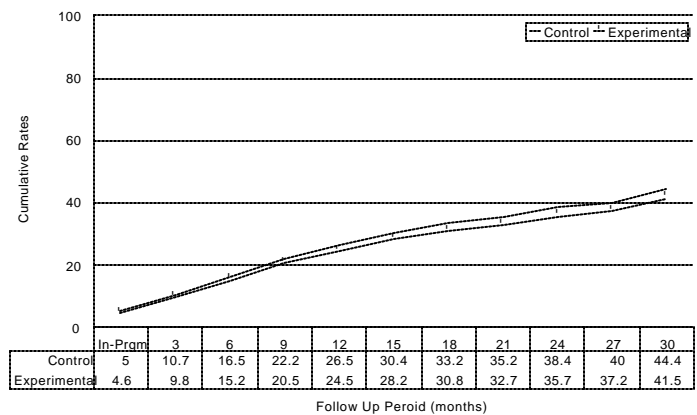


Figure 11c: Predicted parolee return to prison rates by experimental condition, time, control variables, and personality types: interaction – dependents X experimental condition.

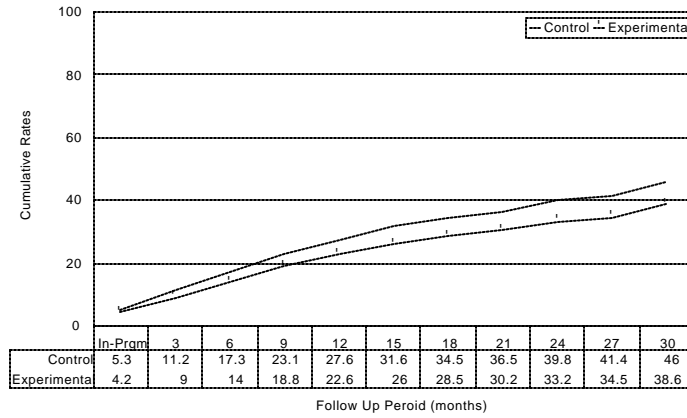


Figure 11d: Predicted parolee return to prison rates by experimental condition, time, control variables, and personality types: interaction – situationals X experimental condition.

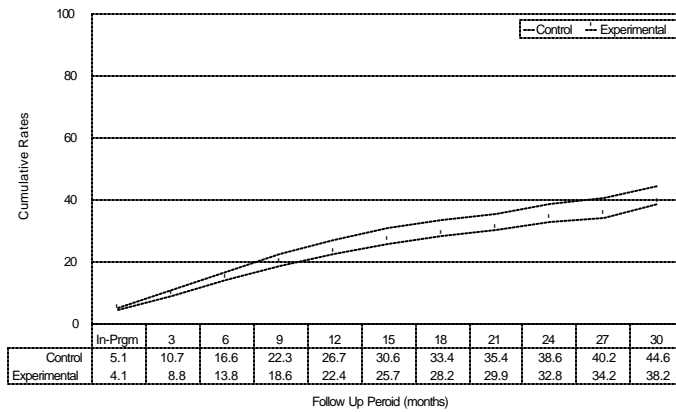


Figure 11d: Predicted parolee return to prison rates by experimental condition, time, control variables, and personality types: interaction – neurotics X experimental condition.

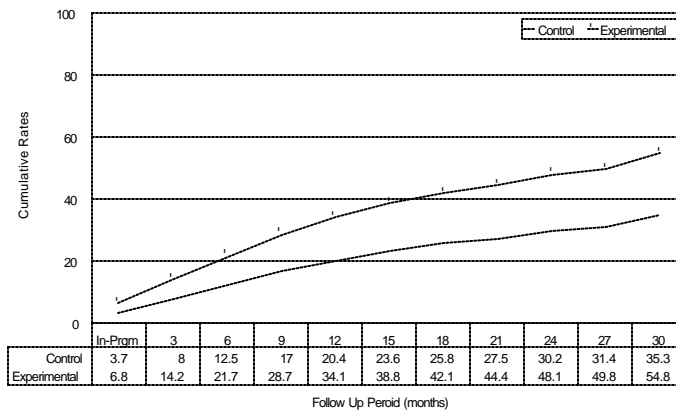


Table 16: Predicted return to prison by experimental condition, time, and interactions: race, age, and personality type X experimental condition.

Independent Variables	Race		Age		Personality	
	Interaction Model		Interaction Model		Interaction Model	
	B	Exp(B)	B	Exp(B)	B	Exp(B)
White X Experimental	-.55**	.58	----	----	----	----
Age 28 to 32 X Experimental	----	----	-.55*	.45	----	----
Neurotic X Experimental	----	----	----	----	.80**	2.22
Experimental condition						
Experimental group	.10	1.11	-.52*	1.31	-.18	.84
Time period ^a						
3 month	.18	1.20	.18	1.20	.18	1.19
6 month	.29	1.33	.28	1.33	.28	1.32
9 month	.32	1.37	.31	1.38	.31	1.37
12 month	.12	1.13	.11	1.13	.32	1.12
15 month	.05	1.05	.05	1.06	.05	1.01
18 month	-.25	.78	-.26	.78	-.25	.78
21 month	-.56	.57	-.57	.56	-.56	.57
24 month	-.01	.99	-.02	.98	-.01	.99
27 month	-.73	.48	-.75	.47	-.73	.49
30 month	.40	1.49	.38	1.43	.41	1.52
33 month	-4.17	.02	-4.21	0.2	-4.17	.02
Medium/high risk	1.11****	3.04	1.08****	3.08	1.08****	3.03
Prior conviction for violent crimes	.03	1.03	.02	1.03	.03	1.03
IQ	.00	1.00	.00	1.00	.00	1.00
Reading above 5 th grade	.00	1.00	.00	1.00	.00	1.00
High school graduate	-.24	.79	-.26**	.801	-.22	.80
Married	-.07	.93	-.06	.95	-.06	.95
Middle class	-.17	.85	-.14	.84	-.16	.85
White	.24	1.28	-.02	.99	-.01	1.00
Age ^b						
Age 18 to 22	.32	1.38	----	1.42	.36*	1.43
Age 23 to 27	-.04	.97	----	1.29	-.04	.96
Age 28 to 32	-.30	.74	-.13	1.01	-.30	.74
Age 33 to 37	.09	1.10	----	1.30	.07	1.07
Employment status ^c						
Employed fulltime	.38**	1.46	.33**	1.43	.38**	1.46
Employed parttime	.03	1.03	.02	1.03	.03	1.03
Unemployed < 6 mo.	.55***	1.73	.50***	1.73	.55***	1.72
Personality type ^d						
Aggressive	-.04	.97	-.02	.97	----	1.36
Dependent	-.05	.95	-.08	.95	----	1.43
Situational	-.08	.92	-.09	.88	----	1.37
Neurotic	----	----	----	----	-.31	.73
I-level 4	-.08	.92	-.12	.94	-.11	.90
Intercept	-4.04****		-3.89****		-3.96****	
Model Chi Square	101.89****		96.39****		103.25****	
-2 log likelihood	2523.85		2529.35		2522.49	
Decrement to Chi Square [#]						
Chi Square (df)	4.83**	(1)	3.22*	(1)	6.28***	(1)

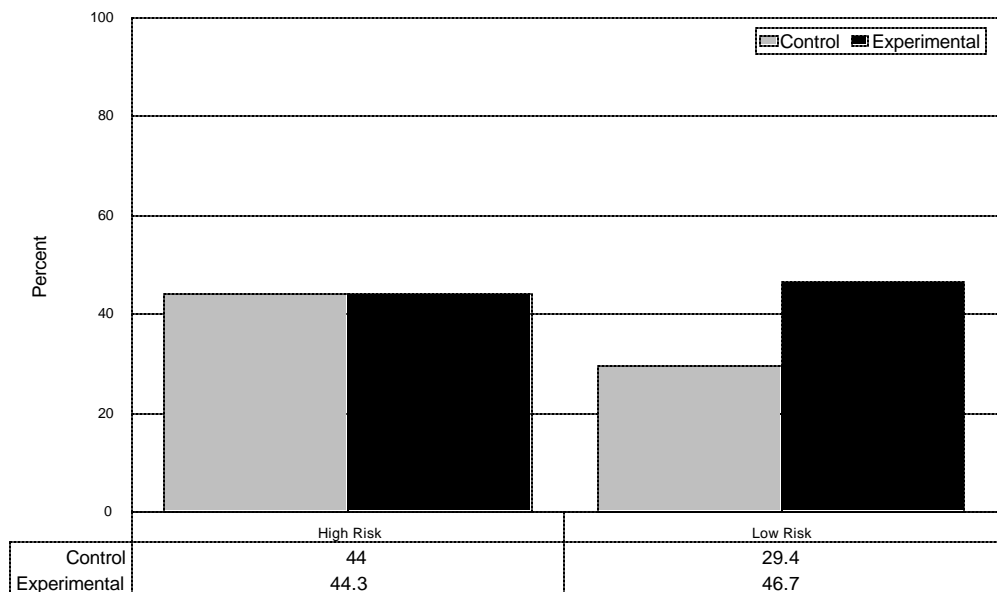
**** p ≤ .001; *** p ≤ .01; ** p ≤ .05; * p ≤ .10

The omitted dummy variables are (a) “during program”, (b) “age 38 and older”, (c) “other”, and (d) “neurotic”. # The decrement to chi square indicates whether the interaction model is a significant improvement over a simpler additive

model.

Technical violations and long-term employment. These analyses were conducted for the 12 month period, only. To account for all three follow-up periods would have entailed numerous tests and a cumbersome interpretation process. The analysis utilized logistic regression and interaction analysis of all independent factors considered in previous analyses reported in this chapter. Only one significant interaction effect was detected, an interaction between risk and technical violations occurring during the 10th to 12th month follow-up period. As shown in Figure 12, low risk experimental parolees still on supervision incurred significantly higher rates of technical violations than low risk comparison group members. Rates were similar for high risk experimental and comparison group members. This logistic regression equation is in Appendix H.

Figure 12: Predicted parolee technical violations at 10 to 12 month follow up by experimental condition, control variables, and risk: interaction – risk X experimental condition.



Summary

Table 17 organizes the findings reported throughout this chapter. They will be discussed further in Chapter 6.

Table 17: Summary of Differential Effects on Program Outcomes.

Parolees				
Group/Outcome	No interactions	Es better than Cs	Es similar to Cs	Es worse than Cs
Pride In Delinquency			Unmarried	Married
COAS – Skills	No interactions	----	----	----
COAS -- Sentiments	No interactions	----	----	----
Returns to Prison		Whites Age 28 to 32	Non-whites Ages < 32 & > 28 Aggressives Situationals Dependents	Neurotics
Rearrests/revocations	No interactions	----	----	----
Technical violations			Medium/high risk	Low risk
Pre-release Males				
Group/Outcome	No interactions	Es better than Cs	Es similar to Cs	Es worse than Cs
Pride In Delinquency	No interactions	----	----	----
COAS – Skills		Married		Unmarried
COAS -- Sentiments	No interactions	----	----	----
Pre-release Females				
Group/Outcome	No interactions	Es better than Cs	Es similar to Cs	Es worse than Cs
Pride In Delinquency		IQ>85	Medium/high risk	Low risk IQ<85
COAS – Skills		Lower class (p<.10)	Middle class	
COAS -- Sentiments		Whites IQ>85 Less than HS (p<.10)	Non-whites	IQ<85 HS grad

Chapter 5

The Characteristics of Effective Programs

“What Works?”

Based on the research on best practices in correctional treatment and on the theoretical framework of R&R, several programmatic factors were anticipated to affect the success of the Georgia’s Cognitive Skills program. Phase II of the Georgia Cognitive Skills Experiment was designed to measure the quality of the cognitive skills classes and to identify those programmatic factors which differentiated effective programs from ineffective ones. The factors considered in these analyses included program completion and class size as well as scales and items derived from participants, coaches, and observers.

Our design for this portion of the analysis differs from that utilized in earlier sections of the report. For parolees, the sample for these tests included all experimental participants along with the 78 parolees who participated from districts that did not participate in the random assignment.²⁴ Comparison group members were not included. Pre-release women consist of only those women who participated in and completed the program. We do not consider pre-release males in these analyses, because the test data are of questionable validity, and analyses of these data up to this point have been unproductive. For the same reason, we are not analyzing programmatic effects on test scores of male parolees.

Programmatic characteristics examined in this chapter derive from multiple sources. As noted in Chapter 2, coaches provided information on participants by completing session evaluation forms which measured participants’ use of skills, enjoyment, understanding, level of participation, and classroom climate. Observers’ evaluation of the coaches’ performance rendered four scales (a) facilitators skills in creating a positive classroom atmosphere (e.g., respect for participants, and ability to confront participants without demeaning them), (b) organization and professional demeanor (e.g., adherence to lesson plan for the day and preparation for the lesson), (c) classroom structure (e.g., evidence and enforcement of rules pertaining to classroom behavior), and (d) skill in dealing with difficult participants (e.g., engagement of most rather than some participants). Participants provided their perceptions of the program through

²⁴ These individuals participated in classes which utilized the same curriculum and completed the same evaluation instruments as the experimental participants.

participant evaluation forms. The scales formed from these included: (a) coaches' adherence to social learning clinical dimensions, (b) relevance of the course to participants, and (c) classroom climate.

These analyses were not without their constraints. Most importantly, our process study for Phase II found that most of these factors showed limited variability (Spruance et al., 2001). Whether it was attributable to the high quality of the training afforded the cognitive coaches, to the well-constructed program manual, or to the qualifications of the coaches, most evaluators (including offenders) rated most coaches and classrooms favorably. As a result, limited variability on the programmatic measures was likely to have reduced the number of significant findings possible for this portion of the analysis.

We begin the chapter with an examination of whether receiving the full program (i.e., graduation) improves parolees recidivism rates over participation in only a portion of the class. These analyses included all of the experimental and control study members. After establishing the importance of completing the program, we exclude the controls from the analyses and focus on how programmatic characteristics affect whether members of the experimental group graduate from the Cognitive Skills program. Subsequent sections of the chapter hold "program completion" constant by limiting the sample to only those parolees who graduated from the program.²⁵ Our final analyses assess the effects of program factors on the test outcomes achieved by pre-release women.

Typically, two levels of measurement were used to test the effects of program characteristics on program completion and recidivism--individual-level and class-level. First, event history analysis using logistic regression was employed to assess the impact of the programmatic factors on individuals' recidivism rates. The question here was whether the class characteristic (e.g., class size) affected the likelihood of an individual recidivating, controlling for relevant individual characteristics. On measures pertaining to offender evaluations, we could view this as an analysis of whether an individual's perception of a classroom characteristic impacted recidivism. Second, linear regression was used to examine how the program characteristics impacted the recidivism rates for classes. This changed the focus from individual perceptions and outcomes to what types of classes achieved what types of outcomes for the class as a whole. In the second set of analyses, the unit of analysis was the class, whereas in the first set of analyses

²⁵ Restricting the analysis to program completers was also mandated by the fact that many program measures were only available for offenders who completed the class.

the individual was the unit of analysis. These individual-level and class-level models were run for both felony arrest/revocations and returns to prison indicators of recidivism.

This chapter is organized in the following manner. First, the effects of program completion on recidivism are presented, followed by an examination of program characteristics that impact program completion. Then, we examine whether and in what ways programmatic characteristics impact rearrests and returns to prison. When addressing the manner in which recidivism is affected by program characteristics, the results are presented for one information source at a time (i.e., coaches evaluations, observers, and participants'). The chapter concludes by summarizing the findings.

Effects of Program Completion: Returns to Prison and Rearrests.

The results of statistical models testing the effects of program completion on recidivism were not surprising. As with the Phase I study (Van Voorhis et al., 2001) significant and substantial reductions in reoffending occurred for those program participants who graduated from the Cognitive Skills program. To examine this issue, the pool of study members were disaggregated into four groups: (a) experimental group members who graduated from the program (n = 280); (b) experimental group members who did not complete the program due to involvement in criminal behavior (e.g., an arrest) (n = 45); (c) experimental group members who did not complete the program for reasons other than criminal behavior (e.g., conflicts in scheduling) (n = 145), and (d) control group members (n = 493). The analyses addressing the effects of program completion, thus, contrasted the four groups of parolees distinguished by whether they received the full dosage of treatment and their reason for not partaking of the full program (i.e., graduates, noncriminal dropouts, criminal dropouts, and controls). Control variables for individual differences in criminal history, demographics, and risk level were also included in the models.

As can be seen in Figures 13 and 14, the recidivism rates for the control group were 36.0% for felony arrests/revocations and 40.8% for returns to prison. When parolees opted out the program for noncriminal reasons, they experienced no gains from the program; in fact, their recidivism rates were similar to those for the control group who received none of the Cognitive Skills Program. By the end of the follow up periods, 35.2% of the noncriminal dropouts were arrested for a felony and/or revoked and 42.7% were returned to prison.

The parolees who graduated from the program recidivated at significantly lower rates than the control and dropout groups. Less than thirty percent of the graduates had a felony arrest/revocation (24.8%) or were returned to prison (28.8%). Hence, receiving the full dosage of treatment led to treatment effects of 11.2% on felony arrest/revocations and 12.0% for returns to prison. The logistic regression table for these analyses is presented in Appendix I.

Figure 13: Predicted parolee return to prison rates by program completion, time, and control variables.

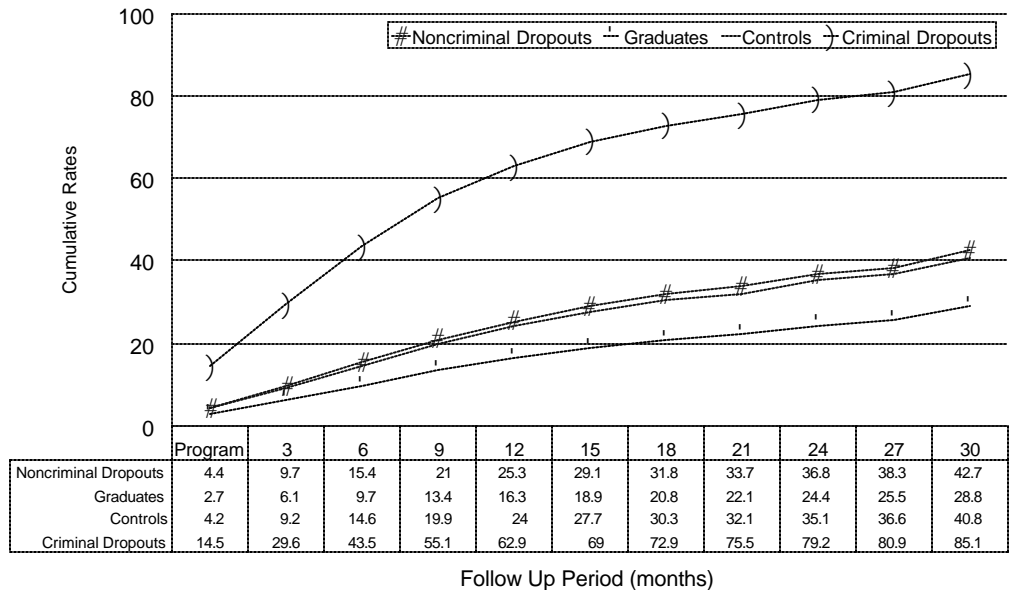
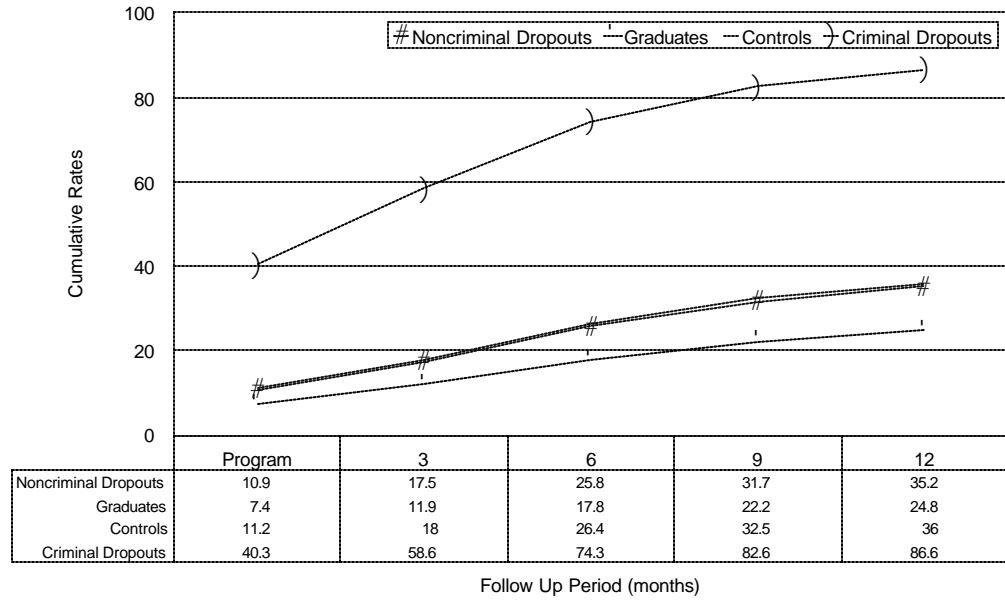


Figure 14: Predicted parolee felony arrest/revocation rates by program completion, time, and control variables.



Effects of Program Characteristics on Program Completion

The above section establishes the importance of program participants completing treatment. In this section we examine whether information available that describes what occurs during the program affects the likelihood of program completion. The following analysis contrasts program graduates from participants who fail to complete the program for reasons other than criminal behavior. By excluding the individuals who dropped out the program due to criminal involvement, we focus on those parolees who arguably could have completed treatment. The question then becomes, given that participants were not removed from the program for criminal behavior, are there characteristics of the program that increased: (a) the likelihood of an individual dropping out of the program and (b) class rates of dropouts. In chapter 4 (p. 82), characteristics of program participants associated with program attrition for any reason were discussed. Program graduates tended to be below middle class, low risk, not neurotic, and either younger than 23 or older than 27.

To explore the programmatic influences on graduating, we first regress the individual-level variable “program completion” (1=graduate, 0=dropout for non-criminal reason) on single program measures while controlling for relevant individual characteristics. Second, we examine the class-level program completion rates via linear regression, again, looking at one program characteristic at a time. Table 18 summarizes the findings from these analyses

Programmatic effects on individual-level program completion Coaches’ assessments of participants’ performance in the program were predictive of program completion. Individuals who reportedly participated in class, enjoyed the program, and understood the lessons were more likely to complete the program. Furthermore, participants’ who apply previously taught skills in later class sessions are more likely to graduate. The group atmosphere as perceived by the coaches influenced whether individuals finished the program; the more cohesive the group, the more likely participants were to graduate.

Characteristics of the coaches and classes as reported by observers of the program lend insight to programmatic qualities that facilitate participants completing the program. The organization and

professional demeanor of the coach impacted individuals' likelihood of finishing the program. This domain of the program was measured as a scale of seven items: (a) adherence to scheduled starting and ending times, (b) adherence to lesson plan for the day, (c) adequacy of preparation for the lesson, (d) organization of the lesson, (e) summary of the lesson at the end of the class, (f) level of cooperation between coaches, and (g) involvement of other coaches. When coaches met the observers' standards for class organization and professional demeanor, the probability of participants completing the program increased²⁶. Two additional aspects assessed by observers of the program that are not incorporated in the above scale affected class attrition. Enthusiastic coaches and classes with high levels of group participation increased the odds of participants graduating from treatment. Class size did not affect program completion.

Participants observation are not included in these analyses, because such data were only available for program graduates.

Programmatic effects on class-level program completion Similar program characteristics impacted class-level rates of program completion as affected individual-level program completion. Classes characterized by positive, cohesive group atmospheres (as reported by coaches) had higher rates of completion. When coaches were organized, professional, and enthusiastic, they were able to retain greater portions of their classes through graduation.

Table 18: Individual-level program completion and class-level rates of program completion by program characteristics: summary of several logistic and linear regression models.¹

Independent Variables	Individual-Level		Class-Level	
	B	Exp(B)	B	Beta
<u>Session evaluation data – reported by coaches²</u>				
Participants' average level of participation.	.59****	.55	NS	NS
Participants' average level of enjoyment.	.55***	.58	NS	NS
Participants' average level of understanding.	.54***	.58	NS	NS
Participants' average use of previously taught skills.	1.42*	.24	NS	NS
Group atmosphere.	.91****	.40	.01*	.26
<u>Observer evaluation data – reported by observers</u>				
Organization and professional demeanor (scale) ³	1.39***	.25	.29**	.39
Enthusiasm of coach (single item)	.971***	.38	.21**	.36
Level of group participation (single item)	.88**	.41	NS	NS

²⁶ The analyses controlled for who observed the classes.

**** $p \leq .001$; *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$; NS = not significant

¹ This table presents the parameter estimates of individual-level and class-level analyses predicting program completion. The individual-level logistic regression models regress “program completion” on program characteristics one at a time, while controlling for parolees’ risk level, prior conviction for violent crimes, IQ, reading level, high school graduation, marital status, socioeconomic status, race, age, and personality. The class-level linear regression models regress class rates of program completion on program characteristics one at a time.

² In the individual-level analyses, average participation, enjoyment, and understanding reflect individual participants’ average scores. In the class-level analyses, these same domains are measured as class averages.

³ The regression models examining the effects of the observer evaluation scale include a control for who observed the classes.

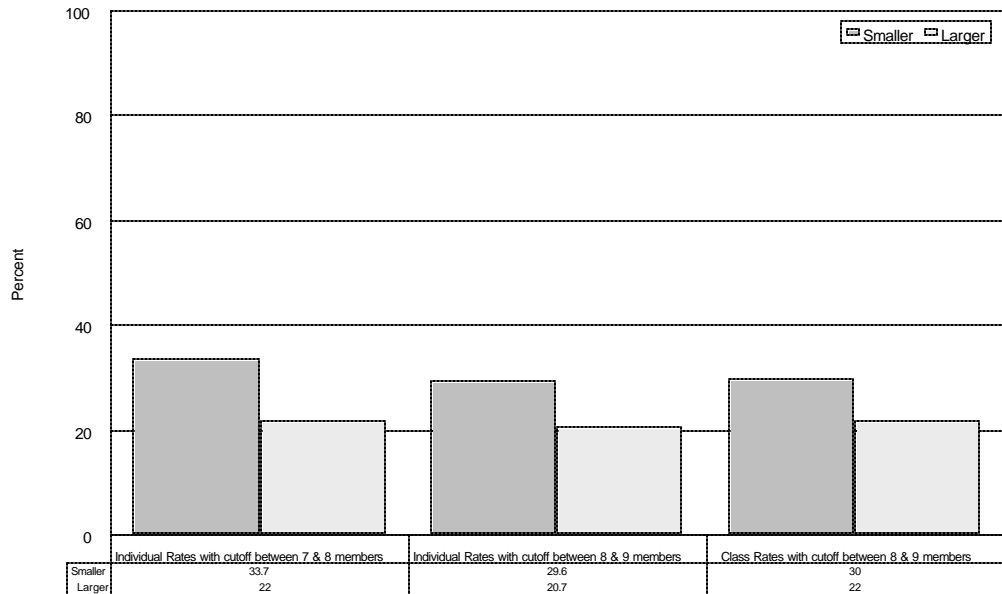
Effects of Class Size: Returns to Prison and Rearrests.

Ross, Fabiano, and Ross (1989) have suggested that R&R classes should range in size from four to eight members. The results of the individual-level and aggregate-level models predicting either felony arrest/revocation or returns to prison indicated that this standard did not translate to better individual or class performance. To determine whether there was an ideal size for Georgia’s R&R classes, we explored the impact of various class size cutoffs on rates of recidivism.²⁷ For both the individual model and the class model returning to prison was not significantly influenced by class size.

Rearrests/revocations were, however, affected by class size at both the individual and class levels of analyses. The effects of class size on rearrests are illustrated in Figure 15. Two cut-off points proved effective for individuals. First, individuals in classes with greater than seven members were less likely to be rearrested/revoked than those who graduated from smaller classes; 22% of graduates in classes of eight or more participants were rearrested, compared to 33.7% of graduates in classes of seven or fewer members. A class size cutoff between eight and nine members also worked; graduates in the larger classes had even fewer rearrests (20.7%) compared to those in the smaller classes (29.6%). The class size cutoff between eight and nine members also affected the felony arrest/revocation rates of classes in the class-level analysis. Classes with nine or more participants achieved a mean of 22% graduates being rearrested, compared to 30% for the smaller classes. The logistic and linear regression analyses are presented in Appendix J.

²⁷ Class size was measured as a ratio level variable and as several dummy variables with the following cutoff points: seven, eight, nine, ten, and eleven members.

Figure 15: Individual- and class-level rates of felony arrest/revocation by class size.



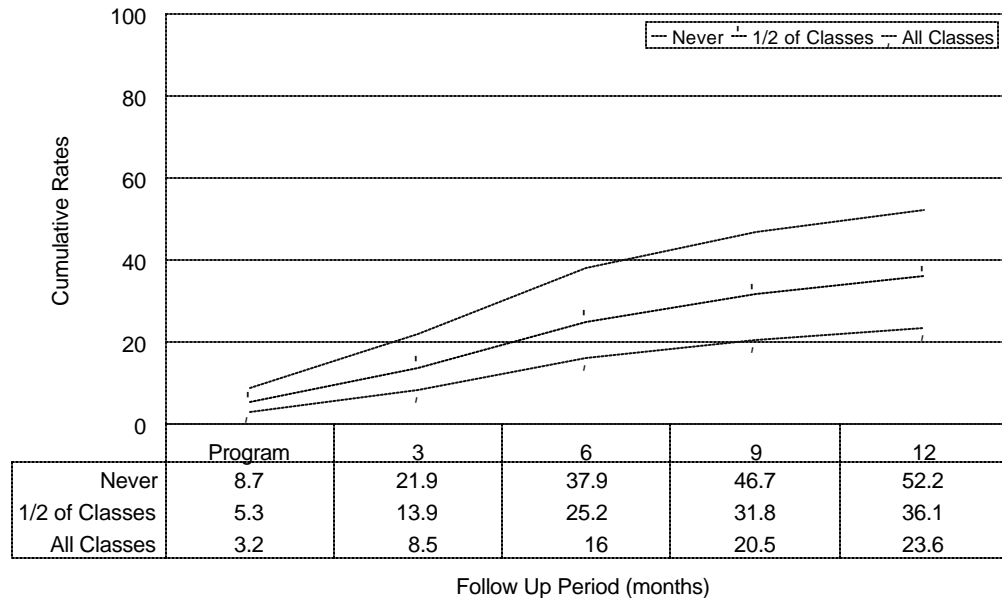
**Effects of Participant Performance as Assessed by Coaches:
Returns to Prison and Rearrests.**

Of the dimensions monitored by the cognitive coaches (e.g., participants’ level of enjoyment, level of understanding, level of participation, use of previously taught skills, and group atmosphere), only one, a measure of participants’ use of previously taught skills, was found to significantly impact recidivism. This factor related to graduates’ likelihood of having a felony arrest/revocation but not to returns to prison. Figure 16 presents the predicted probabilities of having a felony arrest/revocation when graduates used previously taught skills at various frequencies. Over half (52.2%) of those who never used previously taught skills were likely to be rearrested, compared to 36.1% of those using such skills during at least half of the sessions. Only 23.6% of the graduates using skills during all session experienced rearrests. The analysis controlled for relevant individual characteristics. Results of the logistic regression analysis are located in Appendix K.

Other measures of participants’ performance as assessed by coaches were unrelated to the graduates’ likelihood of committing a new offense. Class rates of rearrest and returns to prison were unaffected by session evaluation measures aggregated to the class unit. This finding was not surprising; we

would expect successful practice and skill development to affect individual outcomes far more than class outcomes.

Figure 16: Predicted graduate felony arrest/revocation rates by use of skills, time and control variables.



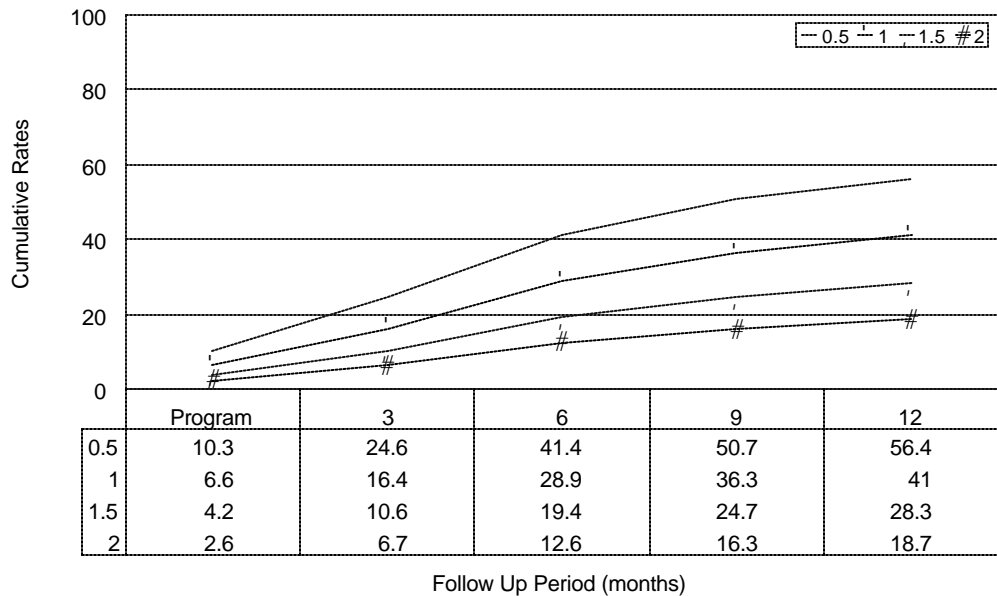
**Effects of Coach Performance as Assessed by Observers:
Returns to Prison and Rearrests.**

One of the three observer scales, coaches adherence to the prescribed structure for the class (e.g., established and enforced rules for missing groups, tardiness, and poor classroom behavior) was significantly related to the arrest/revocation follow up measures. Statistical models examining the effect of observation evaluation scales on individual-level recidivism included controls for who assessed the classes as well as relevant offender characteristics. The classroom structure scale is made up of six items: (a) evidence of rules and sanctions for missing group meetings, (b) evidence of enforcement of rules and sanctions for missing group meetings, (c) evidence of rules and sanctions for tardiness, (d) evidence of enforcement of rules and sanctions for tardiness, (e) evidence of rules and sanctions for poor classroom behavior, and (f) evidence of enforcement of rules and sanctions for poor classroom behavior. Each item was scored from 0 to 2, with “2” representing the most favorable assessments. Hence, the scores of “2” on

the classroom structure scale also reflect superior ratings. The analyses suggest that the presence of clearly stated and enforced standards for classroom behavior and participants' responsibilities in treatment is particularly important in maximizing the benefits of the program for participants. Graduates from program classes characterized by evidence of rules and sanctions and the enforcement of such rules were significantly less likely to experience a felony arrest/revocation. Figure 17 illustrates the relationship between the level of classroom structure and graduates' felony arrests. While only 18.7% of individuals in classes scoring a perfect "2" on the classroom structure scale recidivated, 28.3% of parolees in classes with scores of "1.5" had felony arrest/revocations.

The effect of classroom structure on the overall performance of the classes produced consistent findings. Aggregate class felony arrest/revocation rates lowered with more positive assessments of the class structure, while controlling for who observed the classes. Appendix L provides the logistic and linear regressions discussed here. None of the observer evaluation scales affected the likelihood of returns to prison.

Figure 17: Predicted graduate felony arrest/revocation rates by class structure¹, time and control variables.



¹ Classroom structure is measured on a three-point scale (0 thru 2), with "2" indicating the most positive assessments.

A few individual items not included in the classroom structure scale are related significantly to graduates' probability of recidivating. Our departure from the three scales to the analysis of specific items was conducted for exploratory purposes. Logistic regressions examined the relationships between individual observer evaluation items and individual-level felony arrest/revocation and returns to prison, while controlling for relevant individual characteristics. A separate model was run for each of the items because of (a) the high correlations ($r > .70$) among many of the individual items and (b) the loss of statistical power that results when too many items are included in a model. This approach, however, increases the likelihood of finding significant results due to the sheer number of analyses (45). The relationships between the individual observer evaluation items and graduates' recidivism are summarized in Table 19.

Graduates were less likely to have felony arrests/revocations if: (a) their coach displayed professional detachment (i.e., maintained good boundaries, and did not take matters personally); (b) most rather than some of the participants were engaged; and (c) the level of group participation was high. Coach characteristics related to instances of lower graduate returns to prison included: (a) the adequacy of coaches' preparation for lessons, (b) the level of cooperation between coaches, and (c) the involvement of both coaches. Obviously, the last two items pertain to program classes with dual coaches. Reports of high involvement of both coaches also favorably impacted the class-level rates returns to prison ($r = -.35$, $p = .04$).

Table 19: Graduate felony arrest/revocation and return to prison by Observer evaluation items, time, and control variables: summary of several logistic regression models.¹

Independent Variables	Felony Arrest/Revocation		Return to Prison	
	B	Exp(B)	B	Exp(B)
Professional detachment.	-.75*	.47	NS	
Engagement of most rather than some participants.	-.39*	.68	NS	
Level of group participation.	-.52*	.60	NS	
Adequacy of preparation for the lesson.	NS		-.79**	.46
Level of cooperation between the coaches.	NS		-.80**	.45
Involvement of both coaches.	NS		-.71**	.49

**** $p \leq .001$; *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$; NS = not significant

¹This table presents the parameter estimates of observer evaluation items that were significantly related either to felony arrest/revocation or returns to prison. The analyses involved logistic regression models that examined observer items one at a time, while controlling for parolees' risk level, prior conviction for violent crimes, IQ, reading level, high

school graduation, marital status, socioeconomic status, race, age, and personality.

**Effects of Programmatic Characteristics as Assessed by Participants:
Returns to Prison and Rearrests.**

Qualities of the program as assessed by participants formed three scales: adherence to social learning methods, relevance of skills to participants, and group climate. None of these composite scales affected our long-term outcome measures. Single items, however, were related to (a) graduates' likelihood of rearrest and/or return to prison and (b) class rearrest and/or return to prison rates.

Individual-level analyses. Table 20 presents the logistic regression models showing the effects of graduates' perceptions of the program (single items) on reoffending. Included in these models were controls for individual criminal history and demographic characteristics. Four items from the participant evaluation forms significantly related to graduates' likelihood of both felony arrest/revocation and returns to prison. The probability of recidivism increased when the skills taught in the program were too difficult for individuals and when just a few people seemed to do most of the talking. Unexpectedly, graduates were also more likely to recidivate when they perceived the skills and examples to be less realistic. The number of times a parolee meets with his supervising parole officer during the treatment period increased his probability of recidivism. An additional item is related only to the return to prison dependent variable. When a parolee perceived that his coach(es) understood where he was coming from, the likelihood of his returning to prison was significantly increased. With the exception of meetings with supervising parole officers, the positive coefficients in Table 20, of course, ran counter to our expectations. Increased supervision often leads to increased detection of new offenses.

Table 20: Graduate felony arrest/revocation and returns to prison by Participant evaluation items¹.

Independent Variables	Felony Arrest/ Revocation		Return to Prison	
	B	Exp(B)	B	Exp(B)
The skills and examples seemed pretty realistic.	.71***	2.04	.38**	1.47
The skills were too difficult.	-.39**	.68	-.34**	.71
Just a few people seemed to do most of the talking in the class.	-.36***	.70	-.29**	.75
I felt that the coaches understood where I was coming from.	.31	1.36	.61***	1.83
Not including the Cognitive Skills meetings, how many times per month did you meet with your supervising parole officer while you were in the Cognitive Skills program? (0 = 1 to 2; 1 = 3 or more)	1.08****	2.95	.60**	1.82
Time period ^a				
3 month	1.18**	3.27	1.15	3.15
6 month	1.80****	6.04	2.04***	7.69
9 month	1.30**	3.67	2.31***	10.10
12 month	1.09**	2.97	1.97**	7.16
15 month	----	----	1.83**	6.26
18 month	----	----	1.62*	5.05
21 month	----	----	2.71****	15.08
24 month	----	----	1.29	3.62
27 month	----	----	1.51	4.53
30 month	----	----	2.14**	8.53
33 month	----	----	2.49**	12.12
Risk level	.66*	1.93	.99***	2.70
Prior conviction for violent crimes	-.08	.93	-.02	.98
IQ	.00	1.00	.00	1.00
Reading level	-.01	.99	-.01	.99
High school graduate	.06	1.06	-.09	.91
Married	-.34	.71	.22	1.25
Middle class	.25	1.28	.25	1.28
White	-.67**	.52	-.71**	.49
Age 23 to 27	-.37	.69	-.25	.78
Neurotic	.14	1.15	.55*	1.73
Intercept	-6.21****	.00	-8.01****	.00
Model Chi Square	64.13****		67.96****	
-2 log likelihood	470.32		609.50	

**** p ≤ .001; *** p ≤ .01; ** p ≤ .05; * p ≤ .10; NS = not significant

¹ The participant evaluation items are scored so that higher values correspond with more favorable assessments of the program.

Class-Level analyses. The effects of participant evaluation items on class recidivism rates are summarized in Table 21. A class was more likely to have high rates of felony arrest/revocations when a substantial proportion of its members deemed the skills to be too difficult. Class arrest and return to prison rates both increased when the coaches were perceived by the class as understanding where the members were coming from. Several class characteristics affected class rates of returns to prison and not arrest rates. Classes composed of participants who at first hated going to class had lower return to prison rates than classes with participants who reported not experiencing such initial feelings. Classes with parolees who found the exercises helpful also experienced lower return to prison rates. When at least one of the coaches seemed by the class to take comments too personally the rates of returns to prison increased. Unexpectedly, classes with coaches characterized as enthusiastic about teaching the class had higher return to prison rates.

Table 21: Class felony arrest/revocation and returns to prison rates by Participant evaluation items¹

Independent Variables	Felony Arrest/ Revocation		Return to Prison	
	b	Beta	b	Beta
At first I hated going to class.	.02	.08	.08*	.24
The skills were too difficult.	-.84**	-.37	-.09	-.16
The exercises were helpful.	-.10	-.18	-.25***	-.39
The coaches seemed enthusiastic about teaching the class.	.06	.14	.16**	.34
I felt that the coaches understood where I was coming from.	.13*	.24	.26**	.41
At least one of the coaches seemed to take our comments too personally.	-.06	-.16	-.16**	-.33
Intercept	.77		.29	
F	3.50***		5.70***	
Adjusted R Square	.23		.36	

*** p ≤ .001; ** p ≤ .01; * p ≤ .05; * p ≤ .10; NS = not significant

¹ The participant evaluation items are scored so that higher values correspond with more favorable assessments of the program.

Effects of Programmatic Factors on Test Results for Pre-release Women.

A number of factors were associated with favorable test outcomes for pre-release women offenders. The most meaningful results were seen on COAS performance when we assessed participants' perceptions of the course. As shown in Table 22, participants who rated the class as operating from a social learning format, relevant to their lives, and taking place in a safe and sensitive climate, were significantly

more likely to achieve higher test scores by the end of the class. Observers' ratings showed similar benefits from a positive class atmosphere, adherence to structures, and coaches' ability to deal with difficult participants. Coaches' ratings were difficult to use because the coaches often did not discriminate between the items. Individuals tended to receive the same rating on each item; therefore intercorrelations were high and correlations between the rating and the test outcome were nearly identical for all ratings. A negative correlation between group atmosphere and COAS scores is largely attributable to the fact that the item varied from very good to excellent.

Table 22: Programmatic effects of test (PID & COAS) results, pre-release women.

Programmatic Characteristic	PID Change	COAS Sent	COAS Skill
Class size	.17	.14	.23
<u>Coaches' ratings</u>			
Participation	-.34*	.05	-.04
Enjoyment	-.33*	.00	-.02
Understanding	-.35*	.05	.09
Atmosphere	-.15	-.23	-.25*
Use of skills	.01	.11	.08
<u>Observer's ratings</u>			
Positive atmosphere	-.05	.41**	.31**
Organization, professional demeanor	.23	.26*	.04
Structure (rules)	-.23	.30*	.34**
Skill with difficult participants	-.23	.30*	.34**
<u>Participants' ratings.</u>			
Adherence to social learning methods	-.16	.34**	.30*
Relevance of lessons	-.22	.55***	.39**
Group climate	.03	.53***	.50***

Summary

At the outset of this study, it was not possible to anticipate the severe truncation of many of the programmatic variables. With the occurrence of such a problem, however, many measures were not ideally suited to these analyses.²⁸ We do note, however, that the following programmatic factors were found to be worthy of concern to program facilitators and planners:

Class completion: This seems like a rather obvious finding, however, outcomes were compromised for not only the participants whose completion was prevented by incurring a crime-related incident while in the program but for those who were released in order to pursue employment, childcare, and other seemingly prosocial pursuits.

Class size: Overly small classes, those containing fewer than seven/eight members, were not as successful as the larger classes. Large classes in this regard, were comprised of 10 members on average.

Class structure: Adherence to schedules and rules became an important indicator of successful classrooms.

Successful demonstration of learned skills differentiated successful individuals from unsuccessful ones. This finding not only tagged the behaviors themselves as important, but also the practice and participant engagement required to demonstrate the skills. Classrooms with higher levels of participation also produced more successful individual outcomes.

Detached coaches, those who could facilitate a class without becoming pulled into offender thinking or manipulation graduated parolees and classrooms with lower recidivism rates.

In classes with two coaches, their level of cooperation and the extent to which *both* were involved in the lesson, proved to be an important factor.

Offenders who perceived the classes as too difficult evidenced less favorable outcomes than those who understood the class, even when the analyses controlled for IQ and reading levels.

Finally, test outcomes for women offenders, the only outcome measure available at this time, were impacted by coaches adherence to the curriculum, the relevance of the class, and the group climate. As with parolees, these classes benefited from adherence to and enforcement of rules balanced with a generally favorable class atmosphere.

Other scales, such as adherence to social learning clinical dimension, classroom climate, interaction skills were not significantly related to offender recidivism. Limited variation, due to the fact that the overwhelming proportion of respondents rated these measures favorably, may be one reason for the failure to detect their impact on offender outcomes. In fact, the significant factors listed above were significant for two reasons. First, they really did impact offender success. Second, in some settings these

²⁸ Furthermore, experimentation with such corrective strategies as log transformations of the skewed variables failed to improve this situation.

issues were problematic. If a program shortcoming was not observed in many, or any, settings, it simply could not be studied for its impact on offender performance. Limited variation could also help to understand the few findings which ran contrary to expectation. We were often seeking to determine whether a perfect score on a scale versus an almost perfect score made the difference between a successful or unsuccessful outcome. Unfortunately, the worst problems with variation and the unexpected findings typically involved the participant evaluation forms.

Chapter 6

Discussion and Recommendations

This chapter synthesizes the study findings and discusses their importance to the Georgia Board of Pardons and Paroles as well as to the social science of offender programming. In doing so, we: (a) highlight the most important findings, (b) compare these results, where possible, to relevant studies in the field, (c) offer interpretations, and (d) discuss their importance in terms of future directions for research, policy, and programming.

At the outset, we should note that many things went well for this project, especially at the parole sites. First, GBPP administrators appeared to have adhered to the experimental design set forward in this research. This occurred during both phases of this project. Second, because the parole sample is large (N=1155), most statistical tests were powerful enough to support our analyses. Third, GBPP staff worked hard to facilitate the research. Data collection tasks were approached efficiently. Communication between GBPP and the University of Cincinnati was excellent, and problems were solved definitively and quickly. The Agency used evaluation research constructively. From our very first report, a process study for Phase I, GBPP personnel: (a) utilized the feedback provided by the findings, (b) invited University of Cincinnati research staff to discuss results with staff and administrators, and (c) implemented, where possible and feasible, programmatic changes suggested by the research. Finally, the data maintained in GBPP data basis (OTIS and FLOID) appeared to be consistent and accurate.

As noted in Chapter 2, the pre-release samples were small. Moreover, a comparison between experimental and comparison subjects according to social and criminal background characteristics found a number of differences between the groups. This may have been an artifact of the small size of these groups, however.

Notwithstanding the study's strengths, we have some concerns for the quality of the pretest, posttest data, as well as the participant evaluation forms. Interpretations of these data, especially for parolees and pre-release males offenders, often seemed like exercises in futility. Results were typically

either insignificant or unlikely. Moreover, tests of the validity of the PID and COAS found : (a) both instruments to be invalid among pre-release males; (b) both instruments to be valid for pre-release females; and (c) the PID to be invalid among parolees, and (d) only a few of the COS scales to be valid among parolees. Results for pre-release women, however, were considerably more meaningful. In the recommendations section, we offer suggestions for appropriate test conditions. We turn now to a discussion of the major findings of this study.

Addressing the Research Questions of the Georgia Cognitive Skills Experiment (Phase II).

In the introductory chapter of this report, we outlined the three research questions guiding this project. We were able to address these with varying degrees of success. We hope that the resulting findings will prove particularly meaningful to corrections research as well as to the Georgia Board of Pardons and Paroles. Results are discussed according to the three questions posed at the beginning of the study.

Question 1:

What was the overall effectiveness of the Georgia Cognitive Skills Experiment? How did recidivism rates for parolees participating in the Cognitive Skills Program compare to those who were assigned to the comparison group? How did the groups perform on the intermediate measures of outcome?

For parolees, event history analysis detected a 3.3% difference between experimental and comparison group returns to prison by 30 months following the program. Differences in arrest/revocation rates between the two groups at the end of 12 months was 2.5%. Neither difference was significant ($p \leq .05$). These modest effects are consistent with those found for Phase I of the project, and they remain insignificant even when the Phase I and Phase II samples are combined.

These results, along with those for Phase I, are in line with other large-scale R&R initiatives, which include two experimental studies, one in Canada and another in Colorado. The Canadian study achieved a treatment effect of 5%, among parolees (Robinson, 1995), and the Colorado study found an effect of 3% (Johnson & Hunter, 1992).

This trend has already come to the attention of meta-analysis researchers; two recently published studies report that large cognitive programs for offenders, of which R&R is only one model, achieve treatment effects between 3-5% (see Wilson et al., 2000; Lipsey et al., 2001). The authors further remind us that an earlier small-scale program achieved effects as large as 51% (Ross, Fabiano, & Ewles, 1998). The concern, voiced by many, is that the larger programs may have a more difficult time addressing and maintaining treatment integrity. In other words, they may be more likely to depart from appropriate screening and curriculum protocols.

We also examined the intermediate outcomes for the pre-release and parole samples. As noted above, we had reservations about the validity of pretest and posttest measures for parolees and pre-release males, so they do not receive unwarranted discussion at this point. COAS results for pre-release females, however, were significantly better for e's than c's on four of the ten scales studied. PID pretest, posttest changes were not significantly different between the two groups of women; however, the null findings were primarily attributable to an interaction between risk and treatment assignment.

Long-term outcomes for pre-release groups are not presented in this report, because our follow-up period did not allow a sufficient window of time in the community following release. However, we wish to pursue the matter in an addendum to this report.

Question 2:

Did the program benefit some types of offenders more than others? Did some types of program participants actually do worse than those who did not participate?

Programs usually work better for some types of offenders than for others. In this light, general treatment effects, such as those reported for treatment groups as a whole, often “mask” the larger treatment effects achieved by some types of offenders. This is because the successes of those who are most amenable to the program are cancelled-out by the failures of those who are not. This notion of differential effectiveness, or *responsivity*, is frequently referenced in the corrections literature (Andrews & Bonta, 1998; Bonta, 1995; Gendreau & Ross, 1987; Kennedy, 2000; Palmer, 1974, 2002; Van Voorhis, 1997; Warren, 1983). However, critics of the state of corrections research note that evaluation research almost

never addresses the matter. Even so, studies which have sought to do so, have typically produced extremely important findings (Gendreau & Ross, 1987; Palmer, 1981; Van Voorhis, 1987).

The Georgia Cognitive Skills Experiment identified important areas of offender responsivity. The parolees who achieved the most impressive treatment gains were: (a) white, and (b) between the ages of 28 and 32. Neurotic participants, as classified by the Jesness Inventory (Jesness, 1996), had considerably higher recidivism rates than the comparison group neurotics who did not participate. Responsivity characteristics of women offenders followed the established screening recommendations for R&R: (a) Largest treatment effects were achieved by offenders with IQ scores above 85; (b) low risk experimental offenders achieved worse posttest scores than low risk comparison group members. In addition, a race effect was detected for women as well as parolees; white offenders had much better outcomes than nonwhite offenders.

Race as a responsivity consideration. Even as the fields of mental health and correctional treatment begin to embrace the notion of multicultural counseling (e.g., Baruth & Manning, 1998), examining evaluation results for racial differences is not a common research task. A review of the evaluation research in corrections quickly makes evident that there is little to say about what works for African American offenders, or, for that matter, what works for other ethnic and racial groups.

It is important to note that our study did not find that African American offenders were *harmed* by the Cognitive Skills Program. In both cases, African American participants performed better than African American comparison group members. However, whites in the program achieved much greater treatment effects than African Americans; this was true for both the male parolees and the female pre-release inmates. In the first large-scale study of R&R, Robinson (1995) found relevant results when he differentiated between white and aboriginal Canadian parolees. Aboriginal offenders gained little from R&R.

At this point in time, we are not able to identify the reasons for this finding; any of a number of possibilities are posed. However, while the study was underway, personnel revised some instructional tools to include more examples of African American offenders. Another avenue of inquiry concerns whether African Americans had sufficient representation and visibility among the coaches serving the program. University of Cincinnati staff are currently securing data pertaining to the race of program coaches and will

analyze its impact in future analyses of these data.

Additional avenues of inquiry could cause us to examine numerous aspects of the Southern culture in which the study is imbedded as well as the possibility that race simply hasn't been factored into program planning and staff training for this program. Finally, with such a dearth of research on racial responsiveness, we cannot ignore the possibility that the problem simply may not be unique to cognitive programming in Georgia. There is no way of knowing whether similar findings might exist for other programs as well.

Personality as a responsiveness factor. The present study observed that neurotic offenders performed poorly in the cognitive skills program. In fact, the findings suggest that they would have been better off not being in the program at all. Neurotic parolees were also more likely than non-neurotic parolees to fail to complete the cognitive skills program. This finding is important for a number of reasons. Most importantly, this group of offenders is beginning to appear in a small body of research as a group who may need to be targeted for a distinct type of correctional programming. This small number of studies note the following: (a) neurotics experience much higher levels of depression and stress in prison settings than other inmates (Van Voorhis, 1994); (b) they benefit from programming targeted to the dysfunctional, criminogenic, aspects of their anxiety (Warren, 1983; Palmer, 1974; and (c) they perform poorly in treatments which involve too much confrontation (Warren, 1983; Jesness, 1971; Palmer, 2001). In a recent study examining the long-term recidivism of these offenders, Johnson Listwan (2001) reported that neurotic offenders had higher recidivism rates than non-neurotic offenders. The latter, suggests, of course that high anxiety may constitute a risk factor as well as a responsiveness factor. This is also consistent with findings pertaining to the crime-enhancing potential of "negative emotionality" Caspi, Moffit, Silva, Stouthamer-Loeber, Krueger & Schmutte, 1994; Caspi & Silva, 1995).

We cannot determine precisely why these offenders did so much worse than the other program participants. Some coaches and offenders may have been too confrontative. The curriculum may have involved too much personal disclosure, thus triggering anxiety in a group of offenders whose criminal and acting-out behavior is primarily the result of anxiety. Another possibility is that neurotic personality characteristics are associated with a tendency toward self-destruction. That is "good things" whether they involve relationships, growth enhancing programs, employment opportunities, or any number of other

advantages, are never too good to be immune from sabotage (Watson & Clark, 1984; Bolger & Zuckerman, 1995)

Age as a responsivity factor. Among parolees, age was related to dropping out of the program and to program success or failure. That is, program drop outs, regardless of their reason for dropping out were significantly more likely to be between the ages of 23 and 27, than in any of the other age categories considered in this research. At the same time, offenders in this age group, as well as those 33 and older did not appear to benefit from the program. Recidivism rates for these individuals were similar to those for the comparison group. The most successful parole participants were between the ages of 28 and 32.

We do not wish to speculate as to why this occurred. However, the finding of a curvilinear relationship between age and success occurred in the large Canadian study where the most successful offenders were between the ages of 25 and 39 (Robinson, 1995). Johnson and Hunter's (1992) research in Colorado found that young offenders, 29 or younger, were less successful than older offenders, aged 30 and older. In other words, Johnson and Hunter did not observe the curvilinear effect that was found in the other two studies.

Gender as a responsivity factor. Determination of whether the Georgia Cognitive Skills program worked for women would be premature at this time, because we do not have long term recidivism data for the pre-release offenders. However, women program participants achieved significantly better posttest scores on the COAS than women assigned to the comparison group. In addition, medium to high risk offenders achieved significantly improved pretest, posttest changes in PID scores. Low risk women offenders who were in the program appeared to score worse on the PID than those who were not.

Tests for responsivity factors found that low risk women, with IQ scores lower than 85, achieved less post-treatment gains than medium to high risk women and women whose IQ scores surpassed 85. The results also suggested ($p \leq .10$) that the program was more successful among women of lower socioeconomic standing and less than a high school education. High school graduates in the program did worse than high school graduates in the comparison group, and middle class participants showed minimal improvements over middle class comparison group members. As with the parolees, most of the treatment gains on posttest scores were achieved by white women.

Therefore, while we do not have the recidivism data to fully answer the issue of gender-responsivity (see Chesney-Lind, 2000), our findings offer some preliminary reasons for optimism. Even so, there were several responsivity issues to consider.

Risk as a treatment amenability factor. Risk was studied in Phase I of the Georgia Cognitive Skills Experiment and again in Phase II. We constructed a risk assessment instrument from static risk factors available through OTIS, and it showed impressive predictive validity ($r = .35, p \leq .001$). Regardless of our cut-off points for differentiating low, medium, and high risk, we did not observe a risk effect (i.e., a finding that medium to high risk offenders had greater treatment outcomes than low risk offenders) among male parolees. Moreover, we did not observe a risk effect when we compared low risk e's and c's to high risk e's and c's in an effort to more clearly distinguish the risk groups.

The lack of a risk effect has now been observed in three studies involving parolees: (a) the Phase I study (Van Voorhis et al., 2001), (b) the Canadian study (Robinson, 1995), and (c) the present study. This may be because low risk does not mean the same among parolees as it does among juveniles or community correctional clients. With respect to the Georgia Cognitive Skills studies, it may also be an artifact of the program's screening process. As parole officers screened their "most problematic" offenders into the program, they may have correctly screened out many low risk offenders even without the aid of a risk assessment tool.

Question 3:

What sorts of program and facilitator characteristics differentiated the most effective cognitive programs from the least effective? Do such factors as group size, program completion, attendance, and facilitators' skills impact program completion, test results, and offender recidivism?

Chapter 5 identified a number of instances where outcomes for program participants were much better for some types of programs than others. Even so, this was a difficult research question to address. Doing so requested additional work from busy GBPP personnel and exposed facilitators to a level of scrutiny that they may not have welcomed. It also required GBPP personnel to evaluate their colleagues; a task which some may have created considerable discomfort. Other program measures were sought from

the participants themselves and may not have been obtained under optimal conditions.

These difficulties did not preclude the identification of several important program characteristics. In this rather new climate of program evaluation, however, most programs were favorably evaluated, and many measures did not show variability sufficient to support statistical tests. On some measures, we appeared to be studying the differences an excellent class and a very good class. At the same time, it was not clear that this research problem was necessarily a *problem* for GBPP. Another explanation for this situation could be that the programs really were well-executed. Coaches, after all, had received training or retraining only a month prior to the beginning of Phase II.

Adherence to the program design. Against this background, we are not convinced that the Georgia programs had difficulties with “program drift”, which is the central concern raised with regard to correctional treatment integrity (e.g., Palmer, 1978, 1992; Van Voorhis et al, 1995). Participants told us that, in most instances, they had opportunities to: (a) observe a skill demonstrated, (b) to practice the skill, and (c) to receive feedback regarding how well the skill was performed. They told us, in other words, that the classes employed the clinical dimensions dictated by their underlying cognitive-behavioral theory. Observers gave high ratings on program qualities pertaining to adherence to the curriculum. When their ratings dropped even modestly, however, the reduced scores were significantly associated with lower completion rates.

Relationship qualities and relevance of the material. Both participants and observers rated coaches high in qualities such as empathy and sensitivity of participants. Such findings indicated that the coaches demonstrated high relationship qualities. Without such qualities, even programs that are delivered in a technically correct manner may fail to show a treatment effect (Andrews, 1980). Finally, offender ratings indicated that the material was highly relevant to their lives.

Classroom Structure and Control. Difficulties clustered around the coaches’ ability to keep good discipline and structure within the classroom. Classes that did not maintain and enforce rules regarding tardiness, absences, and unruly classroom behavior were considerably less successful than those that did. In other words, classroom control was an area where evaluations showed the lowest score (Spruance et al., 2001) and one that affected offender outcomes. Perhaps this is not surprising. For the most part, persons

chosen to facilitate correctional programs are not academically trained clinicians or educators. They are lay persons who receive a maximum of two weeks of training and then proceed to lead/teach some of the most difficult students any facilitator or teacher can be asked to instruct. This research also speaks to the notion of “professional detachment” a quality which allows correctional workers to be empathic and approachable, but discourages tendencies to be manipulated by offenders or pulled into offender thinking (thereby modeling criminogenic thinking). Professional detachment was a significant correlate of offender recidivism in the present study; leaders with such skills tended to have more successful classes.

Class Completion. Related to the issue of classroom control, some offenders appeared to have been dismissed from class too readily. We are not speaking here of those who committed a new offense during the program, but rather those who were dismissed to pursue employment opportunities or presented other, non criminal reasons to excuse their participation. Presumably program staff had some discretion over whether these individuals could be dismissed from the class. The offenders dismissed by coaches for non-criminal reasons showed recidivism rates that paralleled those for the comparison group. The recidivism rate for the program graduates, however, was quite low ---24.8% for rearrests and 28.8% for returns to prison. These findings recommend for mandatory rather than voluntary participation.

Some will fault our need to depart from the experimental design in order to address this issue, because separation of the experimental group according to dropout and graduate status rendered the groups incomparable to each other and to the comparison group. Although it is true that the groups were no longer as comparable as they are in the random assignment model, we took precautions to make as many corrections for the fact as possible, by: (a) statistically controlling for demographic and risk characteristics, (b) assessing factors which differentiated program dropouts from program graduates, and (c) not suggesting in any way that these results supplant those presented in Chapter 3 for the randomized control group design.

Class size: The literature on group dynamics puts forward the well established assertion that the ideal size for groups targeting behavioral change should be between eight and 12. According to this body of research, groups larger than 12 tend to incur less participation than smaller groups, and groups smaller than eight can be anxiety provoking, because participants may feel too much pressure to perform. The

manual for the Cognitive Skills Program recommends between four and eight participants. However, our study found that groups smaller than eight did not achieve the treatment gains that groups of eight or larger did.

This should not be interpreted to mean that overly large groups are optimal. The “large” groups in this study averaged 10 members. Even though the Georgia class sizes ranged from five to 21, only 3 groups contained 15 or more members. Therefore, we are not asserting that groups above 15 work.

Practice and offender participation. The final series of significant program characteristics clustered around the extent to which participants were engaged in the program. Participants who were characterized by coaches as unable to demonstrate previously taught skills, and who, themselves, reported that the skills were too difficult were more likely to recidivate than those who could demonstrate and understand the skills. Classes that were characterized by participants as having more rather than less discussions were also more likely to be successful. It didn’t matter who made this observation, coaches, observers, or participants. Classes with more participation were more likely to reduce recidivism.

Classes high in participation, enjoyment, level of understanding, and use of previously taught skill also experienced less attrition. This occurred whether participation was rated by the coach or the observer.

Perhaps this concern also relates to the participant’s ability to practice skills in an active and engaged manner. We can assert, however, that the effects of participation were independent of IQ and reading level, because our analysis controlled for these measures of ability. Therefore, this pattern of findings begins to implicate the amount of time devoted to practice and processing of skills.

The final program characteristic differentiating between success and failure concerned how well coaches who team-taught worked together. Successful programs had coaches who in the estimation of outside observers, cooperated with each other and demonstrated similarly high levels of involvement in classroom activities. At the same time, enthusiastic coaches achieved less class attrition.

At the beginning of these analyses, we thought that it would be important to examine the programmatic variables from both an individual and an aggregate, classroom level. Thus, the analyses allowed for the possibility that individual perceptions might be different from the effects of aggregate ecological factors. In fact, however, both types of models produced consistent results.

Most of the findings reported in this section were observed when the outcome variable was rearrest/revocation, or program completion and did not involve returns to prison. There are few explanations for this other than the possibility that the rearrest/revocation and the program completion models allowed for more degrees of freedom or statistical power, because they contained fewer independent measures pertaining to time periods. In fact, the analyses of program factors involved a smaller sample than other portions of the analysis. Another explanation considers the fact that the arrest/revocation described a shorter time period than the returns to prison data, 12 as opposed to 30. Program attrition occurred over an even shorter time frame. Therefore, it may be that the programmatic factors had short term rather than long term influences. This is highly speculative, however.

Recommendations

3. Plan for important differences among offenders.

“Responsivity,” the consideration of different learning styles among offenders, is a well-touted but mostly undeveloped notion in correctional treatment. Even so, the discovery of an individual attribute that suggests that a certain type of individual (e.g., low risk) is less appropriate for a program than another (e.g., medium/high risk) has two program implications. First, the program could use such findings to screen nonamenable individuals from program participation. Second, they could identify ways to make the program more appropriate to such individuals.

To decide upon a course of action, it might be helpful to consider that our analyses in Chapter 4 detected two types of interaction patterns. The first pertained to participants who performed worse than those offenders in the comparison groups. Among parolees this occurred primarily with neurotic offenders. Low risk participants had slightly higher rates of technical violations than low risk comparisons. Additionally, women participants with low IQ scores less than 85, and those classified as low risk performed worse on outcome tests scores than similar comparison group members. These findings make the strongest case for being exclusion criteria. Some authorities in the field have referred to such situations as instances where the program “harms” a group of participants.

With the exception of the neurotic correlate, all of the other criteria which pointed to those who

may have been harmed by the program have been detected in previous research. Because high anxiety as a responsivity characteristic is new to this study, sources may wish to achieve a more thorough understanding of the reasons for the reduced performance prior to going to the extreme of excluding such participants. For now, we conclude with several warnings: (a) neurotics performed considerably worse than offenders classified into other offender personality types; (b) these offenders are high risk, (c) they do not get better by themselves (Warren, 1983), and (d) most agencies are doing nothing to program for them.

The second type of responsivity-related finding involved instances where the less amendable group performed better than similar offenders in the comparison group but not nearly as well as the more amendable offenders in the experimental group. This was found to be the case with race. Among both women and men, experimental white participants performed much better than white comparison group members. Treatment gains for non-white participants were minimal in comparison. Because nonwhites were not harmed by participation in the program, race-based screening criteria would appear to be unwise. However, questions loom about how to make the sessions more responsive to nonwhite offenders. The findings could warrant examination of: (a) instructional aids such as vignettes, videos, visuals, and role playing situations, or (b) whether there are a sufficient number of African American coaches.

In addition to race, it might be important to consider that there may also have been a class effect among women offenders. The statistical relationship was modest, however, lower class women and those with less than a high school education appeared to have performed better than middle class participants and high school graduates.

Finally, offenders older than 29 as well as those younger than 28 did not benefit from the program. These findings raise questions concerning the maturity level of the material for older participants. Among younger offenders we could inquire about classroom control strategies. Both groups raise questions concerning motivation levels.

In sum, the following recommendations emerge from these findings:

Some offenders should continue to be screened from the program e.g., risk, IQ.

Some factors should be evaluated as possible screening criteria: e.g. neurotic classifications.

Some factors should entail an examination of the curriculum and the facilitation process for ways to improve responsivity: e.g., race, age, neurotic classifications.

4. *The findings continue to recommend against voluntary participation or dismissal for employment or other non-criminal reasons.*

When an offender ceases to participate because he or she has been arrested or revoked, there may be nothing coaches or parole officers can do to keep the offender in the class. However, dismissals for employment, unruly classroom behavior, childcare, transportation, and other non-criminal reasons did not work. As with the Phase I study, such individuals had higher recidivism rates than those who graduated. In fact, findings for the present study find this group had a recidivism rate comparable to the comparison group. A more effective course of action might be to facilitate childcare and transportation needs and to require offenders to schedule work around their supervision and program requirements.

In attempting to reduce attrition it would also be important to recall that certain types of individuals were more likely to drop out than others, including: (a) neurotics, (b) high/medium risk offenders, (c) lower class individuals, and (d) those between the ages of 23 and 27. Moreover, classes incurring the most attrition were those which; (a) were less diligent in adhering to the curriculum, (b) had less enthusiastic coaches, and (c) showed lower levels of classroom participation.

5. *Maintain class sizes at no less than 8 participants.*

Small classes did not achieve optimal outcomes among parolees. However, classes containing 8 to 12 members may be small enough to provide opportunities for interaction, practice, and full participation. They provide more role models and a greater diversity of role models. At the same time, small classes provide participation opportunities only at the cost of participants feeling constantly “on stage.” Moreover, if small groups contain reluctant participants or poor role models there is no fall-back to other class members. Our findings also underscore the fact that coaches are not the only vehicle for change.

4. *Participation and practice to the point of skill mastery appear to be important components of group activity.*

The importance of practice, feedback, and reinforcement for mastery of skills is deeply embedded in cognitive behavioral and social learning theory (Bandura, 1977; Meichenbaum & Jaremko, 1982), the theoretical models underlying the Georgia Cognitive Skills approach. Our findings further advocate for group activity, discussion, role plays, and practice dealing with real-life situations. Groups characterized by more facilitator talk and less active involvement from group members are less likely to be successful as those groups which promote active learning. Individuals experiencing difficulties with the skills or those who are reluctant to become engaged in group work should receive additional attention rather than be allowed to sit on the sidelines. These results not only encourage the provision of group activities but the monitoring of participants for their level of understanding.

8. *Identify ways to help coaches who are having difficulties with classroom control.*

Classroom control was the dimension most likely to be negatively rated by outside observers. Even though observers ratings were not overly negative, a modest level of difficulty decreased program effectiveness. These findings recommend for: (a) ongoing attention to classroom control in GBPP trainings, (b) monitoring of groups by personnel with higher levels of expertise and experience, and (c) in-service assistance to those coaches who are having difficulties. Moreover, classroom control should place a special emphases on expectations, rules, and enforcement of rules. Classes which did not present and enforce rules and expectations, where boundaries between coach and participant may have become blurred, were considerably less effective in promoting offender change than those which attended to these issues.

9. *Ongoing attention to quality assurance and prevention of “program drift” should involve additional training of program observers or clinical supervisors.*

This evaluation was one of the first to utilize observer measures of group quality and adherence to program designs. In doing so, we might have been a bit naive about the skills and mind-sets required of observers. As noted in earlier sections of this report, many classes were rated perfect on all items. While

this may have been a true reflection of the class quality, it may also question observers' ability to identify specific classroom/coach attributes or discriminate between effectiveness and ineffectiveness.

Additionally, we may have observed a reluctance to rate one's colleagues in a negative manner.

Monitoring or clinical supervision should be an ongoing programmatic task, one that occurs even when program evaluations are not in process. GBPP would be wise, in other words, to continue the observation process. However, ongoing training of supervisors and observers is essential. Such training should deal with the precision and characteristics of skills required of coaches and classes, as well as their feelings toward evaluating colleagues.

10. Pretest, posttest, participant evaluations, and observer evaluations will require more attention to the integrity of test conditions.

With the exception of the Pride and Delinquency Scale, the Culture Fair IQ test and the Jesness Inventory, paper and pencil tests administered to offenders throughout the study were relatively untested. The Colorado Attitude Survey had only been used once in a previous study (Johnson & Hunter, 1992), and the participant and observer evaluation forms were developed at the University of Cincinnati for use in this study. Since the beginning of Phase II, however, we have learned that two different assessments would be especially useful for pretest and posttest purposes. These are Criminal Sentiments Scale (Simourd, 1997) and the Psychological Inventory of Criminal Thinking Styles (Walters, 2002). Both appear to have performed better in other studies than our tests in the present research. Therefore, we recommend their use in future testings. Additionally, participant and observer forms are likely to undergo some modifications in order to better measure pertinent program dimensions.

We would be even more critical of the tests used in the present study, had they not worked exceptionally well in *one* of the program sites. The fact that PID, the COAS, and the Participant Evaluation forms proved to be extremely meaningful among women pre-release offenders unavoidably poses some

questions regarding the integrity of these data in other settings. Tests must be administered under carefully controlled conditions. When they are not, they are likely to be invalid.

Some of these conditions include the following:

- There should be no discussion during tests;
- They should be administered under supervision rather than given to offenders to complete at home or in the cell-block;
- Without compromising respondents' confidentiality, test supervisors should scan for situations where offenders might be rapidly checking off the same response to every item;
- Procedures should ensure that offenders know their results are only going to be viewed in aggregate and will not affect release dates, violation proceedings or other case outcomes.
- If tests can be administered by an individual who is not an employee, e.g., an intern, results will be more valid.
- Participant evaluation forms should not be administered at graduation, but rather during the next to last week of class.
- Test conditions should be uniform across groups.

Such precautions are likely to be worth the trouble. Even without an ongoing evaluation, staff, especially coaches, often benefit from some level of feedback regarding the effectiveness of their work. Thus, the value of pretests, posttest, and course evaluations is likely to be ongoing. However, our findings suggest a need for more serious attention to the appropriate and inappropriate ways to administer such instruments.

Conclusion

In closing, useful knowledge appears to have been generated by this research, and we hope that these findings prove meaningful to correctional officials throughout the State of Georgia. It is not common for agencies to commit to the task of identifying ways to optimize the success of their programming endeavors, but this intensive four-year effort did, in fact, produce meaningful results.

References

- Andrews, D. (1980). Some experimental investigations of the principles of differential association through deliberate manipulations of the structure of service systems. American Sociological Review, 45, 448-462.
- Andrews, D. & Bonta, J. (1995). Level of Service Inventory-Revised. North Talawanda, NY: Multi-Health Systems.
- Andrews, D. & Bonta, J. (1998). The psychology of criminal conduct. Cincinnati, OH: Anderson Press.
- Andrews, D., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation: Rediscovering psychology. Criminal Justice and Behavior, 17, 19-52.
- Andrews, D. & Friesen, W. (1987). Assessments of anticriminal plans and the prediction of criminal futures: A research note. Criminal Justice and Behavior, 14, 33-37.
- Andrews, D. & Keissling, J. (1980). Program structure and effective correctional practices: A summary of CaVIC probationers: The effects of selection of officers, supervision by volunteer and supervision practices on recidivism Final report submitted to the Solicitor General Canada, Ottawa, Ontario.
- Andrews, D. Zinger, I., Hoge, R., Bonta, J., Gendreau, P., & Cullen, F. (1990). Does correctional treatment work? A psychologically informed meta-analysis. Criminology, 28, 369-404.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 94, 191-215.
- Baruth, L. & Manning, L. (1998). Multicultural Counseling and Psychotherapy: A Lifespan Perspective. New York, NY: Merrill.
- Bloom, B. & Owen, B. (2002). Gender-responsive strategies: Research, practice, and guiding principles. Washington, D.C.: National Institute of Corrections.
- Bolger, N. & Zuckerman, A. (1995). A framework for studying personality in the stress process. Journal of Personality and Social Psychology, 69, 890-902.
- Bonta, J. (1995). The responsivity principle and offender rehabilitation. Forum on Corrections Research, 7, 34-37.
- Buehler, R., Patterson, G. & Furniss, J. (1966) The reinforcement of behavior institutional settings. Behavioral Research and Therapy, 4, 157-167.
- Caspi, A., Moffit, T., Silva, P., Stouthamer-Loeber, M Krueger, R. & Schmutte, P. (1994). Are some people crime -prone? Replications of the personality crime relationship across countries, genders, races, and methods. Criminology, 32, 163-195.
- Caspi, A. & Silva, P. (1995). Temperamental qualities at age three predict personality traits in young adulthood: Longitudinal evidence from a birth cohort. Child Development, 66, 486-498,
- Cattell, R. & Cattell, A. (1973). Measuring intelligence with the Culture Fair Tests. Institute for Personality and Ability Testing, Inc.
- Chesney-Lind, M (2000). What to do about girls? Perspectives on evaluation of programs for female offenders in community corrections. In M. McMahon (Ed.), Assessment to assistance: Programs for women in community corrections. Lanham, MD: A American Correctional Association.

- Covington, S. (2000). Helping women to recover: Creating gender-specific treatment for substance abusing women and girls in community corrections. In M. McMahon (Ed.), Assessment to assistance: Programs for women in community corrections. Lanham, MD: American Correctional Association.
- Dodson, K. & Khatri, N. (2000). Cognitive therapy: Looking backward, looking forward. Journal of Clinical Psychology, 56, 907-923.
- Dowden, C. & Andrews, D. (1999). What works for female offenders: A meta-analytic review. Crime and Delinquency, 45, 438-452.
- Empey, L. & Erickson, M. (1972). The Provo Experiment: Evaluating community control of delinquency. Lexington, MA: D.C. Heath.
- Fagan, J. (1990). Treatment and reintegration of violent delinquents: Experimental results. Justice Quarterly, 7, 233-263.
- Fulton, B., Latessa, E., Stichman, A. & Travis, L. (1997). "Up to speed"—The state of ISP research and policy implications. Federal Probation, 61, 65-75.
- Gendreau, P. & Andrews, D. (1994). The Correctional Program Assessment Inventory, 4th ed., St. John, New Brunswick: University of New Brunswick.
- Gendreau, P. & Ross, R. (1987). Revivification of rehabilitation: Evidence from the 1980s. Justice Quarterly, 4, 349-408.
- Goldstein, A. & Glick, B. (1987). Aggression Replacement Training. Champaign, IL: Research Press.
- Grant, D. (1965). Delinquency treatment in an institutional setting. In H.C. Quay (Ed.), Juvenile delinquency: Research and theory. Princeton, NJ: Van Nostrand.
- Hair, J., Anderson, R., Tatham, R., & Black, W. (1992). Multivariate data analysis. (3rd ed.). New York: Macmillan.
- Heide, K. (1982). Classification of offenders ordered to make restitution by Interpersonal Maturity Level and by specific personality dimensions. Dissertation Abstracts International. University Microfiles. Albany, NY: State University of New York at Albany.
- Heide, K. (1983). An empirical assessment of the value of utilizing personality data in restitution outcome prediction. In W. Laufer and J. Day (Eds.), Personality theory, moral development, and criminal behavior. Lexington, MA: Lexington Books.
- Hoffman, P. (1994). Twenty years of operational use of a risk prediction instrument: The United States Parole Commission's Salient Factor Score. Journal of Criminal Justice, 22, 477-494.
- Holsinger, A. (1999). Opening the 'Black Box': Assessing the relationship between program integrity and recidivism. Dissertation Abstracts International. University Microfiles. Cincinnati, OH: University of Cincinnati.
- Jesness, C. (1971). The Preston Typology Study: An experiment with differential treatment in an institution. Journal of Research in Crime and Delinquency, 8, 38-52.
- Jesness, C. (1992). The Jesness Inventory (Form G). North Tonawanda, NY : Multi-Health Systems, Inc.
- Jesness Inventory Computer Program Windows™ Version (1996) North Tonawanda, NY: Multi-Health Systems, Inc.

- Jesness, C. F., & Wedge, R. (1983). Classifying offenders: The Jesness Inventory Classification System. Sacramento, CA: California Youth Authority.
- Johnson, G. & Hunter, R. (1992). Evaluation of the Specialized Drug Offender Program. Boulder, CO: University of Colorado, Center for Action Research.
- Johnson Listwan, S. (2001). Personality and criminal behavior: Reconsidering the individual. Dissertation Abstracts International. University Microfiles. Cincinnati, OH: University of Cincinnati.
- Kennedy, S. (2000). Treatment responsivity: Reducing recidivism by enhancing treatment effectiveness. Forum on Corrections Research, 12, 19-23.
- Lipsey, M. (1989). The efficacy of intervention for juvenile delinquency: Results from 400 studies. Paper presented at the 41st annual meeting of the American Society of Criminology, Reno, NV.
- Lipsey, M. (1995). What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In J. McGuire (Ed.), What works: Reducing reoffending. Chichester, England: Wiley.
- Lipsey, M., Chapman, G., & Landenberger, N. (2001). Cognitive behavioral programs for offenders. The Annals of the American Academy of Political and Social Science 578, 144-157.
- MacKenzie, D. (1989). Prison classification: The management and psychological perspectives. In L. Goodstein & D. MacKenzie (Eds.), The American Prison: Issues in Research and Policy. New York, NY: Plenum Press.
- MacKenzie, D. (2000). Evidence-based corrections: Identifying what works. Crime & Delinquency, 46, 457-471.
- Martinson, R. (1974). What works? Questions and Answers about prison Reform. The Public Interest, 35, 22-54.
- Meichenbaum, D. & Jaremko, M. (1982). Stress prevention and management: A cognitive behavioral approach. New York, NY: Plenum.
- Morash, M., Bynum, T., & Koons, B. (1998). Women offenders: Programming needs and promising approaches. National Institute of Justice: Research in Brief. Washington, DC: U.S. Department of Justice.
- Novaco, R. (1975). Anger control: The development and evaluation of an experimental treatment. Lexington, MA: Lexington Books.
- Palmer, T. (1974). The Youth Authority's Community Treatment Project. Federal Probation, 38, 3-14.
- Palmer, T. (1978). Correctional intervention and research: Current issues and future prospects. Lexington, MA: D.C. Heath.
- Palmer, T. (1992). The re-emergence of correctional intervention. Newbury Park, NJ: Sage.
- Palmer, T. (2002). Individualized intervention with young multiple offenders. New York, NY: Routledge.
- Porporino, F., Fabiano, E. & Robinson, D. (1991). Focusing on successful reintegration: Cognitive skills training for offenders. Ottawa, Ontario: Correctional Services of Canada, R-19.
- Reid, N. (1986). Wide-Range Achievement Test – 1984 Revised Edition, Journal of Counseling and

Development, 64, 538-39.

Reitsma-Street, M. & Leschied, A. (1988). The conceptual-level matching model in corrections. Criminal Justice and Behavior, 15, 199-212.

Research and Statistics Branch of Correctional Services of Canada (1991). Effectiveness of the Cognitive Skills Training Program: From pilot to national implementation. Ottawa, Ontario: Correctional Services of Canada, R-07.

Reynolds, C. (1986). Wide-Range Achievement Test (WRAT-R), 1984 Edition, Journal of Counseling and Development, 64, 54-41.

Robinson, D. (1995). The impact of Cognitive Skills Training on post-release recidivism among Canadian federal offenders. Ottawa, Ontario: Correctional Services of Canada, R-41.

Ross, R. & Fabiano, E. (1985). Time to Think: A cognitive model of delinquency prevention and offender rehabilitation. Ottawa, Ontario: Air Training and Publications.

Ross, R., Fabiano, E., & Ewles, C. (1988). Reasoning and Rehabilitation. International Journal of Offender Therapy and Comparative Criminology, 32, 29-35.

Ross, R. & Fabiano, E., & Ross, R. (1989) Reasoning and Rehabilitation: A handbook for teaching cognitive skills. Ottawa, Ontario: Flix Desktop Services.

Sarason, I. & Ganzer, V.(1973). Modeling and group discussion in rehabilitation of juvenile delinquents. Journal of Counseling Psychology, 20, 442-449.

Simourd, D. (1997). Criminal Sentiments Scale-Modified and Pride in Delinquency: Psychometric properties and construct validity of two measures of criminal attitudes. Criminal Justice and Behavior, 24, 52-70.

Spruance, L., Van Voorhis, P., Johnson Listwan, Pealer, & Seabrook, R. (2001). The Georgia Cognitive Skills Program: Process evaluation-- Phase II. Cincinnati, OH: University of Cincinnati.

Sullivan, C., Grant, M. Q., & Grant, D. (1957). The development of Interpersonal Maturity: An application to delinquency. Psychiatry, 20, 373-386.

Van Voorhis, P. (1985). Restitution outcome and probationers' assessments of restitution: The effects of Moral Development. Criminal Justice and Behavior, 12, 259-287.

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

Van Voorhis, P. (1994). Psychological classification of the adult male prison inmate. Albany, NY: SUNY Press.

Van Voorhis, P. (1997). Correctional classification and the responsivity principle. Forum on Corrections Research, 9, 46-50.

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

Van Voorhis, P., Cullen, F., & Applegate, B. (1995). Evaluating interventions with violent offenders: A guide from practitioners and policy makers. Federal Probation, 59, 17-28.

Van Voorhis, P., Murphy, L., & Johnson, S. (1999). Process evaluation of the Georgia Cognitive Skills Program: Phase I. Cincinnati, OH: University of Cincinnati.

Van Voorhis, P., Spruance, L., Listwan, S., Ritchey, P., Pealer, J., & Seabrook, R. (2001). The Georgia Cognitive Skills Experiment outcome evaluation: Phase I. Cincinnati, OH. University of Cincinnati.

Walters, G. (2001). Psychological Inventory of Criminal Thinking Styles. Version 4.0. Unpublished manuscript.

Warren, M. (1971). Classification of offenders as an aid to efficient management and effective treatment. Journal of Criminal Law, Criminology and Police Science, 62, 239-268.

Warren, M. (1983). Application of Interpersonal Maturity Theory to offender populations. In W. Laufer and J. Day (eds.) Personality theory, moral development, and criminal behavior. Lexington, MA: Lexington Books.

Warren, M. & staff of the Community Treatment Project (1966). Interpersonal Maturity Level Classification: Diagnosis and treatment of low, middle, and high maturity delinquents. Sacramento, CA: California Youth Authority.

Warren, M. & staff of the Community Treatment Project (1966). Interpersonal Maturity Level Classification: Diagnosis and treatment of low, middle, and high maturity delinquents. Sacramento: California Youth Authority.

Watson, D. & Clark, L (1984). Negative affectivity: The disposition to experience aversive emotional states. Psychological Bulletin, 96, 465-490.

Wilson, D., Allen, L., & MacKenzie, D. (2000). Quantitative review of cognitive behavioral programs. College Park, MD: University of Maryland.

Yochelson, S. & Samenow, S. (1976). The criminal personality: A profile for change. Vol. 1. New York, NY: Jason Aronson.

Appendix F. Summary of key personality constructs for the original Jesness Inventory personality types (Jesness, 1962).²⁹

Construct	Definition	Relationship to Criminal Behavior
I-Level 2		
Unsocialized aggressive (Aa)	Displays negative attitudes toward convention. Behavior is unpredictable, aggressive, and antisocial.	High self reported criminal behavior.
Unsocialized passive (Ap)	Displays negative attitudes toward convention. Behaviors in inappropriate and often bizarre ways. Has a negative self concept.	Probability of criminal behavior average.
I-Level 3		
Immature conformist (Cfm)	Displays positive attitudes toward convention. Behavior is conforming and dependent (i.e., follower).	Low self reported criminal behavior. Peers may be especially important explanation of criminal involvement.
Cultural conformist (Cfc)	Displays low motivation and negative attitudes toward convention. Feels alienated and hostile toward authority.	High self reported behavior. Above average violent activity.
Manipulator/pragmatist (Mp)	Shows positive attitudes toward convention. Behaves in manipulative and obtrusive ways.	Low self reported criminal behavior but official record indicates high probability. Especially prone to property crimes.
I-Level 4		
Neurotic acting out (Na)	Displays negative attitudes toward convention. Experiences conflicts in relationships. See themselves as somewhat cynical and disenchanting, often exhibiting outspoken and non-conforming behavior.	High self reported criminal behavior. More apt to use drugs.
Neurotic anxious (Nx)	Mostly positive attitudes toward convention. Conforming but also dependent, anxious, and insecure. Do not have criminal orientation.	Low self reported criminal behavior.
Situational emotional reaction (Se)	Positive attitudes toward convention. They are confident, but naïve, rigid, and conforming.	Low reports of criminal behavior. Expect to be caught if he or she breaks the law.
Cultural identified/adaptive (Ci)	High verbal aptitude and positive toward convention. Maintains good interpersonal relationships.	Low reports of criminal behavior.

²⁹ These nine subtypes are used to create the four collapsed types (Van Voorhis, 1994) used in this study (see Figure 3).

Appendix G: Felony arrest/revocation rates and returns to prison by experimental condition and time.

Independent Variables	Felony Arrest/Revocation		Return to Prison	
	B	Exp(B)	B	Exp(B)
Experimental condition				
Experimental group	-.09	.92	-.11	.90
Time period ^a				
3 month	-.51***	.60	.14	1.15
6 month	-.28*	.76	.22	1.25
9 month	-.54***	.58	.24	1.27
12 month	-1.03****	.36	.03	1.03
15 month	----	----	-.04	.96
18 month	----	----	-.34	.71
21 month	----	----	-.63*	.53
24 month	----	----	-.08	.92
27 month	----	----	-.81*	.45
30 month	----	----	.33	1.39
33 month	----	----	-4.30	.01
Intercept	-1.88****		-2.85****	
Model Chi Square	31.82****		25.50***	
-2 log likelihood	2225.80		2621.86	

**** $p \leq .001$; *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$

^aThe omitted dummy variable is “during program” time period.

Appendix H: Parolee technical violations at 10 to 12 month follow up by experimental condition, control variables, and risk: interaction – risk X experimental condition.

Independent Variables	B	Exp(B)
Risk X Experimental	-.73**	.48
Experimental condition		
Experimental group	.74***	2.10
Medium/high risk	.63***	1.88
Prior conviction for violent crimes	-.07	.93
IQ	.01	1.01
Reading above 5 th grade	-.01	.99
High school graduate	-.38**	.68
Married	-.41**	.66
Middle class	-.10	.91
White	-.23	.79
Age ^a		
Age 18 to 22	.02	1.02
Age 23 to 27	-.23	.79
Age 28 to 32	-.33	.72
Age 33 to 37	.30	1.35
Employment status ^b		
Employed fulltime	.14	1.15
Employed parttime	-.06	.94
Unemployed < 6 mo.	.36	1.43
Personality type ^c		
Aggressive	-.30	.74
Dependent	-.35	.70
Situational	-.39	.68
I-level 4	-.36	.70
Intercept	-.88	
Model Chi Square	35.88***	
-2 log likelihood	911.75	
Decrement to Chi Square #		
Chi Square (df)	3.90**	(1)

**** $p \leq .001$; *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$

The omitted dummy variables are (a “age 38 and older”, (b) “other”, and (c) “neurotic”. # The decrement to chi square indicates whether the interaction model is a significant improvement over a simpler additive model.

Appendix I: Parolee Felony Arrest/Revocation and Return to Prison by Program Completion and Time.

Independent Variables	Felony Arrest/Revocation		Return to Prison	
	B	Exp(B)	B	Beta
Completion status ^a				
Graduate	-.47***	.63	-.45***	.64
Dropout for non-criminal reason	-.03	.97	.06	1.06
Fail during the program	1.67****	5.33	1.36****	3.89
Time period ^b				
3 month	-.42**	.66	.23	1.26
6 month	-.10	.91	.37*	1.45
9 month	-.34*	.71	.41**	1.51
12 month	-.83****	.44	.21	1.23
15 month	----	----	.14	1.15
18 month	----	----	-.17	.84
21 month	----	----	-.47	.63
24 month	----	----	.06	1.07
27 month	----	----	-.65	.52
30 month	----	----	.49	1.64
33 month	----	----	-4.14	.01
Risk level	1.06****	2.87	1.03****	2.80
Prior conviction for violent crimes	.02	1.02	.03	1.03
IQ	.00	1.00	.00	1.00
Reading level	-.01	.99	.00	1.00
High school graduate	-.25*	.78	-.21	.81
Married	-.09	.92	-.01	.99
Middle class	-.04	.96	-.11	.89
White	.01	1.01	.01	1.01
Age ^c				
Age 18 to 22	.20	1.22	.31	1.36
Age 23 to 27	.05	1.05	-.07	.94
Age 28 to 32	.01	1.01	-.25	.78
Age 33 to 37	.04	1.04	.11	1.11
Employment status ^d				
Employed fulltime	.25	1.28	.33**	1.38
Employed parttime	.17	1.18	.03	1.03
Unemployed < 6 mo	.38*	1.47	.55***	1.73
Personality type ^e				
Aggressive	-.23	.80	-.10	.91
Dependent	-.29	.75	-.03	.97
Situational	-.25	.78	-.14	.87
I-level 4	-.17	.84	-.08	.93
Intercept	-3.01****	.05	-3.75****	.02
Model Chi Square	159.55****		146.60****	
-2 log likelihood	2069.71		2479.14	

**** p ≤ .001; *** p ≤ .01; ** p ≤ .05; * p ≤ .10; NS = not significant

^a The omitted dummy variable is “control group”.

^b The omitted dummy variable is “during program” time period.

^c The omitted dummy variable is “age 38 and older”.

^d The omitted dummy variable is “other”.

^e The omitted dummy variable is “neurotic”.

Appendix J: Felony Arrest/Revocation by Class Size: Graduate (Logistic Regression) and Class (Linear Regression) Levels of Analyses.

Independent Variables	Graduates Arrest Rates: Class Size Cutoff between 7 & 8 Members		Graduates Arrest Rates: Class Size Cutoff between 8 & 9 Members		Class Arrest Rates: Class Size Cutoff between 8 & 9 Members	
	B	Exp(B)	B	Exp(B)	B	Beta
Class size cutoff between 7 & 8	-.52**	.59	----	----	----	----
Class size cutoff between 8 & 9	----	----	-.43**	.65	-8.08E-02*	-.236
Time period ^a						
3 month	.51	1.67	.51	1.67	----	----
6 month	.91***	2.49	.92***	2.50	----	----
9 month	.48	1.62	.47	1.60	----	----
12 month	.12	1.12	.11	1.12	----	----
Risk level	.48*	1.61	.46	1.59	----	----
Prior conviction for violent crimes	-.06	.94	-.07	.94	----	----
IQ	.01	1.01	.01	1.01	----	----
Reading level	-.03	.97	-.04	.97	----	----
High school graduate	-.06	.95	-.03	.97	----	----
Married	-.04	.96	-.06	.95	----	----
Middle class	.08	1.08	.10	1.10	----	----
White	-.59**	.56	-.57**	.57	----	----
Age 23 to 27	-.17	.84	-.20	.82	----	----
Neurotic	.41	1.50	.40	1.49	----	----
Intercept	-3.25***	.04	-3.49***	.31	.30***	
Model Chi Square	26.37**		25.60**		----	----
-2 log likelihood	640.06		640.83		----	----
F	----	----	----	----	3.06*	
Adjusted R Square	----	----	----	----	.04	

**** $p \leq .001$; *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$; NS = not significant

Appendix K: Felony Arrest/Revocation by Graduates' Use of Skills¹.

Independent Variables	B	Exp(B)
Use of skills	-1.06**	.35
Time period ^a		
3 month	.58	1.78
6 month	.99***	2.70
9 month	.55	1.74
12 month	.19	1.21
Risk level	.51*	1.67
Prior conviction for violent crimes	-.06	.95
IQ	.01	1.01
Reading level	-.03	.97
High school graduate	-.03	.98
Married	-.17	.84
Middle class	.06	1.06
White	-.57**	.57
Age 23 to 27	-.20	.82
Neurotic	.31	1.37
Intercept	-2.91***	.05
Model Chi Square	26.06**	
-2 log likelihood	632.89	

**** p ≤ .001; *** p ≤ .01; ** p ≤ .05; * p ≤ .10; NS = not significant

¹ Use of skills is measured as the mean of participants' use of skills (y/n) during four class sessions.

Appendix L: Felony Arrest/Revocation by Class Structure: Graduate (Logistic Regression) and Class (Linear Regression) Levels of Analyses.

Independent Variables	Individual Arrest Rates		Class Arrest Rates	
	B	Exp(B)	B	Beta
Classroom structure scale	-.97**	.38	-.18**	-.45
Who observed the class	.95*	2.58	.21**	.41
Time period ^a				
3 month	.51	1.67	----	----
6 month	.92***	2.50	----	----
9 month	.49	1.63	----	----
12 month	.13	1.14	----	----
Risk level	.44	1.55	----	----
Prior conviction for violent crimes	-.07	.94	----	----
IQ	.00	1.00	----	----
Reading level	-.04	.96	----	----
High school graduate	-.03	.97	----	----
Married	-.08	.93	----	----
Middle class	.04	1.04	----	----
White	-.45	.64	----	----
Age 23 to 27	-.21	.81	----	----
Neurotic	.35	1.42	----	----
Intercept	-3.37***	.03	.22*	
Model Chi Square	28.31**		----	----
-2 log likelihood	638.12		----	----
F	----	----	2.97	
Adjusted R Square	----	----	.07	

**** $p \leq .001$; *** $p \leq .01$; ** $p \leq .05$; * $p \leq .10$; NS = not significant

Class structure is measured through a scale of items where 0 = needs improvement, 1 = satisfactory, and 2 = very satisfactory.