Psychometric Evaluation of the
Correctional Program Assessment Inventory (CPAI)

Aleksandra Nesovic

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Psychology

Carleton University
Ottawa, Ontario
May, 2003

Copyright © 2003 Aleksandra Nesovic

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-83525-1

Our file Notre référence

ISBN: 0-612-83525-1

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Canada

The undersigned recommend to the
Faculty of Graduate Studies and Research
Acceptance of the thesis:

"Psychometric Evaluation of the Correctional Program Assessment Inventory (CPAI)"

submitted by

Aleksandra Nesovic

in Partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Chair

External Examiner

Thesis Supervisor

Carleton University
February, 2003

Abstract

The Correctional Program Assessment Inventory (CPAI), an inventory designed to assess the extent to which correctional programs meet principles of effective correctional interventions, summarizes many of the treatment elements empirically linked to the reduction of recidivism. This dissertation examined the psychometric characteristics of the CPAI. Three separate, methodologically distinct studies were conducted to evaluate interrater reliability/internal consistency, face/content validity and predictive validity of this inventory.

In the first study two independent raters scored the CPAI for twenty federally and provincially delivered correctional program based on the CPAI-Q (Correctional Program Assessment Inventory Questionnaire). Additionally, two raters scored audiotaped interviews conducted with treatment staff for thirteen of these programs. Obtained interrater reliability estimates indicated satisfactory interrater reliability for both methods on all CPAI sections, treatment and total scores. Internal consistency estimates indicated that the CPAI represents a homogenous scale.

The face and content validity study examined the degree to which samples of students ($n=50$), treatment staff ($n=34$) and psychologists ($n=50$) agreed that the CPAI is an instrument that could assess the potential of correctional programs to reduce reoffending. The level of agreement was high within and across the three samples, indicating that this inventory has satisfactory face and content validity.

The third study examined the predictive validity of the CPAI by establishing a link between the CPAI scores and effect size, a statistical measure that expresses the difference in outcome between treatment and control group. One hundred and seventy three articles of treatment effectiveness that yielded 266 effect sizes were scored using the modified version of

the CPAI scoring manual. Results indicated that the higher number of CPAI items present, the higher the effect sizes were. This correlation was statistically significant across different CPAI sections, treatment and total scores within both the control and treatment groups. Treatment CPAI scores maintained significant correlation with effect size under a variety of statistical and methodological considerations.

Overall, the results suggest that the CPAI has satisfactory psychometric properties. The limitations of the study, their implications for correctional practice and suggestions for future research were discussed.

Acknowledgments

I would like to thank my advisor, Dr. Don Andrews, who introduced me to the psychology of criminal conduct, and guided my work throughout the school years. Only he knows how much patience and discussion it took to combine my clinical psychology background with correctional empirical research. Thank you for making me aware of a new, exciting path to modern corrections.

I am deeply grateful to Dr. Adelle Forth for taking the time to discuss with me many issues throughout my school years. Her office was always open and she provided so much needed support and directions when at times I did not see the solutions. I would like to thank Dr. John Weeks for his insightful comments, help and friendly support.

My sincere appreciations to Dr. Peter Barrett and Dr. Witold Poplawski from Millbrook Correctional Centre, for their friendship and guidance in a challenging correctional setting. I learned a lot from you and for that I am very thankful. Debbie Carrierre provided much needed support and encouragement by endlessly reviewing this study. At important times she helped me believe that I could complete this research when I was so discouraged. Thank you, Deb.

I would like to express my gratitude to all participants in this study who were so knowledgeable, friendly and simply great. A special thank you to the Ministry of Correctional Services and the Correctional Service of Canada.

Many thanks to Craig Dowden and Sharon Dunley who helped with scoring and to Tzvetanka Dobрева-Martinova for great research suggestions and her unconditional friendship. And to Lynn Giff and Kim Cook for their patience, understanding and support over many years of my graduate work. I am very grateful.

Finally, I would like to thank my family, my husband Zeljko and my son Karlo who were always supportive even at the time when the end of my studies looked so far away. Without their support and love, this dissertation would not have been possible.

Table of Contents

Abstract	iii
Acknowledgments	v
Table of Contents	vii
List of Tables	ix
List of Appendices	x
Introduction	1
Meta Analysis: Breakthrough in Correctional Research	5
The Correctional Program Assessment Inventory	10
Principles of Effective Correctional Intervention	17
Risk Principle	18
Need Principle	21
General Responsivity Principle	22
Specific Responsivity Principle	24
Appropriate Treatment	26
Treatment Integrity	28
Staff Characteristics	35
Summary	39
The Present Study	44
Study 1: Interrater and Intermethod Reliability	46
Goals	46
Method	46
Sample	47
Hypotheses	49
Results	49
Scoring the CPAI-Q Responses	49
Interrater reliability estimates	51
Mean differences	55
Summary levels for CPAI scores	56
Summary	59
Scoring Audiotaped Interviews	59
Interrater reliability estimates	62
Mean differences	64
Summary levels for CPAI scores	65
Summary	66
Comparison of Two CPAI Scoring Methods	66
Qualitative Analysis of the CPAI Items	69
Program implementation	69
Preservice assessment	70
Program characteristics	71
Staff characteristics	71
Evaluation	72
Other	72
Discussion	73
Study 2: Face/Content Validity Study	77

Face Validity Study – Goals	77
Content Validity Study-Goals.....	78
Method	79
Sample	81
Hypotheses	82
Results.....	82
Sample/Demographics	82
Intraclass Correlation	85
Mean Rating Values.....	87
Discussion	89
Study 3: Predictive Validity Study	92
Goals	92
Method	93
Contribution of the Present Study.....	93
New CPAI Coding Procedures	94
CPAI Coding Manual: Modifications	96
Scoring the CPAI: Missing Data	98
Sample	100
Hypotheses	101
Results.....	101
Sample Descriptives	101
Mean Effect Sizes	103
CPAI Sections – Descriptives.....	104
Summary Levels across CPAI Scores	106
Predictive Validity Estimates.....	108
Regression Analyses	112
Individual CPAI Items	113
Treatment Level.....	116
Discussion	123
Conclusions and Implications for Future Research	127
References.....	132

List of Tables

1	Frequencies and Percentages of Coded Categories (CPAI-Q)	50
2	Cross tabulation of the Coded Categories (CPAI-Q)	51
3	Kappa, Pearson's Correlation and Chronbach's Alpha Coefficients (CPAI-Q).....	52
4	T Test for Difference Between Means for CPAI scores (CPAI-Q)	55
5	Summary Levels for CPAI Sections (CPAI- Q).....	57
6	T test for Difference Between Means (Institutional/Community Programs)	58
7	Frequencies and Percentages of Coded Categories (Interview)	61
8	Cross Tabulation of the Coded Categories (Interviews).....	62
9	Kappa, Pearson's Correlation, and Cronbach's Alpha Coefficients (Interviews)	63
10	T Test for Difference Between Means for CPAI Scores (Interviews).....	65
11	Summary Levels for CPAI Sections (Interviews)	66
12	Comparison of Two CPAI Scoring Methods.....	68
13	Intraclass Correlation: Two Way Random Effect Model (Absolute Agreement)	85
14	Mean Ratings for CPAI Sections.....	87
15	CPAI Descriptives for Treatment and Control Groups	104
16	Bivariate Correlations of CPAI Scores with Effect Sizes	109
17	Individual CPAI Items: Mean Effect Sizes, F Tests and Eta.....	115
18	Mean Effect Sizes by CPAI Treatment Level and Control Variables	119
19	Mean Effect Sizes by CPAI Treatment Level and Treatment Targets	122

List of Appendices

A	Correctional Program Assessment Inventory (CPAI) - Scoring Guide	164
B	Correctional Program Assessment Inventory Questionnaire (CPAIQ)	177
C	List of the Participating Program Sites	200
D	Interrater Reliability Indicators - Two Coded Categories	201
E	Kappa Coefficients for Individual Program Sites.	202
F	Ranges within the CPAI Sections (CPAI-Q)	203
G	Ranges within the CPAI Sections (Interview)	204
H	Effective Treatment Elements Questionnaire (ETEQ)	205
I	ETEQ Individual Items: Descriptive Statistics	213
J	Meta-analysis- CPAI Scoring Manual	229
K	Values for Eta for Individual CPAI Items	243
L	Ranges within the CPAI Sections Scores (Meta-analysis)	246
M	Informed Consent - (CPAIQ).....	247
N	Informed Consent Form	248
O	Informed Consent (Student sample) - (ETEQ)	249
P	Informed Consent (Psychologist and Treatment Staff Sample) - (ETEQ)	250
Q	Correctional Program Assessment Inventory Questionnaire-Debriefing	252
R	Effective Treatment Elements Questionnaire-Debriefing.....	253

Introduction

Throughout history the interpretation and explanation of the nature and causes of criminal behavior have influenced treatment allocation for persons who violated the law. The oldest approach of the justice system is use of the deterrence and retribution models for the processing and/or management of offenders. This approach to the management of offenders has been ineffective and in some instances even linked to an increase in recidivism (Andrews et al., 1990; Bonta & Gendreau, 1992). The political beliefs influencing correctional practice have been antagonistic to the knowledge about effective correctional programs and have instead led to the development of deterrence, sanction and punishment based correctional approaches (Leschied, 2002). Despite the evidence for the effectiveness of correctional programs, a systematic disregard of empirical findings, as indicated by Gendreau, Goggin and Smith (2002), has led to the correctional policies that did not support the rehabilitative approach. According to Andrews and Bonta (1998), numerous research studies continued to provide impressive evidence that the delivery of human service in the treatment of offenders is associated with a reduction in criminal behavior under a variety of judicial sanctions (diversion, parole and custody). Therefore, the attempt to reduce recidivism rates of offenders should concentrate on rehabilitation (treatment) within the context established by justice models as there is no empirical evidence that reoffending can be reduced by criminal sanctions (Andrews, 2002).

Studies that have explored the effectiveness of correctional treatment programs have frequently used different outcome criteria. Correctional programs were considered effective if they influence psychological/behavioral variables (e.g., recidivism reduction, number of institutional misbehaviors, interpersonal adjustment). Bonta, Boyle, Motiuk, and Sonnischsen

(1985) argued that the viability of a program cannot rest solely on a single outcome, regardless of how important that outcome is. Some interventions may have no effect on recidivism but can still be effective in the reduction of institutional misbehaviors and/or psychological variables. Nevertheless, recidivism reduction was often considered the ultimate criterion for program effectiveness. Most of the research on program effectiveness has concentrated on explaining the ways in which programs influence recidivism rates of offenders.

Despite extensive research findings offering strong support that correctional treatment “works”, the antirehabilitation sentiment led to a systematic disregard for the empirical evidence. However, the “Nothing works” debate was thought provoking and has led to an increase in the knowledge of what constitutes an effective correctional treatment program. Although many programs have reported a reduction in the recidivism rates of offenders, the effectiveness of correctional programs has been frequently questioned and disputed. The publication of Martinson’s (1974) review of 231 studies on treatment effectiveness published from 1945 to 1965 had a lasting negative effect on public confidence in correctional treatment and a strong impact on correctional policy. Based on this review, he suggested: “with few and isolated exceptions, the rehabilitative efforts that have been reported so far have had no appreciable effect on recidivism”(p.25). The widespread acceptance of this “nothing works” position influenced the change from treatment and rehabilitation to a deterrence model (Thornton, 1987).

Many authors have argued that Martinson was biased in his conclusion and that his interpretations were questionable (e.g., Cullen & Gendreau, 1989; Gendreau & Ross, 1987; Lipton, Pearson, Cleland, & Yee, 1997; Palmer, 1975; Thornton, 1987) as he analyzed studies

that reported on different treatment methods with different outcome measures. The programs described in these studies had different durations and intensity levels and were delivered to heterogeneous offender samples. The influence of possible moderator variables that can explain the link between treatment and outcome was not explored and his interpretation did not take into account methodological qualities of the studies reviewed (Ross & Fabiano, 1985).

Lipton, Martinson, and Wilks's (1975) book that reported on the same research did not express such a pessimistic view on treatment effectiveness. They maintained that treatment can work under certain conditions. The re-examination by Thornton (1987) of the same studies reviewed by Martinson (1974) determined that many of them have positive effects on recidivism reduction. The reviews of controlled outcome studies on treatment effectiveness (Baily, 1966; Gendreau & Ross, 1979, 1987; Kirby, 1954; Logan, 1972) were reanalyzed by Andrews and Bonta (1998). The authors maintained that the majority of the studies that were included in these reviews reported that correctional treatments can reduce criminal behavior of offenders. More specifically, from 40% up to 80% of the studies have indicated that the delivery of correctional treatment was linked to recidivism reduction.

Contrary to what Martinson (1974) concluded, the evidence that correctional programs can influence changes in recidivism rates was emerging rapidly. Effective programs were reported across institutional (Wexler, Falkon, & Lipton, 1990) and community (Lee & Olejnik, 1976) settings and with different treatment interventions such as employment and vocational (Saylor & Gaes, 1992), educational (O'Neil, 1990; Walsh, 1985), and cognitive behavioral (Henning & Frueh, 1996) programs. The positive effects were found among juveniles (Fagan, 1990) and adults (Lindforss & Magnusson, 1997), among sex offenders

(Davies & Leitenberg, 1987; Marshal, & Barbaree, 1988), and substance abusing offenders (Annis, 1983; Eisberg & Fabelo, 1996).

The central question of treatment effectiveness literature has been gradually shifted from “Do correctional programs work?” to “Under what conditions do correctional programs work?”. There was a need to explore the “black box” of treatment programs (Cullen & Gendreau, 1989; Gendreau, 1996) and to determine the treatment elements that distinguish between effective and ineffective programs. In recent years, different treatment types and treatment elements that influence program effectiveness have been identified. The increasing knowledge pertaining to the elements of successful programs has led to a systematic evaluation of correctional programs using process evaluation. The purpose of this approach was to establish whether correctional programs incorporated potentially important treatment elements. A search for a psychometrically sound approach in process evaluation resulted in the construction of the Correctional Program Evaluation Inventory (CPEI, Gendreau & Andrews, 1990) and its later version, the Correctional Program Assessment Inventory (CPAI, Gendreau & Andrews, 1996). The CPAI is still in an experimental phase, with potential as an instrument that provides timely feedback on the recidivism reduction potential of correctional treatment programs. However, at the present time, there is little data on the psychometric properties of the CPAI.

The goals of this dissertation are: a) to evaluate psychometric properties of the CPAI, and b) to explore the value of this inventory for the evaluation of treatment programs. Before summarizing the literature of treatment effectiveness, the instrument itself and the rationale on which it is based will be described. Since the main research findings pertaining to the treatment effectiveness were the result of meta-analytic studies, the advantages and

disadvantages of this approach will be discussed in the next section. Finally, the methodology and empirical findings of the psychometric evaluation of the CPAI will be presented.

Meta Analysis: Breakthrough in Correctional Research

Despite the large number of outcome studies reporting on the effectiveness of correctional programs, the comparison of these programs within narrative reviews was generally fragmented due the variety of treatments, methods, samples, and outcome measures. Wolf (1986) maintained that narrative reviews frequently had selective inclusion criteria, and subjective interpretations of results. In addition, these reviews cannot take into account moderating variables that can have an effect on program efficacy and in some instances, they even neglect different methodological characteristics of studies. These limitations have led to the acceptance of an alternative way of reviewing the literature, namely, meta-analysis.

The meta-analytic literature review is an objective and systematized way of comparing the results among different outcome studies (Glass, McGrow, & Smith, 1981; Rosental, 1991; Wolf, 1986). Meta-analysis aggregates the statistical results of a number of studies into a single database in which the results are analyzed collectively rather than individually and can be replicated by different researchers (Lipsey, 1995). Using this approach, both the methodological issues present in primary studies and potential moderating variables that influence program outcome, can be controlled for. The results from individual studies are expressed in terms of a statistical index called an effect size, which indicates the strength of the relationship between treatment and its outcome. The effect size allows for the comparison of results across a number of studies and averaging across different studies (Andrews & Bonta, 1998). The mean effect sizes may be interpreted as the difference in recidivism rates between treatment and control groups, with a positive difference indicating that the recidivism

of the treatment group was lower than that of the comparison group. For example, an effect size of .20 would indicate a recidivism rate of 40% in the treatment group and 60% in the comparison group.

There are many potential threats to the meta-analytic research as indicated by Rosenthal (1991). The first potential threat is a heterogeneous sample of primary studies that could prevent the combination of the results. However, this threat could be managed by formulating specific inclusion criteria for meta-analytic studies. Secondly, a poor quality of primary studies might influence the results. The defenders of meta analytic research argue that the quality of a study is an arbitrary characteristic. Finally, there was a concern that meta-analytic research includes a limited number of unpublished studies that could lead to overestimated values of the effect size. The assumption was that negative findings were less likely to be published. The recent meta-analyses have included an increasing number of unpublished studies. Despite these limitations meta-analytic research represents an advantage in comparison to the traditional narrative review of the literature. Coding different study characteristics enables the researcher to examine the relationships among variables which can be tested under a variety of control conditions. Also, conclusions are based on statistical analysis rather than subjective interpretation of the results. Finally, the meta-analytic research can be repeated by independent researchers.

The extensive meta-analytic reviews of the treatment effectiveness literature had some contradictory results regarding effectiveness of correctional programs. While one of the early meta-analysis (Whitehead & Lab, 1989) concluded that there is no evidence for the effectiveness of correctional treatment, many others (Andrews, 1996; Andrews & Bonta, 1998: Resource Note 10.1; Andrews, Dowden, & Gendreau, 1999; Andrews et al. 1990;

Cleland, 1997; Dowden, 1998; Garrett, 1985, Gottschalk et al. 1987 a,b; Hill, Andrews & Hoge, 1991; Izzo & Ross, 1990; Lipsey, 1992; Losel, 1993) reported positive effects of treatment on recidivism reduction. The overall effect has been mildly positive across different meta-analyses. An exact estimate of the overall general mean effect cannot be obtained because the sets of primary studies overlap in the various meta-analyses (Losel, 1995). Although obtained values are not very high, they still indicate a reduction in recidivism rates that has practical significance. For example, a mean effect size of .10 indicates that a recidivism rates for a treatment group is 45% compared to the control group with a 55% recidivism rate (Losel, 1993). There are several possible explanations for the relatively small effect sizes obtained in meta-analytic studies (Losel, 1995). First, control groups usually consist of participants undergoing different types of treatment rather than no treatment at all. Second, measurement of treatment and outcome variables may be insufficiently reliable. Third, the presence of dichotomized independent (treatment) and dependent (recidivism) variables may reduce statistical power. Fourth, the meta-analyses that include many unpublished studies (e.g., Gensheimer et al., 1986; Lipsey, 1992) tend to have lower effect sizes, but they probably set a lower limit for estimated effect sizes. Finally, the type of treatment determines effect size.

Whitehead and Lab (1989) conducted a meta-analysis of 35 studies of juvenile programs and concluded that there is no support for the rehabilitative potential of programs. However, Andrews et al. (1990) reanalyzed these results paying particular attention to what made some programs effective and others not. This meta-analysis was conducted on 154 treatment comparisons obtained from a sample of juvenile and adult programs that also included Whitehead and Lab (1989) sample. They found that treatment is effective when

based on relevant principles of risk, need and responsivity (see next sections for a discussion of risk, need and responsivity principles). The same pattern of results was reported by Andrews (1996), Andrews and Bonta (1998: Resource Note 10.1) and Dowden (1998). These meta-analyses based on the Carleton University data bank, extended Andrews et al. (1990) meta-analysis. Dowden (1998) study was based on 374 treatment comparison gathered from 225 articles of adult and juvenile treatment effectiveness studies.

Garett (1987), in a meta-analysis conducted on the juvenile offender literature, found that treatment programs (especially behavioral programs) have an effect on recidivism reduction. Lipsey (1992), on a large sample of juvenile offender programs, found that more than half of the studies reported positive effects of treatment programs on recidivism. Many other meta-analyses have found a positive influence of treatments in reducing recidivism (Antonowicz & Ross, 1994; Gottschalk et al., 1987 a.;b; Izzo & Ross, 1990). Based on these results, it could be argued that there is strong empirical evidence that correctional treatment “works” in reducing the recidivism of offenders.

Some authors (e.g., Lipsey, 1995; Losel 1995) identified few concerns with meta-analytic research of correctional treatment effectiveness. Lipsey (1995) argues that the absence of specific program details in primary studies results in coding broad treatment categories that are not clearly defined and distinct. He suggests that detailed analysis and coding of different treatment types and their distinct characteristics in primary studies should be conducted. The results from Lipsey’s (1992) meta-analysis had 3 to 4 times higher variability than expected if the primary studies reflected the same underlying treatment effects. This led some authors (e.g., Antonowicz & Ross, 1994; Losel, 1995) to conclude that multimodal treatment combinations can be more important than any particular treatment per

se. The coding of appropriate treatment based on risk, need and responsivity factors used in meta-analyses reported by Andrews (1996), Andrews and Bonta (1998: Resource Note 10.1), Andrews et al. (1990), Dowden (1998), and Hill, Andrews and Hoge (1991) avoids these limitations.

Losel (1995) questions the generalizability of the results obtained on a sample of predominantly juvenile offenders to adult offender treatment. The majority of the primary studies (over 80%) included in the meta-analyses have reported on programs delivered to juvenile offenders. Indeed some studies have reported that the treatment effectiveness found in juvenile samples is not found among adults (Cleland, 1997). However, other meta-analyses reported that similar patterns in treatment effectiveness were found in both juvenile and adult samples (Andrews, 1996; Andrews et al., 1990; Dowden, 1998; Losel, 1993). An additional concern is that meta-analytic and primary studies do not systematically address the question of specific program elements that are linked to program effectiveness. Losel (1995) argued that better controlled primary studies would be desirable in order to determine the impact of program elements on treatment effectiveness. Many relevant program characteristics were underexplored as a result of the lack of studies with clear empirical indicators of program effectiveness. For example, treatment integrity is an important precondition for the success of adequate programs (Hollin 1993; Quay 1987; Rezmovic 1984). Evaluation of poorly implemented programs can lead to Type II error conclusions that treatments were ineffective, when in fact the inadequate treatment implementation precluded conclusive evidence of its effectiveness. However, some recent meta-analyses (Andrews, 1996; Andrews & Dowden, 1999; Dowden 1998; Hill et al. 1991) have explored different treatment related elements (e.g., treatment integrity) that may impact the evaluation of program effectiveness.

In conclusion, the results across different meta-analyses indicate an overall positive effect of correctional programs. Despite its limitations, meta-analysis is a more systematic and objective method of summarizing research than a traditional narrative review of studies. Whereas meta-analysis might not provide a definitive answer as to what constitutes the “best” correctional programs, it indicates which treatments are more likely to be effective. Gendreau, Goggin, and Smith (2002) cautioned that, despite meta-analytical findings being presented as definite answers, the lack of essential details in many primary studies indicates that additional research is still needed. Narrative studies, with tendency to underestimate the effect of treatment, might still be an option for the issues that have been explored by a limited number of studies. This type of studies, is still warranted, especially for underexplored and underreported treatment elements that would enhance our understanding of correctional treatment. However, meta-analysis has contributed immensely to knowledge about effective treatment and these findings should be accepted and incorporated into current correctional practices. One way to translate these findings into correctional practice is to incorporate the number of elements suggested by research on treatment effectiveness in design and delivery of correctional programs and to evaluate the effectiveness of these programs. The CPAI follows evidence based correctional practice and has provided the opportunity to evaluate the potential of correctional treatment programs to reduce reoffending.

The Correctional Program Assessment Inventory

A review of the literature discussed in the previous section revealed that there is strong empirical support for the effectiveness of correctional programs. The increasing knowledge pertaining to the elements of effective programs has led to the acceptance of process evaluation as a method of choice for the evaluation of correctional treatment. Process

evaluation of correctional treatment programs assesses whether programs contain the treatment elements found to be associated with treatment effectiveness. The basic assumption underlying this approach is that programs which encompass treatment elements known to be linked to positive outcome are more likely to be successful than the programs that do not include them. The construction of the CPAI was an attempt to develop a psychometrically sound instrument that would sample indicators of the effectiveness of correctional programs (Andrews, 1995a). To our knowledge there is no any other instrument designed to evaluate the effectiveness of correctional programming.

The CPAI is designed to assess the potential of correctional programs to reduce the recidivism both in community and institutional settings. According to Andrews (1995b), this inventory can give a sense where the program” fits in terms of evidence-based effective practice” (p.27). *Psychology of Criminal Conduct* (Andrews & Bonta, 1994; 1998), with its emphasis on human service delivery, risk classification, criminogenic needs, responsivity and therapeutic integrity, shaped the content of this inventory. The development of the CPAI was and still is an ongoing process and the inventory was frequently modified to reflect the current knowledge about effective correctional treatment. The first version, the Correctional Program Inventory, was replaced by the Correctional Program Evaluation Inventory (1990) and consequently by the CPAI (1996). Although the inventory consisted of the same sections across different versions, the number of individual items increased with each new version. Copies of the previous versions are not available in a structured format as they were draft versions developed over the years. This research evaluated psychometric properties of the 6th version of the CPAI.

The CPAI (1996; 6th version) consists of items grouped into 7 subcomponents (see Appendix A for complete scoring manual):

Program demographics: provides general information about correctional programs. It includes the name of the program, its philosophy and budget, setting, years in operation, number of offenders and staff, and governmental or private agency status.

Program implementation: contains eleven items related to the program director (his/her qualifications and previous experience, whether s/he is the originator of the program and whether s/he trains staff and conducts some aspects of the programs) and program implementation (literature review and pilot program done before the program began, cost effectiveness and documented need for program; consistency with the existing values in institution/community).

Client pre-service assessment: eleven items of this subcomponent are elements closely related to program effectiveness: appropriateness of client category, a rational exclusion criteria for clients and a reasonable survey of risk, need, and responsivity factors that were measured by recognized psychometric scales with defined risk, need, and responsivity levels.

Program characteristics: this section consists of twenty-two items that assess whether the program uses adequate treatment modalities and targets criminogenic attitudes and behaviors. It also indicates how clients are prepared for their return to the community.

Staff characteristics: includes eight staff related characteristics: staff stability, their education and professional field, staff experience and personal qualities, presence of assessment, supervision and training of staff and staff input on programs.

Evaluation: seven items in this component assess program functioning based on: quality assurance control, consumer satisfaction, assessment of client on program target behaviors, and follow-up of clients with reported evaluation of program functioning.

Other: this subcomponent consists of 6 miscellaneous items related to the existence of clients' records, ethical guidelines, program change, program funding, program community support and the presence of a program advisory board.

The first subcomponent (program demographics) is non-scorable, while the remaining subcomponents are scorable. Each of the scorable items is scored as 1 (element present), 0 (element absent), n/a (not applicable), or n/k (not known). The obtained scores can be expressed in three different ways. First, they can be expressed as the percentage of elements present in each subcomponent relative to the number of elements assessed. Sections are scored as “unsatisfactory” (less than 50%), “satisfactory” (50-69%) and “very satisfactory” (over 70%). Secondly, an overall total CPAI score can be calculated. Finally, an overall CPAI Treatment score can be calculated based on subcomponents directly related to program effectiveness (client pre-service assessment, program characteristics). The advantage of this scoring method is its flexibility. Any combination of subscales can be used, depending of the research objectives.

It should be noted that there is a new version of the CPAI (2000) which incorporated new treatment elements in 8 dimensions: organizational culture, program implementation/maintenance, management/staff characteristics, client risk/need practices, program characteristics, dimension of core correctional practices, inter-agency communication and evaluation. The new items were mainly related to the core correctional practices focussing on staff characteristics and treatment practices. The data collection for this study was completed

prior to the publication of the new CPAI version therefore this dissertation focussed on psychometric properties of the CPAI (1996, 6th edition). Nevertheless, the implication of the present findings for the new version will be discussed in the final part of the dissertation.

Despite the interest that the CPAI has been attracting in recent years, there are few published studies that use this inventory. One of few studies, that used the CPAI in the evaluation of correctional programs was conducted by Gendreau, Goggin, and Annis (1990) on 170 substance abuse programs that were community and institution-based. These programs were operated by the Correctional Service of Canada (CSC) or contracted by external agencies. An earlier version of the instrument, the CPEI, had been adapted for survey purposes and sent to different program sites. The results were very discouraging regarding the percentage of program elements found at different program sites: the mean percentage of the treatment elements present for all programs was 25%, which is fairly low. The community and contracted programs scored significantly higher on the CPEI. Very few programs had follow-up data and conducted program evaluations. Gendreau et al. suggested that a possible explanation for the lack of evaluation is that 33% of the programs were within the first year of operation. These results indicated that there is a great need to evaluate programs in order to identify areas in which they should be improved.

Hoge, Leschied, and Andrews (1993) conducted a survey of 135 programs for young offenders in the province of Ontario using the Young Offender Service Inventory (YOSI), a questionnaire/interview that gathers information on different aspects of young offender programs. The CPEI was scored based on the information obtained from YOSI. The results indicated that only 10% of the programs had more than 50% of the CPEI elements present. The majority of the programs (88%) had more than 50% of elements of the other

subcomponent present, while only 13% of the programs obtained 50% or more on the evaluation subcomponent. Low scores were also obtained on the client assessment, integrity, and relapse prevention subcomponents. No significant differences on the CPEI scores were found between programs that were delivered within low and high risk facilities (facilities designated to accommodate low or high risk offenders). Higher scores on the CPEI were found for specialized facilities (facilities that have at least 50% of young offenders) compared to mixed facilities (facilities that have a minority of young offenders) and dedicated facilities (facilities that have exclusively young offenders).

Two recent studies conducted in the United States using the CPAI obtained higher scores. Latessa, Jones, Fulton, Stichman, and Moon (1999) used the CPAI to evaluate 28 of the juvenile justice programs in Ohio. They reported that 39% of the programs had less than 50% of the elements present while 10.7% of the programs had more than 70% of the elements present. The Program Evaluation and Other sections were on average scored in the very satisfactory range while the Program Characteristics, Pre-Service Assessment and Evaluation sections were mostly in the unsatisfactory range. This study provided only descriptive data on the sample included, and did not report any recidivism data of program participants

In his unpublished master's thesis Grey (1997) used meta-analytic approach to explore the link between coercion and treatment effectiveness. The link between the indicators of program quality and the effect size was also examined. The ratings of the program quality were based on the Quality Intervention Index, a 20 item short version of the CPAI. This study reported negative correlation ($r = -.17$) between coercion and effect size. Lower effect sizes were associated with the programs that used more coercion. Furthermore, there was positive relationship between the program quality ratings and effect size ($r = .31$ for Rater 1;

$r=.41$. for Rater 2). Overall, the programs that included more items from the Quality of Intervention Index, were associated with larger effect sizes.

Latessa, Holsinger, Jones, Fulton, Johnson, and Kadleck (1999) evaluated nine programs (cognitive, behavioral, and cognitive/behavioral) for juvenile offenders delivered in secured and unsecured residential facilities and group homes in Ohio. The CPAI was scored based on interviews with selected program staff and obtained scores were used as indicators of program quality. Interrater reliability indicators were not reported in this study. The CPAI scores were correlated with the recidivism rates, The analysis revealed that the higher scores led to lower recidivism rates. An examination of subcomponent scores detected variability that ranged from 40% to 77% of elements present within subcomponents. Over half of the programs scored above average on the Program Characteristics, Implementation, Program Evaluation and Other sections, while the staff characteristics and pre-service client assessment subcomponents among half of the programs were below average.

The higher scores on the CPAI and its components in Latessa, Hollsinger et al.(1999) study in comparison with programs evaluated by Gendreau et al. (1990) and Hoge et al. (1993), may suggest that in recent years correctional programs are incorporating more and more of the treatment elements that were found to be empirically linked to program efficacy. However, the small sample size ($N=9$) in Latessa et al. study and relatively short follow-up periods (3 and 6 months) limit the conclusions that may be drawn regarding program effectiveness.

Holsinger (1999) used the same set of nine programs and conducted a more detailed analysis of the link between program integrity measured by the CPAI score and different recidivism measures. Outcome measures used in his study were: any new court contact, any

new felony, new adjudication, personal offence (i.e., violent sexual, involving a weapon or arson type offences), and return to secure facility (incarceration). CPAI scores were entered into a logistic regression equation predicting treatment outcome. Significant negative relationships were found with all recidivism measures, with the exception of new court contact. More specifically, the subjects from the programs that had lower CPAI scores were significantly more likely to be arrested for new felony, have personal offence, have new adjudication and are more likely to return to secure facility than the offenders from programs that scored higher on the CPAI.

In summary, the CPAI appears to be a cost effective instrument that can give timely feedback on the quality of correctional treatment programs. Additionally, the changes in program elements can be detected by re-administration of this inventory. However, the CPAI provides only the number (percentage) of treatment elements present in a program but it is not able to address the nature of the interactions among these elements. In order to determine whether the CPAI adequately addresses the elements of effective correctional treatment, a comprehensive literature review on the factors that can influence program efficacy will be conducted in the next section. Empirical evidence for the principles of effective correctional interventions will be presented. Finally, the discussion will focus as to where the CPAI fits in terms of evidence based correctional practices.

Principles of Effective Correctional Intervention

The importance of individual differences in criminal behavior was emphasized by the interdisciplinary psychology of criminal conduct (Andrews & Bonta; 1994, 1998) which recognizes the importance of number of factors linked to criminal offending: personal, interpersonal, biological, cultural, political and situational (Andrews, 2002). The empirical

research findings from the Carleton University meta-analytic sample (Andrews et al. 1990, 1996; Andrews & Dowden, 1999; Andrews, Dowden & Gendreau, 1999; Dowden, 1998) have demonstrated that a human service approach has a positive effect on treatment outcome when offered within the justice context of community, residential and diversional interventions ($r=.12$) while the approaches using punishment led to an increase in recidivism ($r=-.03$). The lack of empirical evidence for the effectiveness of criminal sanctions, reported also by Gendreau, Goggin, Cullen, and Andrews (2002), suggests that instead of using approaches that are not effective (i.e., sanctions), correctional treatment programs should focus on the design and delivery of clinically relevant and psychologically appropriate human service. This type of service has consistently led to more effective correctional treatment programs.

Appropriate treatment defined as human service delivery based on principles of risk, need and responsivity, was found to have significantly higher mean effect size (.30) in comparison to unspecified correctional service (.13), inappropriate treatment (-.06), and criminal sanctioning (-.07) in Andrews et al. (1990) meta-analytic study. These results show that the variation in the treatment outcome can be explained by the delivery of correctional treatment based on the principles of risk, need, and responsivity. Effective correctional programs emphasize the importance of delivering service to high risk cases, targeting criminogenic needs, and using styles and modes of treatment that are matched with client needs and learning styles. Given the importance of risk, need, and responsivity empirical evidence for each of these principles will be examined separately.

Risk Principle

Risk factors are offender related characteristics whose presence can predict the occurrence of criminal behavior. According to Andrews and Bonta(1998), there is strong

empirical evidence linking criminal behavior to antisocial cognition and emotional states, antisocial peers, history of antisocial behavior, antisocial personality, substance abuse, lack of family support, weak self-control and problem-solving skills, and lack of vocational skills and jobs. The importance of these elements, especially antisocial attitudes, associates, antisocial behavior and antisocial personality identified by Andrews, Bonta, and Hoge (1990) and Andrews and Bonta (1994; 1998), in the prediction of criminal behavior was confirmed by the results of a few meta analyses (Bonta, Law, & Hanson 1998; Gendreau, Little, & Goggin, 1996; Hanson & Bussiere, 1998; Simourd & Andrews 1994). A similar set of predictors was found among adults and juveniles, among male and female offenders and among mentally disordered and sex offenders. It should also be emphasized that some of the risk factors were found to be specific for subpopulations of offenders, e.g. compliance with antipsychotic medication for mentally disordered offenders, deviant sexual arousal for sexual offenders.

The risk principle states that correctional programs should be delivered to high risk offenders (Andrews & Bonta, 1998). There is empirical evidence that high risk offenders benefit from correctional programs (e.g., Porporino & Robinson, 1992) while the impact of treatment for low risk offenders is neutral or even negative (e.g., Andrews & Friesen, 1987; Andrews & Kiessling, 1980). Meta-analytic studies have also explored the influence of risk factors on program efficacy. In a study that lacked the definition of risk criteria, Antonowicz and Ross (1994) reported that risk level is not associated with treatment outcome. However, other meta-analyses indicated that risk factors are at least moderately important for program efficacy (Andrews, 1996; Andrews & Bonta, 1998; Resource Note 10.1; Andrews et al. 1990; Dowden, 1998; Hill et al. 1991; Lipsey, 1989).

It should be emphasized that there is a limited number of studies that reported findings for lower and higher risk cases within the studies (e.g., Andrews, 1996; Dowden, 1998). The majority of studies used “aggregate estimate” to determine the risk level of offenders in meta-analytic studies by taking into consideration whether the majority of the total sample had prior criminal records and/or whether the majority of the sample had formally penetrated the justice system (e.g., Andrews, 1996; Dowden, 1998; Lipsey, 1989). The appropriate treatment was found to be most effective when delivered to high risk cases (Andrews, 1996; Hill et al. 1991). Andrews and Bonta (1998: Resource Note 10.1), using aggregate estimate approach, reported a weak although significant effect of risk factors on treatment outcome in the appropriate treatment category (.26). The obtained mean effect size within the high risk samples was .11 while low risk samples were associated with a lower mean effect size (.02). A similar pattern of results for the aggregate estimate of risk was obtained by Dowden (1998). He found a significant relationship between risk level and effect size (.17) that increased to .46 when within sample analysis was done. These results reveal that risk factors are associated with program efficacy. More specifically, programs that target high risk cases yield higher effect sizes compared to programs targeting low risk offenders.

There are some suggestions that not all of the high risk offenders (e.g., psychopaths) benefit from correctional programs (Hare, 1996). Rice, Harris, and Cormier (1992) have found that a therapeutic community increased violent recidivism of psychopaths and decreased violent recidivism of nonpsychopaths. These findings suggest that an offender’s characteristics can modify the link between risk factors and program efficacy. These characteristics will be addressed later in relation to the responsivity principles.

Need Principle

Correctional treatments can target criminogenic and/or noncriminogenic needs (Andrews & Bonta, 1998). Criminogenic needs are factors that are found to be related to criminal behavior (e.g., antisocial attitudes, values, beliefs, antisocial peers, antisocial personality, substance abuse, and self-control) and whose change is likely to influence a change in the criminal activities of offenders. The changes in noncriminogenic needs (e.g., low self esteem and emotional distress) are not likely to have a significant impact on criminal behavior as the relationship between them is weak at best.

The need principle states that correctional programs should target criminogenic needs as intermediate treatment goals. The main postulate of this principle is that the change in criminogenic needs will have an effect on recidivism reduction (Andrews & Bonta, 1998; Andrews, Bonta, & Hoge 1990; Gendreau, Cullen, & Bonta 1994). The meta-analytic results (Andrews, 1996; Andrews & Bonta, 1998: Resource Note 10.1; Andrews, Dowden, & Gendreau, 1998; Andrews et al., 1990; Cleland, 1997, Dowden 1998) reveal that targeting criminogenic needs influences program effectiveness. In their extended meta-analysis, Andrews and Bonta (1998: Resource Note 10.1) reported that programs targeting mainly criminogenic needs had higher mean effect sizes (.25) compared to the programs targeting noncriminogenic needs (.00). The role of the need principle in treatment effectiveness has been examined separately in Dowden's (1998) meta-analysis and similar results were obtained. The programs that targeted criminogenic needs (i.e., antisocial attitudes, substance abuse, antisocial peers, family supervision, self-control, vocational skills and jobs) had a higher mean effect size (.19) compared to (-.01) the programs that targeted noncriminogenic needs (i.e. self esteem, fear of official punishment, physical activity, respect for criminal

thinking). In addition, 75% of targeted criminogenic needs had a significant positive relation with effect size while the only noncriminogenic need that was significantly, although negatively, related to effect size was fear of official punishment (-.25). Both meta-analyses (Andrews & Bonta, 1998: Resource Note 10.1; Dowden, 1998) revealed that the higher number of targeted criminogenic needs was associated with higher effect sizes. The number of targeted criminogenic needs was positively related to treatment effectiveness in Andrews and Bonta's (1998: Resource Note 10.1) and in Dowden's (1998) studies (.54 and .55) while the number of targeted noncriminogenic needs was inversely related to treatment effectiveness (-.19 and -.18). The importance of targeting criminogenic needs was also reported in Cleland's (1997) meta-analysis of substance abuse programs. The correlation between appropriate treatment category and outcome dropped considerably when the influence of need principle was partialled out. The strong empirical support for the relationship between the need principle and program efficacy suggests that the more criminogenic needs that are targeted the more effective programs are likely to be. Andrews (2002) suggested further experimental research that will explore the effect of need principle on reoffending by statistically controlling for changes in needs targeted.

General Responsivity Principle

Andrews and Bonta (1998) maintain that the types of treatment influence program effectiveness (general responsivity principle). In addition, the style and mode of treatments should be matched with the learning styles of offenders (specific responsivity principle). Analysis of treatment type as a moderator variable in the meta-analyses (e.g., Andrews et al. 1990; Garrett et al. 1985; Lipsey, 1992) has resulted in some significant findings regarding the types of treatments associated with program effectiveness. One of the most extensive meta-

analyses in recent years was a meta-analysis conducted by Lipsey (1992) on a sample of 443 studies of treatment effectiveness of juvenile delinquency programs. The studies included in this meta-analysis adopted several outcome measures: recidivism rates, psychological outcome, interpersonal adjustment, school participation, and academic and vocational performance. Positive effects of the treatments were found in 64.5% of the experimental groups. The delivery of cognitive/behavioral treatments was associated with the biggest changes in recidivism reductions.

The programs based on cognitive/behavioral theoretical models were also found to be effective in reducing recidivism in a study conducted by Anotonowicz and Ross (1994). In this study, the authors examined 44 studies of rehabilitation programs for juvenile and adult offenders that had random/nonrandom designs. The results indicated that 75% of the successful programs were cognitive/behavioral in nature, while only 35% of unsuccessful programs were cognitive/behavioral. They also found that all behavioral programs that did not include a cognitive component were unsuccessful. Antonowicz and Ross (1994) concluded that the presence of a cognitive components is crucial for the effectiveness of behavioral programs. This is in accordance with reported results by Izzo and Ross (1990), Garrett (1985), Genevie, Margolies, and Muhlin (1986), and Mayer et al.(1986).

Programs with a behavioral orientation have stronger recidivism reduction potential than non-behavioral intervention, according to results of Dowden's (1998) meta-analysis. There were no additional data available regarding the type of programs using non-behavioral intervention. Within the behavioral treatment category, Dowden (1998) made a distinction among social learning, cognitive/behavioral and radical behavioral programs. Social learning programs are defined as programs that use modeling, role playing and graduated practices.

Cognitive/behavioral programs use techniques that influence feelings or behavior through thought patterns. Radical behavioral programs are based on classical/operant conditioning and token economy. Dowden (1998) reported that the social learning approach produced the strongest treatment effects. The mean effect size for social learning approach was higher (.29), in comparison to cognitive/behavioral (.21), radical behavioral (.16), and non-behavioral (.04) treatment.

Andrews et al (1996) reported that the higher mean effect sizes were obtained for human service (.34), and behavioral/social learning service (.47). The human service that included behavioral approaches had a greater mean effect size (.25) than human service that did not use a behavioral technique (.08). The same pattern of results pertaining to the effective treatment types was obtained in Cleland's (1997) meta-analysis. The treatment programs using behavioral or cognitive behavioral strategies produced significantly higher mean effect sizes than programs that used non-behavioral modes of intervention. In summary, the results of different meta-analyses suggest that theoretically well founded cognitive behavioral, social learning, and radical behavioral treatments are likely to be effective.

Specific Responsivity Principle

The specific responsivity principle takes into account not only the type of programs, but also offenders' and staff characteristics in an attempt to optimize treatment. Programs delivered to high-risk offenders that target a number of criminogenic needs might not be effective unless the type of treatment and staff characteristics are matched with the learning styles and characteristics of offenders. In order to improve program efficacy, treatment should be matched with learning styles and personality of offenders, offenders matched with

therapists and therapists matched with the type of program (Gendereau, Cullen, & Bonta, 1994).

Specific responsivity factors are not necessarily treatment targets, they are offender characteristics that can influence the treatment process if not addressed adequately (Andrews & Bonta, 1998; Bonta, 1995). Some offender characteristics (e.g., depression, anxiety) although not strongly linked to criminal behavior, can determine an offender's response to treatment. Kennedy and Serin (1995) grouped specific responsivity factors into cognitive (inadequate problem solving skills, concrete oriented thinking, poor verbal skills), internal (poor social skills psychopathy, motivation, self esteem, anxiety, depression and mental disorder), demographics (age, gender, race, ethnicity) and external (setting and therapist related characteristics). Bonta (1995) maintains that responsivity factors should not be equated with risk and need factors, although they can overlap in some cases (e.g., psychopathy can be both a risk and responsivity factor while one of its components, impulsivity can be a criminogenic need).

The usefulness of matching clients with the provided treatment, the treatment providers and the treatment environment has been reported in relation to Interpersonal Maturity level (Harris, 1998; Van Voorhis, 1985), Conceptual-level (Brill, & Reitsma, 1980, Reitsma-Street & Leschied, 1988; Leschied, Jaffe, & Stone, 1985; Reitsma-Street, & Street 1981 Reitsma-Street, 1984) and motivation (Stewart and Milson, 1995; Baxter, Marion, & Gogen, 1995; Gillis & Grant, 1996). The offender's characteristics and learning styles should be taken into account while designing and delivering the program in order to enhance its effectiveness. Until now, the meta-analytic research did not systematically address the issue

of specific responsivity principle, mainly as a result of small number of primary studies examining the relationship between specific responsivity and treatment outcome.

Some authors (e.g., Losel, 1993) maintain that correctional programs should also take into account protective factors (i.e., the factors whose presence can lead to a more positive outcome even when unfavorable conditions are present). Originally identified in psychiatric and developmental psychology literature (e.g., Rutter, 1985, 1990, Cohler, 1987, Masten, Best, & Garmezy 1990) their influence on the occurrence of criminal behavior have been explored recently (e.g., Bender, Bliesner, & Losel, 1996; Hoge, Andrews, & Leschied, 1996). The strong link between protective factors and criminal behavior was suggested by some authors (e.g., Bender et al. 1996; Farington & West, 1993; Stouthamer-Loeber et al., 1993) while several authors (e.g., Garmezy, Masten, & Tellegen, 1984; Luthar, 1993; Rutter, 1987) even maintain that there is an interaction between risk and protective factors. There is no empirical evidence that the presence of protective factors will change the effect of risk factors on criminal activity. Nevertheless, the protective factors defined as personal and interpersonal strengths (e.g., intellectual level, family and social support, good coping strategies) could enhance treatment effectiveness and the new generation of risk/need instruments is taking into account the assessment of strength factors (Andrews, 2002).

Appropriate Treatment

Dowden (1998) created the New Appropriate Treatment, a composite variable based on counting the number of principles (risk, need, and responsivity) that programs adhered to. The results of this meta-analytic study shows that the mean effect sizes for the combinations of the risk, need, and responsivity principles were higher than the mean effect sizes obtained for individual principles. The higher number of principles included in the treatment is linked

to higher mean effect sizes (-.02 for none of principles present, .02 for one principle present, .18 for two principles present, and .26 for all three principles present). Each of these principles contributes independently to effect size when the other two principles were controlled for. The need principle had the highest and the risk principle the lowest contribution to effect size. Andrews, Dowden, and Gendreau's (1999) meta-analytic study indicated that the contribution of human service, risk, need and general responsivity were incremental in predicting effect size. While the multiple correlation with all four variables was .58, the partial correlations were .07 (Any human service); .15 (Risk); .37(Need) and .18(General Responsivity). Need principle was confirmed as the most important principle. The New Appropriate Treatment variable was tested under variety of control condition, and it remained the most powerful indicator of treatment effectiveness. These results suggest that each of these principles should be adequately applied in order to maximize treatment effectiveness.

Andrews (2002) argues that the assessment of risk/need/responsivity factors should be conducted in a systematic way as the adherence to the principles of effective programming will depend upon the reliable and valid assessment for the assessment of risk/need/responsivity principles. Assessment tools are needed to identify appropriate intermediate treatment targets and to guide intensity of treatment interventions.

The setting in which treatment is delivered can influence program effectiveness. While some studies (Anotonwicz & Ross, 1994) reported no difference in effectiveness between institutional and community treatments, other meta-analyses (Andrews, 1996; Andrews, Dowden, & Gendreau, 1999; Andrews et al. 1990; Dowden, 1998; Izzo & Ross, 1990; Lipsey, 1992; Withead & Lab, 1989) have found higher effect sizes for community as compared to institutional programs. More specifically, Andrews et al. (1990) found that in community

settings the effect of appropriate treatment was higher (.38) while the effect of inappropriate treatments in institutional settings was more negative (-.10).

The consideration of the influence of setting type on treatment effectiveness needs to take into account a few additional methodological issues. Losel (1995) maintains that both institutional and community treatments are broad categories that include different programs. Programs offered within the community vary from diversion to electronic monitoring, while institutional treatments range from maximum security to halfway houses. There are even some overlapping treatments in between. In addition, the negative influences of institutions could be minimized by many factors (presence of positive institutional and educational climate, staff characteristics, and offender personality). These elements were too specific and were not taken into consideration although they can have a potential effect on treatment efficacy.

The presence of risk, need, and responsivity factors is not a sufficient condition for program efficacy. There is a wide range of “background variables” that are not directly related to the effectiveness of programs but whose presence enhances programs and indirectly influences their effectiveness. This means that the programs that incorporate risk, need and responsivity factors are more likely to be effective if these “background” elements are also present. These variables are frequently referred to as treatment integrity variables. The concept of treatment integrity will be discussed separately in the next section.

Treatment Integrity

Controlled outcome studies have often identified program content and procedures as a source of variance in treatment effectiveness while minimizing the relevance of other variables (Kazdin, 1986). Program content and procedures although important, are likely to

interact with other variables and jointly influence program efficacy. The search for an optimal set of variables that will enhance programs is not unique to the correctional field. The literature has numerous examples of well conceived but poorly implemented programs or programs poorly sustained but well implemented (Leschied, 2002). Treatment outcome literature from psychotherapy, medicine, clinical psychology, and social work has also been influenced by the research on treatment fidelity as a potential factor contributing to treatment efficacy (Moncher & Prinz, 1991). Treatment fidelity relates to the variables that could influence the internal, external validity and statistical power of outcome research although they are not closely related to program content and procedure (Yeaton & Sechrest, 1981). Treatment fidelity consists of two distinct categories: treatment integrity and treatment differentiation. Treatment integrity is not “an all or nothing phenomenon” (Kazdin, 1986) and it is not an unitary construct (Dobson & Shaw, 1988; Moncher & Prinz, 1991; Quay 1977, 1987). It refers to the degree to which treatment is implemented as intended (Yeaton & Sechrest, 1981). Treatment differentiation refers to whether treatments can be differentiated from other treatments (Dobson, & Shaw, 1988; Kazdin, 1986). Moncher and Prinz (1991) conducted an extensive evaluation of 359 treatment outcome studies from clinical psychology, behavior, marital and family therapy and psychiatry that were published from 1980 to 1988. The purpose of this study was to establish the extent to which investigators are adequately attending to the issues of treatment fidelity. Results indicated that 55% of studies ignored treatment fidelity. Although in the more recent studies increased attention was paid to these elements, only one out of eight studies is checked for treatment fidelity. The latter refers to the extent to which a program is implemented and delivered as intended

The empirical evidence suggests that once treatment begins its integrity needs to be addressed. Only a few studies addressed and/or assessed whether treatment is delivered as planned (e.g., Wormith, 1984). An increased interest in issues of treatment integrity has been noticed in recent correctional literature on treatment effectiveness (e.g., Davidson, Redner, Blakely, Mitchell, & Emshoff, 1987; Hollin, 1996). The results of Hill et al. (1991) meta-analysis revealed that the influence of program integrity was minimal when assessed individually but its presence increased the effectiveness of appropriate treatments and decreased the effectiveness of inappropriate treatments. Some authors (e.g., Gordon, Andrews, Hill, & Kurkowsky, 1992 as cited in Andrews & Bonta, 1998) suggest that therapeutic integrity can contribute to treatment effectiveness even when appropriate treatment is controlled for. Andrews and Dowden's (1999) meta-analysis explored the influence of program integrity on recidivism reduction. The indicators of program integrity in this meta-analysis were: involved evaluator, the presence of a specific model, selection of workers, training of workers, clinical supervision of workers, training manuals, monitoring of service process and/or intermediate gain, adequate dosage, new/fresh program and size of program. These indicators of treatment integrity (except monitoring and adequate dosage) were positively associated with effect sizes and with appropriate treatment category. When appropriate treatment is controlled for, specific model, selection of staff, training of staff, clinical supervision of staff, involved evaluator and sample size remained significantly associated with effect size. However, none of the indicators of integrity was correlated with effect size under inappropriate treatment conditions. These results show that inappropriate treatment, even if accompanied with a high level of program integrity is not significantly correlated with treatment outcome. Number of treatment elements, whose presence can

influence program integrity and have an impact on program efficacy, will be discussed in the next sections.

Treatment drift and non-compliance are considered to be possible threats to treatment integrity (Hollin, 1995). Program drift refers to the digression from planned treatment techniques and procedures by program delivery staff. Johnson (1981) evaluated the effectiveness of a therapeutic community in a pretrial detention center. His results suggested that the nature of staff-inmate interaction tended to shift over time. Staff drifted from the therapeutic procedure and focused their attention on the immediate needs of prisoners. This study identified program drift as a potential threat to the integrity of treatment, although it did not determine the extent of its influence on program efficacy.

Program non-compliance refers to the omission and/or change of program elements by delivery staff. In a study that evaluated the effectiveness of cognitive multisystem therapy Henggeler, Melton, Brondino, Scherer, and Hanley (1997) explored treatment adherence. The authors identified two elements as potentially linked to program effectiveness. First, low therapist's adherence to multisystem principles was associated with higher arrest rates. Second, non-productive therapy sessions were associated with more arrests. These results suggest that the therapy sessions and their dynamics should be monitored. Raynor and Vanstone (1996) recommended monitoring program sessions in order to determine whether staff have followed program design and whether any digressions from planned intervention have occurred.

Intensity and duration of treatment are elements of the treatment process that can be linked to treatment integrity and program efficacy. There are some indications that the relationship between risk level and program efficacy can be modified by the intensity and

duration of treatment (Eisber & Fabelo, 1996; Wexler, Falkin, & Lipton, 1990; Wormith 1984). For example, the programs that last from 3 to 9 months, in which high risk offenders are spending 40-70% of their time, are likely to be more effective according to Gendreau et al. (1994). The effects of duration and intensity of treatment on its efficacy were explored in several meta-analyses. While the results of the meta-analysis done by Izzo and Ross (1990) indicated that the duration of treatment was not significantly related to its effectiveness, other meta-analyses (Gensheimer et al. 1986; Gottschalk et al. 1987a,b; Lipsey, 1995) have reported that a longer duration of treatment was associated with a more positive outcome. Lipsey (1995) argued that both duration and dosage of treatment were associated with treatment outcome (i.e., high dosage treatments that last longer than 26 weeks, that have two or more contacts per week and/or more than 100 hours of total contact are likely to be more effective). In Andrews' (1996) meta-analysis, the dosage of treatment (assessed as the total number of contact hours) was found to be related to effect size (.38). However, in additional analysis of program integrity Andrews and Dowden (1999) found that when appropriate treatment was entered into the regression analysis, the link between the dosage and effect size was no longer significant.

The degree to which the duration of a program influences its effectiveness was explored by Wexler et al. (1990) study. The results of their study will be presented in more detail because they provide information that meta-analytic studies have not taken into account. The authors evaluated the effectiveness of the "Stay n'out" program in reducing recidivism (assessed as parole violation and time until arrest). The participants were divided into five groups based on the amount of time they spent in treatment. The general data pattern revealed that longer program duration has led to an increased program efficacy (less parole

violation and longer time until arrest). However, the positive influence of program duration was not detected after one year. The authors offered two possible explanations for this drop in program efficacy after one year. First, during the course of this program offenders were assuming increasing responsibilities, and by the end of the first year offenders had usually achieved the highest responsibility level with no possibility to progress further. Second, few offenders within the program had parole denied at that time. Although these results can suggest that adequate duration of treatment is not the same for all programs and that different types of programs might have a different optimal time, the lack of control for risk factors in this study limits the generalizability of these results.

Effective rehabilitation programs are guided by theory, according to Martin, Sechrest, and Render (1981) as the atheoretical approach to ad hoc program development usually leads to poorly conceptualized and implemented programs. Antonowicz & Ross (1994) reported that effective programs are associated with a sound conceptual model. According to these authors, theoretical conceptualization of criminal behavior that is associated with program design determines the choice of intervention and techniques that are used. Andrews and Bonta (1998) maintain that in order to influence criminal behavior, correctional programs should be based on psychologically relevant and empirically sound theory of criminal conduct and suggest that research on treatment effectiveness should be guided by theoretical considerations. Andrews et al. (1996) explored the influence of the specificity of treatment model on treatment efficacy. The results suggested that it was negatively related to effect sizes - .23 for treatment in general and .30 for appropriate treatment. However, its contribution was not significant beyond the contribution of appropriate treatment category. In Andrews and Dowden's (1999) meta-analysis, the specificity of the treatment model remained significantly

associated with effect size even when control is introduced for appropriate treatment. The importance of specificity of treatment model for program effectiveness was also reported by Lipsey (1989).

The threats to treatment integrity can be reduced and/or eliminated by using treatment manuals (Dobson & Shaw, 1988). These authors reported that while prior to 1977 hardly any programs had treatment manuals, in recent times many programs have used them. The main advantages of using manuals according to Dobson and Shaw (1988) are that they facilitate staff training and enhance internal validity of research studies by ensuring that a specific set of identifiable treatment procedures can be repeated. Hollin (1995) points out that treatment manuals contain guidelines for planning program sessions, their content and the techniques/procedures to be used. The general treatment literature has revealed that the use of treatment manuals enhances program efficacy (e.g., Moncher & Prinz, 1991). Some meta-analytic studies (Andrews 1996; Andrews & Dowden, 1999; Dowden 1998; Lipsey, 1989) have found that printed manuals are associated with program effectiveness. However, when treatment type was controlled for, the contribution of program manuals to effect size was not significant (Andrews 1996; Andrews & Dowden, 1998; Dowden, 1998).

The involvement of an evaluator/researcher in program design and implementation can lead to the higher treatment integrity and more successful programs (Lipsey, 1995). There is evidence suggesting that evaluator involvement in program design/delivery is associated with higher effect sizes (Andrews, 1996; Andrews & Dowden, 1998; Andrews, Dowden, & Gendreau, 1998; Dowden 1998; Gensheimer et al. 1986; Gottschalk et al. 1987). In addition, there is evidence indicating that the combination of treatment type and evaluator/researcher involvement will result in programs that are more effective than programs that did not have an

involved researcher (Andrews, 1996; Hill et al., 1991; Lipsey, 1992). The relationship between evaluator involvement and effect size was present even when treatment modality was controlled for (Lipsey, 1992). According to Andrews and Dowden's (1999) and Andrews et al. (1998) meta-analytic results, having an involved evaluator was significantly associated with effect size even when appropriate treatment is controlled for.

Andrews and Bonta (1998) argue that agency/community conditions can be a source of variability in treatment outcome. In her dissertation Leddermen (1986) explored the conditions under which consultation within an agency setting is successful. A sample of consultants was asked to identify the elements that contributed to the success of the consultation process. The results suggested that successful consultation is associated with the agency that is stable and supportive. Consultants need to have strong support of management and good communication across different departments within the agency. In addition, agencies with good assignment of resources will also have a positive influence on consultation. Andrews (2002) indicates that the effectiveness of a program will be reduced if the host agency is not understanding or interested in clinically relevant approaches. The influence of these elements on correctional treatment efficacy has not been explored in detail in the controlled outcome studies. However, surrounding agency/community conditions can influence program implementation, the number of treated clients and the procedures for hiring staff (Andrews, 1995b).

Staff Characteristics

Staff related variables were identified by many authors (e.g., Andrews & Bonta, 1998; Andrews et al. 1990; Andrews, 1996; Hollin, 1995; Losel, 1995) as related to treatment

outcome. There are two broad categories of staff related variables that can influence treatment effectiveness: interpersonal factors and skill factors

The relationship principle of treatment effectiveness (Andrews & Bonta, 1998) states that interpersonal relationship elements are important in influencing behavior. The evidence from both clinical experience and empirical studies suggest that staff interpersonal qualities influence the success of a program. Many staff related characteristics and attributes were linked to program success: for example, interpersonally warm, tolerant, flexible, democratically authoritative (Andrews & Kiessling, 1980), emphatical (Hengeler, Melton & Smith, 1992; Wexler, Falkin & Lipton, 1990), interpersonally sensitive (Genderau et al., 1994), warm, genuine, with sense of humor, enthusiastic, self confident, empathic, respectful, flexible and mature (Andrews & Carvell, 1998) staff. In Dowden's (1998) meta-analytic study, programs in which staff were selected based on their characteristics and relationship factors were associated with higher effect sizes (.34) than programs that did not report such selection practices (.07). The influence of the selection of staff based on relationship factors on program effectiveness was explored in Andrews and Dowden's (1998) meta-analysis. The association between the selection of staff and effect size remained significant even when appropriate treatment was controlled for.

The contingency principle maintains that the content and direction of what is learned influences offender's behavior. As the majority of the interaction within prison is found to be procriminal (Buehler, Patterson & Furniss, 1966), the programs should be based on anticriminal modeling (Andrews & Bonta, 1998; Andrews & Kiesling, 1980, Wexler, Falkin, & Lipton, 1990), and oriented towards problem solving (Andrews & Bonta, 1998; Wexler, Falkin, & Lipton, 1990). Andrews and Carvell (1998) maintain that when staff exhibit

directive, solution focused, structured, and nonblaming behavior, and when staff use effective reinforcement, effective disapproval, prosocial modeling, problem solving, rehearsal, coaching and cognitive restructuring, the treatment outcome is likely to be more positive. Some of these elements were tested in Dowden's (1998) meta-analysis. The results revealed that programs reporting the presence of these factors were associated with significantly higher effect sizes (.27) compared to the programs that did not report them (.06). The higher effect sizes were associated with programs that used core correctional practices of effective modeling (.29), used reinforcement (.31), effective disapproval (.18), problem solving (.19) and structured learning procedures (.30). These effect sizes were found to be significantly higher when compared to groups that did not report these practices. When composite measures were taken into account, programs using one or more of these principles were associated with higher effect sizes (.25) compared to the rest of the programs (.06).

The influence of staff training on treatment effectiveness has been debated recently. Gendreau et al. (1990, 1994) argue that programs whose staff are trained in behavioral treatment interventions are more likely to be effective. Although some studies did not find evidence that trained staff (Henngeler, Melton, & Smith, 1992) and trained volunteers (Wormith, 1984) have an impact on program efficacy, the results of few meta-analyses (e.g., Andrews, 1996; Andrews & Dowden, 1999; Dowden, 1998) indicate that training of the staff is associated with treatment effectiveness. Andrews's (1996) meta-analysis explored the influence of trained therapist, supervised therapist, and trained supervisor on treatment effectiveness. Each of these elements was found to be mildly related to effect sizes (.12 for trained therapist, .25 for supervised therapist, and .10 for trained supervisor). The association of these elements with effect sizes was slightly higher within the appropriate treatment

category, but none of them contributed significantly to treatment outcome beyond the contribution of the appropriate treatment category. Dowden (1998) found that the presence of trained supervisors and trained front line staff was associated with higher effect sizes but was not important once appropriate treatment was taken into account. However, the meta-analysis done on the same set of studies that excluded the studies addressing the effects of punishment (Andrews & Dowden, 1998) revealed that the presence of trained staff and staff that are clinically supervised is significantly and positively associated with effect size and the appropriate treatment category. In addition, when appropriate treatment is controlled for, the contribution of these elements remained significantly associated with effect size.

Although the importance of treatment delivery by paraprofessionals has been debated (Berman & Norton, 1985; Durlak 1981) some effective programs have used paraprofessionals with adequate training (Davidson et al., 1987; Emshoff & Blakely, 1983; Gordon & Arbuthnot, 1988). The existence of detailed manuals for the training of paraprofessionals was reported in several studies (Bush 1995; Davidson et al. 1987; Henning & Frue, 1996).

Literature on treatment effectiveness reported the importance of some program elements for treatment efficacy but the empirical support for these elements is still limited. Quay (1987) has argued that staff's belief that the programs have an impact on recidivism could be an element influencing treatment integrity and program effectiveness. He analyzed Kassebaum, Ward, and Wilner's (1971) unsuccessful counseling program. It was found that more than two-thirds of the staff delivering the program did not believe that the program would lead to recidivism reduction and that the program leaders were not adequately trained. According to Quay (1987), these are the reasons that could potentially explain why this program was not successful. Andrews (2002) summarized staff related characteristics that

were underexplored in correctional literature. He suggested that they should be included in the treatment effectiveness studies: support of human service activity, belief that core correctional practices work, belief that offenders can change, belief that staff have skills to deliver programs effectively and that correctional treatments reduce reoffending of offenders.

Summary

What is known at the present? The programs that use principles of risk, need, and responsivity were found to be effective among different offender populations and across different settings. More specifically, programs that target a higher number of criminogenic needs using radical behavioral/social learning/cognitive behavioral treatments delivered to high risk offenders are more likely to be effective. At the same time, the presence of these “fundamental” elements is not a sufficient condition for program effectiveness because treatment outcome can be influenced by too many unforeseen variables. Factors related to treatment integrity can moderate the effects of risk, need, and responsivity factors on program efficacy and influence the way in which offenders will respond to treatment. The combination of these elements can create the optimal treatment conditions for different groups of offenders in which an offender’s response to treatment will be maximized. Different authors use different characterizations of treatment integrity, but in general treatment integrity can be influenced by treatment process related, program related, and agency/community related treatment elements. These elements influence program efficacy by providing necessary conditions for program effectiveness.

Higher intensity and dosage treatments delivered to high risk offenders are more likely to be effective. There is some indication that productive treatment sessions and the adherence to treatment procedures and techniques can also influence treatment outcome. Programs that

are based on a sound theoretical model, that are delivered within the community, that have the evaluator/researcher involved in the program design and implementation, and that use treatment manuals in which contents and procedures of treatments are outlined, were found to be associated with positive treatment outcome. A wide range of staff related characteristics are found to be important for treatment effectiveness. Staff that has training in behavioral interventions, and that is clinically supervised seems to be present in more effective programs. Warm tolerant, flexible, democratically authoritative, interpersonally sensitive, emphatical, enthusiastic are interpersonal characteristic found to be related to positive treatment outcome. In addition, skill related factors could also influence program effectiveness: using effective modeling, reinforcement, effective disapproval, problem solving and structured learning procedures. The agency/community that is supportive of treatment, has good communication among program and managerial staff, and has a productive exchange of information among different organizational levels can improve program, treatment process and staff related characteristics and have an indirect influence on treatment effectiveness.

In his 2002 study, Andrews emphasized that integrity depends on and is enhanced when a specific, empirically sound theory is employed, when printed manuals are available, when staff selection, training and supervision take into account skills and characteristics required for effective program delivery, when the clinical supervisor is trained, when intermediate targets are achieved, when the dosage is adequate when researcher is involved in the design, delivery and evaluation of the program. Research findings demonstrate (Andrews & Dowden, 1999) that although many treatment integrity indicators are positively linked to treatment outcome, just few of them have a significant contribution to program effectiveness beyond the combined influence of risk, need, and responsivity principles (i.e., involved

evaluator, staff selected by relationship principle, trained staff, staff supervised by trained supervisors and the presence of a specific model). When appropriate treatment was delivered the mean effect sizes were significantly higher for high than low integrity conditions. The link between the indicators of program integrity and the effect size was not found under inappropriate treatment conditions, suggesting that inappropriate treatment even if delivered under a high program integrity will not contribute to treatment effectiveness. Overall, the contributions of program integrity were limited to enhancement of appropriate human service based on the principles of risk, need and responsivity.

The effectiveness of correctional programs can be partly explained by the effects of risk, need, and responsivity factors. However, the individual contributions of risk, need, and responsivity to program efficacy can be deflated if the program was not implemented and delivered as intended, i.e. unless the program has satisfactory treatment integrity. The way that program elements interact is still not clear although some authors (Henggeler, Melton, & Smith, 1992) maintain that the interaction is probably synergetic, rather than additive. Nevertheless, two or more treatment elements when joined might or might not function well together. The ultimate goal of the search for treatment elements that are associated with program efficacy, is the identification of a “good” combination of treatment elements. Presently, a considerable number of treatment elements whose presence will enhance program efficacy has been identified, although the nature of their interactions remains unknown.

Despite many treatment elements reported to be linked to treatment outcome, the final decision about treatment itself should be given to a therapist. The presence of all “effective” elements within programs will not make a program effective unless the program is designed to correspond to individual offenders. Ultimately, offender related factors determined by both

objective assessment and clinical interviews, should be the factors that will determine the choice of the program delivery staff. The specific responsivity principle seems to have a central role in program efficacy, but is unfortunately less explored compared to risk and need principles.

Today, there is an extensive knowledge about the effective treatment elements and the correctional programs should be designed and implemented based on this empirical evidence. At the same time, there is a need for a consistent and continuous effort to evaluate the effects of the interventions based on the treatment effectiveness literature. The CPAI is an instrument that could be used not only for program design but also for the evaluation of ongoing correctional treatment programs. The review of the literature on program effectiveness presented in the previous sections indicates that the CPAI items cover the majority of the treatment elements that were identified as being associated with program effectiveness. The Preservice Assessment and Program Characteristic subcomponents of the CPAI received the strongest empirical support. This was expected, given that these subcomponents were derived from the main postulates of the “Psychology of Criminal Conduct” (Andrews & Bonta, 1994, 1998). Meta-analytic studies (Andrews, 1996; Andrews & Bonta, 1998: Resource Note 10.1; Andrews et al. 1990; Andrews & Dowden, 1999; Andrews, Dowden & Gendreau, 1999; Cleland 1997; Dowden, 1998; Hill et al. 1991) reported that risk, need and responsivity factors were significantly associated with treatment effectiveness. In addition, the Program Characteristics section of the CPAI (6th edition) received strong support from the same meta-analytic studies that identified number of treatment integrity indicators contributing to the treatment effectiveness (e.g., intensity durations of the program, matching staff, clients and programs, presence of printed manual). The Staff Characteristics section of the CPAI also

covered the majority of the staff related characteristics that were found to be associated with program efficacy: interpersonal qualities and skills, clinically supervised and trained staff. The three remaining components of the CPAI (Program Implementation, Program Evaluation and Other) were not reported to be strongly related to treatment outcome. However, they can improve treatment integrity and indirectly enhance treatment effectiveness. For example, involved evaluator was found to contribute to treatment effectiveness beyond contribution of appropriate treatment.

Although the CPAI (6th edition) items cover the main factors linked to the program efficacy, several treatment elements that can potentially influence treatment outcome were not included: First, motivation for change among offenders can be included not only as an intermediate treatment goal but also as a special responsivity consideration. Besides the elements of risk, need, responsivity, and treatment integrity that have relatively strong empirical support, there is a group of motivational factors that can potentially influence treatment outcome. Changes in the motivation for criminal behavior can potentially lead to changes in the criminal activities of offenders. Additionally, the degree of motivation for treatment can potentially determine the extent to which an offender's criminal behavior will change. Furthermore, all of the program integrity related factors and staff related treatment elements can have an effect on motivation for change of offenders. Unfortunately, there is no empirical evidence on the extent to which motivational change can influence program effectiveness. Second, the presence of treatment adherence was suggested by Henggeler et al. (1997) and Raynor and Vanstone (1996) to be related to treatment integrity. Third, productive therapy sessions can be considered as important part of an effective program (Henggeler et al. 1997). Fourth, setting in which treatment is delivered was found to be important (Andrews et

al. 1990; Dowden, 1998; Izzo & Ross, 1990). The nonscorable demographic subcomponent of the CPAI included this element. However, it is suggested that this element be included in the program subcomponent of the CPAI that is scorable. Fifth, staff beliefs in the recidivism reduction potential of correctional programs are suggested by Quay (1987) as potentially important for program efficacy. Sixth, smaller treatment groups with higher staff ratio are potentially important for program effectiveness according to Jesness (1971/1972). Seventh, favorable conditions regarding treatment delivery within agency/institution/community can be associated with positive treatment outcome (e.g., Ledermen, 1986).

The influence of these program elements is probably limited to the development of optimal conditions for program delivery. None of these elements, with the exception of setting, was systematically explored in controlled outcome studies and the extent and ways in which they can influence treatment outcome remain to be determined. In summary, recent empirical evidence has identified a number of treatment elements that can explain variations in treatment outcome and the majority of these elements were summarized in the CPAI, as described in the introductory section.

The Present Study

The CPAI is based on the process evaluation of correctional programs and it covers a wide range of treatment elements that were found to be related to program efficacy. The purpose of this inventory is to assess the recidivism reduction potential of correctional treatment programs. The review of the program effectiveness literature conducted in the previous sections suggested that although the CPAI (6th edition) has covered the majority of the treatment elements found to be linked to program efficacy, a few elements that can be potentially important were omitted (i.e., motivation for change, treatment adherence,

productive therapy sessions, setting, staff beliefs in recidivism reduction potential of programs, size of treatment group, and favorable agency/community conditions). However, the new version of the CPAI (2000) have included majority of these elements. These additional items will be included in reliability and validity studies that will be described later in the text.

The continuing use of this inventory in program evaluation and the possible implications of its results for program delivery require the determination of satisfactory psychometric characteristics of the CPAI. The purpose of this dissertation is to assess the reliability and validity characteristics of the CPAI. The first part of the study will mainly focus on assessing interrater reliability and will also explore different scoring methods of the CPAI. Face and content validity will be explored in the second part of the study while predictive validity will be addressed in the third part. Finally, summary of the findings and suggestion for future research will be given.

Study 1: Interrater and Intermethod Reliability

Goals

The main goal of Study One was to assess the inter-rater reliability of the CPAI. Internal consistency reliability, expressed as the relationship between individual items and the total score, was also explored. Since the CPAI is still in the developmental phase, an additional research question pertaining to different scoring and administration methods, was also addressed in this study. So far, the CPAI has been scored in three different ways: reviewing program manuals and program files, interviewing the program directors during site visits (e.g., Latessa, Jones, Fulton, Johnson et al., 1999), and finally by sending surveys to different program locations to be filled-out by program staff (e.g., Gendreau et al., 1990). This study compared the interview and survey based scoring approach.

Method

This research was conducted in two phases. In the first phase the CPEI questionnaire used by Gendreau et al. (1990) was modified for the purpose of this study and a new questionnaire the Correctional Program Assessment Inventory Questionnaire (CPAI-Q) was constructed. A number of treatment elements identified in the previous sections as potentially important for program effectiveness, were also included. Prior to conducting the study, the CPAI-Q was piloted. Two program directors and three treatment staff were asked to review the questionnaire and note any problems they might have in answering or understanding questions. As a result of their feedback minor modifications to the questionnaire were made. The final version of the CPAI-Q can be found in Appendix B. The CPAI-Q was sent to twenty program locations. The program directors were asked to fill out the questionnaire and return it to the researcher within a month.

In the second phase, the researcher visited the same program sites and conducted an audiotaped CPAI based interview with the program directors and treatment staff. When available, treatment manuals and accompanied documentation were reviewed. There was on average two month period from the time that the questionnaires were sent to the time the interviews were conducted. Two set of raters scored the CPAI based on a) answers on the CPAIQ and b) audiotapes of the interviews.

Sample

For the purpose of this study correctional treatment programs are defined according to Whithead and Lab (1989) as any intervention aimed at reducing recidivism. The inclusion criteria specified that the programs should have been a) in existence for at least one-year b) delivered under federal or provincial jurisdiction c) delivered in the institution or in the community and d) delivered to male or female offenders.

The programs that have met selection criteria were identified and selected from the comprehensive list of programs provided by the Ministry of Correctional Services (MCS) and the Correctional Service of Canada (CSC). Every tenth program from the list of provincially run programs was included. A number of the programs delivered at the federal level could not be included in the sample due to operational issues (e.g., low staff numbers) and the Regional Trainer identified the programs that would likely be able to participate.

All selected programs agreed to participate and the questionnaires (CPAI-Q) were sent to the program sites. Within two weeks eleven programs (55%) contacted the principal researcher inquiring whether the CPAI-Q could be completed by the treatment team as the questions were too detailed to be answered by the program director as originally requested. The instructions were revised and the treatment team at the individual program sites identified

the most appropriate team members for the completion of the questionnaire. All program sites also agreed to participate in the second part of the study in which the interviews with the program director and (or) treatment staff were conducted and audiotaped. However, originally planned third part of the study in which program documentation would be reviewed and the CPAI scored based on the documentation could not be implemented. Notwithstanding the availability of some printed documents at the program sites, the majority of the documentation pertaining to design/delivery was located in administrative headquarters and would have required considerable staffing resources and time to identify relevant materials. Furthermore, staff related confidentiality issues were an obstacle in obtaining an access to hiring/training/evaluation information. This project has already used considerable staff resources, and after consulting with administrative managers at the MCS and CSC, it was decided that the principal researcher would have an access to the documents available at the individual program sites.

The final sample consisted of twenty programs with half of the programs being delivered at the provincial level and the other half at the federal level. Sixty five percent of the programs were institution based, while the rest were community-delivered programs. Among participating institutions, the majority (55%) were medium secure facilities while maximum and minimum-secure facilities were equally represented (23%). The majority of the programs were delivered to adult, male offender populations while the rest was delivered to female (three programs) and young (one program) offenders. See Appendix C for a complete list of program sites.

Hypotheses

Hypothesis 1: It is expected that the CPAI will have high interrater and internal consistency reliability. A systematic and objective scoring method has been developed and the ratings of the individual CPAI items are likely to be independent of rater related variables.

Hypothesis 2: It is hypothesized that the different administration and scoring methods will yield different total, treatment, and CPAI sections scores. Notwithstanding an increasing number of programs that have comprehensive program manuals/files, the amount of information contained in this documentation might not be sufficient to score the CPAI. The questionnaire based scoring method will likely provide more information on treatment elements present within individual correctional programs. However, this method might score the way in which the programs were reportedly delivered but not necessarily the way they are actually delivered. It is possible that these two methods will reflect the “paper quality” of the programs while an in-depth interview with program director/treatment staff will provide more information on actual delivery of the program.

Results

Scoring the CPAI-Q Responses

The principal researcher and second rater independently scored the CPAI-Q using the CPAI scoring manual. The second rater was an Honors student in Psychology who had undertaken this research as a part of her thesis and was trained in scoring the CPAI by the principal researcher. Two questionnaires were scored and scoring differences compared and discussed. After the clarification of scoring difficulties, the raters scored the rest of questionnaires.

Based on the responses on the CPAI-Q, treatment elements were scored as present (1) or absent (0). The score not known/not applicable (8) indicated that information was either not applicable for particular program or absent/ not sufficient to score the item. Table 1 summarizes frequencies and percentages of coding categories given by two raters for all items (72) across all programs (N=20).

Table 1:
Frequencies and Percentages of Coded Categories (CPAI-Q)

Coded categories	Rater 1		Rater 2	
	f	%	f	%
“0”	458	31.8	422	29.3
“1”	730	50.7	729	50.6
“8”	252	17.5	289	20.1
Total	1440	100	1440	100

The results indicate that the raters coded a similar number of items as present (50.7% Rater 1 vs. 50.6% Rater 2) and absent (31.8% Rater 1 vs. 29.3% Rater 2). There was a considerable number of items that were coded as “8” (i.e., not known or not applicable). Rater 1 coded 252 (17.5%) items as not known/not applicable while rater 2 coded 289 (20.1%) as not known/not applicable.

Additional analyses were conducted to determine how frequently the raters agreed in scoring the items. Table 2 represents cross tabulation of the scores for the two raters.

The results demonstrate that the raters agreed in 90% of the cases whether to score the items as present, absent or not known/not applicable while they disagreed as to the coding category for particular items in only 10% of the cases. In calculating inter-rater reliability, not

Table 2:
Cross tabulation of the Coded Categories (CPAI-Q)

		Rater2			
Coded Categories		“0”	“1”	“8”	Total
Rater 1	“0”	389	40	29	458
	“1”	20	678	32	730
	“8”	13	11	228	252
	Total	422	727	289	1440

know/not applicable category was also taken into account as it was expected that the raters might not agree as to the amount information that was sufficient to score the item as present/absent. The analysis indicated that rater 1 scored 24 items as not know/not applicable while rater two found that there was sufficient information to score these items as present (1) or absent (0). Rater 2 scored 61 items as not know/not applicable while the rater 1 scored the same items as present (1) or absent (0). Subsequent analysis of the items scored as not know/not applicable revealed that information provided were ambiguous, vague, unclear and frequently distantly related to the question. The most frequent problems were vague and ambiguous responses that left an opportunity to infer the presence/absence of the element.

In summary, these results indicate that there is a high level of agreement between raters with 90% of the items being agreed upon by both raters. In order to determine whether such a high level of agreement is due to a chance, statistical indicators of interrater reliability were calculated.

Interrater reliability estimates

The level of the agreement between the two raters was assessed in three different ways: a) Kappa coefficient, b) Pearson coefficient of correlation, and c) t test for differences

between means. Table 3 shows the level of agreement between the two raters across total CPAI scores and individual CPAI section scores expressed by Kappa Coefficients and Pearson's Correlation Coefficients. It also provides values for the measure of internal consistency i.e., Cronbach Alpha Coefficient.

Table 3:
Kappa, Pearson's Correlation and Chronbach's Alpha Coefficients (CPAI-Q)

CPAI sections (N=20)	Items (N)	Kappa	Pearson's r	Alpha
Program Implementation	11	.87**	.95**	.01
Preservice Assessment	15	.84**	.97**	.77
Program Characteristics	24	.86**	.90**	.61
Staff Characteristics	9	.73**	.58**	.04
Program Evaluation	7	.71**	.85**	.55
Other	6	.95**	.97**	.50
Treatment Score	33	.85**	.95**	.83
Total score	72	.84**	.96**	.88

**p<.01

The results indicated that the levels of agreement between the two raters measured by Kappa coefficients were moderate to high and were statistically significant ($p<.01$) for all CPAI sections, treatment and total scores. The highest level of agreement was obtained for the sections Program Implementation (.87) and Other (.95). Sections Staff Characteristic and Evaluation had slightly lower kappa coefficient values (.73 and .71 respectively) in comparison to other sections however they remained statistically significant ($p<.01$).

Additional data on reliability were obtained by calculating Pearson's Correlation Coefficients. The percentages of the items scored as present (out of all scorable items) on each of the CPAI sections within individual programs were calculated for both raters and these scorings were correlated. All correlations were statistically significant with values as high as .97 for the section "Other" and as low as .58 for the section "Staff characteristics". Two sections Staff Characteristics and Evaluation had the lowest values for both Pearson r and Kappa coefficient. The possible explanation for these values is that these sections included a number of open-ended questions that led to scoring differences.

When category 8" was collapsed into category "0" (element not present,) all correlations across different CPAI sections, treatment and total score remained statistically significant and their magnitudes were slightly higher in comparison to values obtained by the analysis that included all three scoring categories. Appendix D contains the interrater reliability indicators expressed as kappa coefficient and Pearson r when only two coding categories were taken into account.

Additional analyses pertaining to the reliability indicators within individual program for the total CPAI score were conducted. The level of agreement between two raters for coding of the individual programs was calculated for the total CPAI score. Appendix E provides the list of individual program sites and level of agreement between two raters expressed as kappa coefficient. All values reached statistical significance indicating that there was a high level of agreement in scoring individual programs. There were noticeable differences among the programs with kappa coefficient ranging from .67 to .93. The variability could be partly explained with different responding styles of the staff completing the CPAI-Q. While some program sites provided concise relevant information, there was a number of sites

that provided vague information that probably led to different scoring. The responses of two programs with the lowest Kappa coefficients (program # 16 and program #18) were analyzed. It was found that staff completing the questionnaires at these program sites provided contradictory information that probably led to different scoring by two raters. The treatment staff from these program sites responded to the questions but have indicated that additional information from different sources (e.g., administrative headquarters) was needed to address some of the program related issues (e.g., type of risk/need assessment conducted, funding/staff related issues). As a result, additional information were obtained from the administrative headquarters. However, the information obtained from these sources was different from the information provided by the treatment staff and two raters had higher disagreement rates while combining these information. The interrater reliability estimates for these programs sites were higher when the researcher had an opportunity to interview the treatment staff and ask detailed questions.

Internal consistency was assessed by Cronbach's Coefficient Alpha and the obtained values ranged from .01 for Program Implementation to .77 for Preservice Assessment. Alpha coefficient for treatment score was .83 and the highest value was obtained for the total score (.88). The inter item correlation ranged from -.01 to .18 with the mean value of .09. The widely accepted cut off in the social science for satisfactory internal consistency is that the alpha coefficient should be .70 or higher, Therefore the CPAI could be considered as an inventory with high internal consistency. It should be indicated that the sample size ($N=20$) was too small for the number of items and it would be desirable to repeat this research with a larger number of program sites. The lowest alpha values were obtained for the Program Implementation and Staff Characteristics sections, likely due to the number of heterogeneous

items included in these sections (e.g., Program Implementation contained items pertaining to the program director, literature review, pilot project, and cost efficiency).

Mean differences

The percentages of the CPAI items that were scored as present for individual CPAI sections were calculated for the individual programs. The mean percentage values for the CPAI sections, treatment and total score were calculated. A t test was used to test the difference between the means. Table 4 summarizes means and standard deviation for CPAI scores and t -test values and their level of significance.

Table 4:
T Test for Difference Between Means for CPAI scores (CPAI-Q)

CPAI section (N=20)	<u>M</u>		<u>SD</u>		<u>t-test</u>	<u>Sig</u>
	Rater 1	Rater2	Rater1	Rater2		
Program Implementation	82.64	81.26	12.37	13.57	.34	.74
Preservice Assessment	57.13	59.01	25.79	27.39	-.22	.99
Program Characteristics	52.75	60.25	16.75	19.37	-1.31	.20
Staff Characteristics	62.40	63.84	13.98	15.31	-.31	.76
Program Evaluation	52.53	45.07	28.28	32.70	.77	.44
Other	62.23	63.90	24.56	26.11	-.21	.84
Treatment score	55.11	59.79	19.08	20.14	-.75	.46
Total score	60.95	62.98	14.87	16.25	-.41	.68

T test analyses did not reveal any significant differences between means on CPAI sections, treatment or total score. These results indicated that raters scored the items within different subcomponent in a similar way that could not be explained by chance alone. It

appears that the CPAI has a detailed manual with clearly defined items that enabled raters to score the items in a consistent manner.

The Program Implementation was the CPAI section with the highest mean values (82.64 Rater 1; 81.26 Rater 2) indicating that the programs included on average 80% of the elements from this section. Mean values for total CPAI score revealed that programs had on average 61% of the elements present. The section with the lowest mean percentage values was Program Evaluation (52.53, Rater 1; 45.07, Rater 2). These results point out that there is a lack of controlled outcome evaluation studies across participating programs.

Standard deviation values indicated that there was a considerable variability within individual programs. While some of the programs were highly structured programs that incorporated the majority of the elements, other programs had a limited number of elements included. For example, the ranges for the sections Other and Evaluation were 0 to 100 indicating that while some of the programs had none of the elements from this section, the other programs had all elements present. The summary of the ranges for different CPAI sections for both raters can be found in Appendix F.

In summary, the variability among programs expressed by both standard deviation and range values was very high, with some programs following the principles of effective correctional programs and other programs having a low number of treatment elements present.

Summary levels for CPAI scores

The CPAI manual outlines that programs with less than 50% of the items present can be coded as being “unsatisfactory”, programs with 50-69% are considered “satisfactory”, and programs with over 70% of the elements present were deemed as “very satisfactory”. These

categories were used, not only for describing individual CPAI sections, but also for the overall assessment score. In order to determine how closely participating programs are meeting known principles of effective correctional treatment, summary scores for the program were calculated for all CPAI sections, treatment and totals scores based on mean scores from both ratings. For example, if one program was scored by Rater 1 as having 45% of the CPAI items present (“unsatisfactory”) and by Rater 2 as having 55% of the items present (“Satisfactory”), the mean score of two ratings was calculated. The calculated mean score (50%) would place the program in the satisfactory category.

Table 5:
Summary Levels for CPAI Sections (CPAI- Q)

CPAI Sections (<u>N</u> =20)	Unsatisfactory	Satisfactory	Very satisfactory
Program implementation	0	3	17
Pre service assessment	6	7	7
Program characteristics	6	9	5
Staff Characteristics	3	10	7
Program Evaluation	10	4	6
Other	3	9	8
Treatment score	6	7	7
Total score	4	8	8

Results indicated that majority of the programs were scored as being in the satisfactory (40%) and very satisfactory (40%) range on the total CPAI score. The analysis of the Preservice and Program Characteristics scores revealed that 30% of the programs were within unsatisfactory range while the rest of programs were in the satisfactory and very satisfactory range. Half of the programs were scored as unsatisfactory for the Evaluation section

indicating that the evaluation of the programs was not a priority to half of the surveyed programs. Program implementation was the subcomponent in which majority (85%) of the programs fall in the very satisfactory range. In addition, only 15% of the programs were scored as unsatisfactory on the Staff and Other sections. We should keep in mind that the scoring was based solely on responses from questionnaires and it is possible that the results reflect the “paper quality” of the program rather than its “real quality” that might be obtained by visiting program sites and interviewing treatment staff. This approach will be examined in the next section.

Institutionally delivered programs were compared to community delivered programs. Table 6 summarizes descriptives for community/institutional delivered programs and t test values. The difference between total CPAI scores of institutional (M=62.45) and community

Table 6:
T test for Difference Between Means (Institutional/Community Programs)

CPAI section (N=20)	<u>M</u>		<u>SD</u>		<u>t-test</u>	<u>Sig</u>
	Inst	Com2	Inst	Com2		
Program Implementation	81.49	82.64	13.35	12.84	-.19	.85
Preservice Assessment	61.52	52.75	28.69	23.31	.73	.48
Program Characteristics	57.52	54.97	19.54	15.32	.31	.76
Staff Characteristics	60.72	66.72	11.31	15.32	-1.01	.33
Program Evaluation	51.12	45.32	30.85	28.59	.42	.67
Other	59.91	67.80	27.43	22.20	-.68	.51
Treatment score	59.81	53.91	22.33	14.54	.66	.52
Total score	62.45	61.26	17.90	11.87	.16	.87

(M=61.26) programs was not statistically significant. Additionally, no statistically significant

difference between means was found on any of the CPAI sections or treatment CPAI score. Analysis of the summary scores revealed that 33% of the institutional programs and 12% of community delivered program were in unsatisfactory range on the total score. Treatment score results indicated that 42% of the institutional and 25% of the community programs were in the unsatisfactory range and did not contained treatment elements most frequently linked with treatment effectiveness.

Summary

All three statistical methods (Kappa Coefficient, Pearson's r , t test) demonstrated high levels of agreement between raters indicating that the CPAI has good interrater reliability. The impressive values of Kappa Coefficients and Pearson Correlation Coefficients show that the CPAI scoring manual has been developed to such extent that the scoring is objective and not influenced by rater related variables. However, the lower but statistically significant values for the Staff Characteristics section suggest that the presence of a number of open ended questions might influence the scoring results. The internal consistency estimates indicated that while the CPAI represents a homogenous scale overall, the individual CPAI section had lower alpha values and they might contain a number of heterogeneous items. Results also indicated there was great variability among the programs in terms of the quality of correctional program offered. However, the majority of the participating programs contained a high number of treatment elements that placed them in "satisfactory" range.

Scoring Audiotaped Interviews

In the second part of the interrater reliability study, data were collected through structured interviews with selected program staff. Interviews were conducted at the same program sites that participated in the CPAI-Q research described previously. While nine

program sites had the program director providing the information to the interviewer, the remaining programs choose to have the complete treatment team present (program director, treatment staff) during the interview. The staff that completed the CPAI-Q in the first part of the research participated in the interview phase of this study. The principal researcher conducted and audiotaped all interviews. Additional data gathering was conducted by reviewing available program materials. All twenty programs sites were visited and their staff interviewed, however two tapes which could not be scored due to poor sound quality and were excluded from the sample.

Interrater reliability was assessed by comparing independent coding of the CPAI by two raters who reviewed the audio-tapes. The first rater was the principal researcher while the second rater was a Ph.D. student who was familiar with the CPAI and treatment effectiveness literature. Coding was based on the information that was gathered from the interview and the supporting documentation (policy and procedure manuals, files) provided to the researcher. As mentioned previously, the printed documents were not available at many program sites and could not be used to score the CPAI based solely on the documentation as originally planned. Information obtained by the CPAI-Q was not taken into account in scoring the interviews. As a result of budgetary constraints only thirteen programs were reviewed. The interrater reliability estimates for the interview part were based on the scoring of these programs.

CPAI treatment elements were scored as present (1) or absent (0) based on the information obtained from the interviews. The score not known/not applicable (8) indicated that information was either not applicable for particular program or absent /not sufficient to score the item. It was the same scoring method used for scoring the CPAI-Q responses in the

first part of the research. Table 7 summarizes frequencies and percentages of scores given by two raters for all items (72) across all programs ($N=13$).

Table 7:
Frequencies and Percentages of Coded Categories (Interview)

Coded categories	Rater 1		Rater 2	
	<u>f</u>	<u>%</u>	<u>f</u>	<u>%</u>
“0”	340	36.3	339	36.2
“1”	449	48.0	449	48.0
“8”	147	15.7	148	15.8
Total	936	100.0	936	100.0

The results demonstrated that the two raters coded the same number of items as present (48.0% Rater 1 vs. 48.0% Rater 2) and absent (36.3% Rater 1 vs. 36.2% Rater 2) and not known/not applicable (15.7% Rater 1 vs. 15.8% Rater 2). It is surprising that the number of items in not known/not applicable category was similar to the numbers that were obtained from the CPAI-Q responses. Even though the interviewer had an opportunity to obtain additional information by asking detailed questions, some of the interviewed treatment staff were not familiar with all aspects of the program and could not provide the answers.

Additional analyses were conducted to determine how frequently the raters agreed in scoring the items. Table 8 represents cross tabulation of the coding categories for two raters.

The results revealed that in 91% of cases the raters agreed whether to code the items as present, absent or not known/not applicable. The discrepancy in coding categories was present in small percentage (9%) of cases. The analysis of not know/not applicable category indicated that Rater 1 scored 15 items as not know/not applicable while Rater 2 found that there is

Table 8:
Cross Tabulation of the Coded Categories (Interviews)

		Rater2			Total
Coded categories		“0”	“1”	“8”	
Rater 1	“0”	307	28	5	340
	“1”	21	417	11	449
	“8”	11	4	132	147
	Total	339	449	148	936

sufficient information to score these items as present (1) or absent (0). Rater 2 scored 16 items as not know/not applicable while the rater 1 scored the same items as present (1) or absent (0).

The high level of agreement in coding categories (91%) indicate that two raters scored the items similarly and additional statistical analyses were conducted to determine whether such high level of agreement is due to chance.

Interrater reliability estimates

Kappa coefficients and Pearson Correlation Coefficients were calculated to determine the level of agreement between the two raters across CPAI sections, treatment and total scores. The interrater reliability indicators were based on the ratings of thirteen programs and were summarized in Table 9

All Kappa coefficients were statistically significant at $p < .01$ level and the values, ranging from .64 to .93 were moderate to high for all CPAI sections, treatment and total score. Staff Characteristics (.64) and Evaluation(.80) sections had lower values in comparison to the rest of CPAI sections, however they were still statistically significant. The high statistically significant values of Kappa Coefficient indicated that the similar coding was not due to

Table 9:
Kappa, Pearson's Correlation, and Cronbach's Alpha Coefficients (Interviews)

CPAI sections (<u>N</u> =13)	Items (<u>n</u>)	Kappa	Pearson's <u>r</u>	Alpha
Program implementation	11	.86**	.92**	.47
Pre service assessment	15	.92**	.97**	.88
Program characteristics	24	.87**	.93**	.65
Staff characteristics	9	.64**	.89**	.54
Program Evaluation	7	.80**	.97**	.79
Other	6	.93**	.96**	.84
Treatment score	33	.89**	.96**	.85
Total score	72	.86**	.98**	.93

** $p < .01$

chance alone but rather that the scoring manual has provided sufficient information to score the items without many discrepancies.

The percentage of the items scored as present on each of the CPAI sections for both raters was obtained and Pearson Correlation Coefficients were calculated to establish the degree of associations. All Pearson r values were statistically significant and all of them were in the high range with the lowest value obtained for the section Staff Characteristics(.89) and the highest for the total CPAI score (.98).

Internal consistency was assessed by Cronbach's Coefficient Alpha and its values ranged from .47 for Program Implementation to .85 for Preservice Assessment. The obtained alpha values for treatment score was .85 and the highest alpha was obtained for the total scale (.93) with mean inter item correlation of .17. Internal consistency estimates were higher than values obtained for the scoring of the CPAI-Q response however they, followed the same

pattern with the sections Program Implementation and Staff Characteristics having the lowest alpha values in comparison to the rest of CPAI sections. Overall, alpha for the total score indicate that the CPAI can be considered a homogenous scale.

Mean differences

Mean ratings for the different subcomponents, treatment and total CPAI scores were obtained and the differences were compared using t test for difference between means. Table 10 summarizes mean ratings, standard deviations and t test values on different CPAI sections for both raters. There were no significant differences between raters on CPAI sections, treatment and total CPAI scores. The sections with the highest mean values for both raters were Other and Program Implementation with 79% (Other) and 72% (Program Implementation) items presents. As noted in the first part of the research, the Program Evaluation section with an average of 41% elements present was the section with the lowest mean value. Programs had on average 56% of items present for the CPAI total score indicating that they were on average in the satisfactory range.

The values of standard deviation and range indicated considerable variability within the sample. The analyses of the total scores across different programs revealed that while some of the programs had as few as 29% of the elements present, others had as many as 77% of the items presents. Some of the programs had a range of CPAI scores from 0% to 100% on individual CPAI sections (e.g., Other). Appendix G summarizes the ranges for different subcomponents for both raters.

Table 10:
T Test for Difference Between Means for CPAI Scores (Interviews)

CPAI sections (N=13)	<u>M</u>		<u>SD</u>		<u>t-test</u>	<u>Sig</u>
	Rater 1	Rater 2	Rater1	Rater2		
Program Implementation	73.07	70.32	17.87	17.16	.40	.70
Preservice Assessment	47.21	46.68	23.11	22.95	.06	.95
Program Characteristics	50.64	54.49	13.66	14.26	-.70	.49
Staff Characteristics	64.74	67.33	21.59	20.25	-.32	.76
Program Evaluation	43.81	38.86	32.47	33.19	.38	.70
Other	79.23	79.48	27.32	27.35	-.02	.98
Treatment score	49.07	51.40	16.25	15.21	-.38	.71
Total score	56.26	56.63	16.66	14.66	-.06	.95

Summary levels for CPAI scores

The summary scores were calculated for all programs on all CPAI sections. Table 11 provides information on summary levels for individual CPAI sections, treatment and total scores.

The analysis of the table indicated that the majority of programs had CPAI scores in the satisfactory and very satisfactory range on the following sections: Preservice Assessment (69%), Program Characteristics (61%), Program Implementation (85%), Staff (69%), and Other (92%). Additionally, 61% of total and 54% of treatment CPAI scores were in satisfactory/very satisfactory range. Program Evaluation was the only CPAI section on which the majority (69%) of the programs scored in unsatisfactory range. These results were similar to the results obtained by scoring the CPAI based on the responses from the CPAI-Q.

Table 11:
Summary Levels for CPAI Sections (Interviews)

CPAI sections (N=13)	Unsatisfactory	Satisfactory	Above satisfactory
Program implementation	2	3	8
Pre service assessment	4	7	2
Program characteristics	5	7	1
Staff Characteristics	4	3	6
Program Evaluation	8	1	4
Other	1	3	9
Treatment score	6	6	1
Total score	5	5	3

Summary

Interrater indicators calculated for interview based CPAI scoring indicated a high level of agreement between the two raters. The Kappa and Pearson r were moderate to high and statistically significant. There were no statistically significant differences between raters on mean CPAI scores across CPAI sections, total and treatment score. In respect to the summary CPAI scores, the analysis revealed that the majority of the programs were at the satisfactory and very satisfactory level for all individual CPAI section (with exception of Program Evaluation) and for total and treatment CPAI score. Internal consistency estimates suggest that the CPAI is a homogenous scale.

Comparison of Two CPAI Scoring Methods

In order to determine whether the different CPAI scoring methods yielded different results, CPAI scores obtained by coding interviews with treatment staff for thirteen programs were compared to scoring of corresponding programs based on their responses from the

CPAI-Q. It was hypothesized that two methods will yield different results, as the questionnaires might indicate that treatment elements were “likely” present while the interview will enable the researcher to ask in detail the questions and to determine whether the treatment elements were “actually” present. While responses to the questionnaire could indicate the “paper quality” of the program, it is expected that the interviews will reveal the extent of program adherence to the principles of effective programming.

Notwithstanding high interrater reliability between the two raters indicated by statistically significant and moderate to high Kappa and Pearson coefficients, the analysis of the differences between different scoring methods was based on the average ratings. The average scores obtained from the thirteen programs that participated in the interview part of the research were compared to the average scores obtained for the same programs based on the responses on the CPAI-Q (Table 12).

The analysis of the results revealed higher CPAI scores for the scoring based on the responses from the CPAI-Q in comparison to the scoring that was based on the information obtained from the interviews conducted with the treatment staff. These differences were found for treatment and total CPAI scores and all CPAI sections with exception of Staff Characteristics and Other which had slightly higher values when scoring was based on the interview based information. However, when the differences between means were tested using t tests, the only difference that had reached statistical significance was for the section Program Implementation ($p < .05$). Higher mean values on this section for the scoring based on the CPAI-Q responses were likely a result of several questions that contained yes or no answers. When the same information were explored during the interview part, detailed questions about these treatment elements gave an opportunity for detailed exploration of the

Table 12:
Comparison of Two CPAI Scoring Methods

CPAI sections (N=13)	CPAI Scoring Method				t-test	sig	r
	Interview		CPAI-Q				
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>			
Program Implementation	71.74	17.21	85.05	9.95	2.14	.02	.34
Preservice Assessment	46.94	22.88	65.04	24.90	1.93	.07	.75**
Program Characteristics	52.56	13.93	61.55	16.21	1.48	.15	.43
Staff Characteristics	66.03	20.25	63.71	13.89	-.37	.72	.39
Program Evaluation	41.33	31.71	57.26	27.99	1.30	.21	.79**
Other	79.35	27.08	62.16	26.15	-1.65	.11	-.08
Treatment score	50.23	15.75	63.03	16.58	2.00	.06	.77**
Total score	56.44	15.57	66.46	12.86	1.79	.09	.80**

**p<.01

presence of the elements. There is a possibility that responses on the questionnaire might have a tendency to portray programs in a positive light while the interviewer has an opportunity to ask additional questions to determine whether the elements were actually present.

Additionally, the information obtained from the CPAI-Q were frequently limited, vague and unclear and the interviews enabled the clarification of these inconsistencies.

The correlations between scores obtained when scoring the CPAI-Q and the interview expressed as Pearson Correlation Coefficient were found to be significant for Preservice Assessment (.75), Program Evaluation (.79), Treatment (.77) and Total (.80) CPAI score. Correlations for the rest of the CPAI sections were in low to moderate range and did not reach statistical significance.

Despite higher scores on the survey version in comparison to the interview version, the absence of a statistically significant difference was an indicator that both the interview and the questionnaire yielded similar results and that the CPAI could be scored reliably by either method. Our second hypothesis was not confirmed as there was no significant difference between different scoring methods. The fact that the same staff completed the CPAI-Q and participated in the interview might be one of the reasons for these results. Additionally, the time between the completion of the questionnaire and the interview might have been too short and influenced the final results.

Qualitative Analysis of the CPAI Items

In the final part of the interrater reliability study individual CPAI sections were reviewed and the scoring problems identified by the raters were discussed. The purpose of this analysis was to suggest improvements in the scoring manual and/or CPAI items based on difficulties encountered while scoring the CPAI based on the CPAI-Q responses and audiotaped interviews.

Program implementation

The most problematic part in scoring this CPAI section was the lack of clearly defined program director/managerial position for provincially and federally delivered programs. The importance of the involvement of the program director in program design, implementation, and supervision was reflected by several CPAI items pertaining to the role of the program director. However, the position of program director was often an administrative role (e.g., Deputy Superintendent of Correctional Programs in the provincial system and Assistant Warden of Correctional Programs in the federal system). The clinical supervision was done by the Chief Psychologist/Chief Social Worker in the provincial system and Regional Trainer in

the federal system. The fragmented and limited role of the program director and different levels of supervision, revealed an emerging bureaucratic structure in which the information was not distributed to all levels. The treatment staff reported unfamiliarity with elements related to program implementation. For example, the treatment staff were not able to answer questions pertaining to literature used in the design of the program or whether pilot research was conducted. The program designer was identified as the person that would be able to provide these answers.

The participants indicated that the questions pertaining to cost effectiveness were difficult to answer adequately as clear criteria were not set to indicate exactly what a cost effective program means. They suggested that a short questionnaire pertaining to the cost effectiveness should be formulated, piloted and implemented in order to explore staff perception of the program's cost effectiveness.

Preservice assessment

The assessment pertaining to risk, need, and responsivity factors were not usually conducted at the program sites. In the federal system all initial assessments were concentrated in the Millhaven intake unit. Based on this assessment a correctional plan was formulated and correctional treatment programs recommended. The risk assessment within the provincial correctional system was usually conducted in detention centers and based on these assessments offenders were placed in correctional or treatment centres where individual program plans were developed. The risk assessment was not repeated at the program sites unless the assessment was more than 6 months old in which case an updated assessment was done. As a result of this assessment process, treatment staff were frequently unaware of the

specifics of the assessment process and the questions were referred to the classification/intake staff.

Despite recent research findings which suggest that motivation should be addressed in treatment design and delivery, the assessment of motivation factors was underrepresented. Federal programs, however, recognized the importance of including motivation assessment for treatment effectiveness and incorporating motivation interviewing in the program design. Despite the lack of the adequate tools for the assessment of motivation, considerable research activity in recent years is encouraging.

Program characteristics

The questions pertaining to punishment evoked strong feedback from the subjects. The participating treatment staff expressed a belief that these items do not reflect the current knowledge about corrections and that deterrence is “an obsolete concept”.

Participants overwhelmingly stated that although the matching staff /client, client/ program, and staff/program were desirable characteristics of a correctional program, financial restraints and low staffing levels would prevent its implementation. Only one of the participating programs was scored as having these treatment elements present. Furthermore, despite acknowledgement that the completion criteria is an important treatment element, only a few programs had completion criteria defined as “achieving target behavior” as stated by the CPAI item. For the majority of the programs, the duration of the program was used as the completion criteria

Staff characteristics

Recent research had indicated that staff related characteristics could influence treatment outcome. The selection of program staff has become an increasingly important

element in correctional programming. The federal system has formulated criteria for an effective program deliverer which was used for hiring program delivery officers. However, the selection process was centralized in the regional headquarters and the treatment staff frequently did not have information about the hiring /selection procedures. They indicated that Human Resources and Regional Trainers would be appropriate contacts for adequate information. The provincial correctional system did not have a treatment designated position, as the treatment delivery was identified as a one aspect of a psychologist/social worker job description. There was no hiring criteria specific for treatment delivery, although the selection criteria for psychologist/social worker staff took into account some of the characteristics important for an effective treatment deliverer. Given that the hiring/selection of the staff was considered confidential, many of the staff related issues were not available to the principal researcher.

Evaluation

Evaluation of treatment outcome has been underrepresented among the majority of the participating programs. Only one of the surveyed provincial programs had a controlled outcome study conducted. Although a number of federally delivered programs had controlled outcome study completed, the evaluation methods and results were not known to the treatment staff indicating lack of communication between researchers and treatment staff. The questions pertaining to the program evaluation were referred to the researchers as the most reliable source of information for this topic.

Other

Treatment staff did not have difficulties answering questions pertaining to the items from the section Other. The only item that was identified as vague and not clear was "Ethics

of the intervention”. Participants indicated that even though they provided answers, their understanding of the question was limited at best.

Discussion

High values of Kappa Coefficient and Pearson Correlation Coefficient obtained for both scoring methods indicate a high level of agreement between raters when the CPAI is scored based on information from a) interview with treatment staff and b) responses from a structured CPAI based questionnaire. The fact that high interrater reliability was obtained by the two sets of raters indicated that regardless of the level of experience in scoring the CPAI, the training of the raters and use of the CPAI scoring manual is sufficient to lead to reliable results.

Furthermore, the high level of interrater agreement was obtained not only for the overall total CPAI score, but also for individual CPAI section scores. The highest level of interrater agreement expressed by Pearson Correlation Coefficient was obtained for the section Pre Service Assessment and these results were found for both scoring methods. The high interrater reliability for this CPAI section was likely a result of clear and concise scoring criteria for risk, need and responsivity assessment that were included in this section. None of the raters had problems scoring these items.

Both scoring methods indicated that the section Staff Characteristics had the lowest interrater reliability estimates. This CPAI section contained a number of items that were scorable only if certain strict criteria were met. For example, in order to score the item pertaining to staff hiring, 5 out of 8 listed characteristics should be specified. However, the open ended format of the questions left a lot of space for vague formulation that could be coded differently by two raters. In order to eliminate these difficulties, both CPAI-Q and

interview methods should ask detailed question about all of 8 listed hiring criteria, instead of leaving it to the respondents to provide their own criteria. This will likely lead to higher interrater reliability for this CPAI section. Notwithstanding the difficulties in scoring these items, interrater estimates for this section were in moderate range and statistically significant.

Internal consistency estimates indicated that the CPAI is a homogenous scale and the high values of alpha coefficients for the CPAI total and treatment scores were obtained for both interview and questionnaire based scoring of the CPAI. The lowest values of alpha coefficients were obtained for Program Implementation and Staff Characteristics sections for both methods. The analysis of these sections indicated that they consisted of a number of heterogeneous items. For example, the Program Implementation section consisted of items related to the qualification of program director but also to the literature review, pilot project, cost efficiency. Overall, the first hypothesis was confirmed as the CPAI has high interrater and internal consistency reliability and the scoring of the individual CPAI items is likely to be independent from rater related variables

There were no statistically significant differences between scoring methods indicating that regardless of the scoring method, the CPAI appears to be a reliable instrument. Still, CPAI- Q scoring method yielded higher mean scores than the interview method on the majority of the CPAI sections, however none of these differences reached statistical significance. The lack of confirmation of our second hypothesis, as noted previously is likely a consequence of the same staff completing the CPAI-Q and participating in the interview within a short period of time.

Given the high reliability for both methods, it should be emphasized that these result do not suggest that the questionnaire method is equally useful as the interview method. The

amount information that was obtained in the interview part was much greater and this information would be valuable in the evaluation of individual programs. The goal of this study was not to evaluate individual programs but rather to provide a data of how confident we should be the CPAI scoring is reliable and independent from rater related characteristics.

The number of the items that was scored as missing and not applicable by both methods indicates that some improvements in the scorings approach is needed. Contradictory information provided by the treatment team and administration at the some program sites and their impact of the scoring of the CPAI discussed in the previous sections, indicate that there is a need to have staff knowledgeable about different aspects of program implementation, design, delivery and evaluation. It was suggested that in addition to treatment staff and the program director, a number of other staff should be interviewed in order to obtain pertinent information. Most specifically, program designers should be consulted about program implementation, human resources staff should be asked about hiring/selection of the staff, researchers should provide information about any evaluation studies conducted. Undoubtedly, it will require more resources, however, it will provide more information about the program itself. Finally, it was evident that there was considerable variability with the programs and even though it was not initially planned, some additional analyses were conducted to determine how closely programs have met principles of effective programming. The results demonstrated that the majority of the programs scored in the satisfactory and very satisfactory range on the overall total score and on all CPAI sections with exception of Program Evaluation. This was the section that had the highest number of programs scored in unsatisfactory range indicating a need to incorporate evaluation related elements in

correctional programming. There were no differences in the quality of delivered programs assessed by the CPAI between institutional and community delivered programs.

In summary, the CPAI represents a homogenous scale that could be scored reliably by different raters regardless of whether the CPAI is scored based on the questionnaire or the interview. It seems that the interview with a number of treatment staff that are knowledgeable about different aspects of the program, accompanied with a review of the program documentation will be the best method of data gathering. The main limitation of this study was the unavailability of the program documentation that could be used for scoring the program. Additionally, it would be beneficial to have the evaluator observe program delivery and score some of the staff related characteristics: e.g., skills, adherence to program intervention, productivity of the sessions. The combination of different data gathering methods will likely provide invaluable information as to where the program fits in terms of evidence based correctional practice. Future research will hopefully provide additional information about the contribution of different administration/scoring methods.

Study 2: Face/Content Validity Study

The increasing use of the CPAI in process evaluations of correctional programs (Andrews, 1995a) requires the assessment of its validity. This dissertation addressed face, content and predictive validity in two separate studies.

The main goal of Study Two was to establish the face and content validity of the CPAI. A systematic analysis of the instrument's content and face validity is necessary to establish whether the CPAI could be used for planning, implementing, modifying, and evaluating treatment programs. Without satisfactory validity data, the extent to which the instrument covers treatment elements empirically linked to treatment effectiveness and its potential usefulness in the evaluation of effectiveness of correctional programs will remain unclear.

Face Validity Study – Goals

Face validity, although not validity in the technical sense, is a desirable feature of a test. It is important to determine face validity characteristics of the CPAI because it pertains to whether the test “looks valid” to the administrative personnel who decide on its use and to the general public. In a provocative article Nevo (1985) called attention to the lack of available research on face validity, despite its probable influence on the attitudes of the general public and nonprofessionals toward a test. In his opinion it is an important feature of a test that has implications for its practical use. It is suggested that even though a test appears to be pertinent and related to the area that it is suppose to assess, the test should not only be valid but should also appear valid in order to enhance its use. Nevo (1985) suggested a quantitative assessment of the face validity by having administrative personnel rate the suitability of a test for its intended use. A direct measure of face validity can also be obtained by asking the general

public to rate the validity of the test as it appears to them and to evaluate individual items. The face validity can be considered a coherent construct only if a reasonable level of agreement exists among raters (Nevo, 1985).

The main goal of this study was to obtain quantitative indicators of face validity of the CPAI by surveying a sample of subjects about whether they consider the CPAI to be an instrument that could assess the potential of correctional treatment programs to reduce recidivism rates of offenders. In the preliminary phase of this research, administrative personnel working for the CSC and the MCS were identified as an adequate sample for the face validity study. It was expected that they will not be familiar with the treatment effectiveness literature and therefore could give an opinion as to whether the CPAI appears to be a valid instrument. After both organizations indicated that it would be very difficult to include administrative personnel in the research due to organizational issues, a new sample consisting of first year psychology students from Carleton University was included. This choice was based on an assumption that the students could represent the general population due to their unfamiliarity with the literature on treatment effectiveness.

Content Validity Study-Goals

Content validity is another important consideration when evaluating the value of the CPAI items in correctional programming. Content validity is built into the test through the choice of appropriate items that provide adequate representation of the conceptual domain the test is designed to cover (Kaplan & Saccuzzo, 1993). The most frequently used method to ensure content validity is a review of the literature in the area of interest in order to determine whether the test covers the most important elements linked to the conceptual domain. The review of the literature conducted in the previous section indicated that the CPAI covered the

majority of the elements found to be empirically linked to treatment effectiveness.

Furthermore, a few additional treatment elements potentially associated with treatment outcome were also identified. Although the literature review summarized in the previous sections indicated that the CPAI has satisfactory content validity, a quantitative assessment of content validity of the complete instrument was also conducted.

The quantitative assessment of the content validity of the CPAI was conducted by surveying psychologists and treatment staff working within the CSC and the MCS as to their opinion of whether the CPAI is covering treatment elements that are relevant for treatment effectiveness. An additional group of psychologists and criminologists who worked in academic/research settings and taught/researched correctional issues was also surveyed. The assumption underlying the sample selection was that these three samples were “experts” with extensive knowledge about program delivery and treatment effectiveness literature. They should be able to provide information about whether this instrument adequately covers treatment elements that are linked to effective programs (i.e., programs that reduce reoffending rates).

Method

A questionnaire that addressed the elements of effective correctional programs identified by the CPAI was constructed. The Effective Treatment Elements Questionnaire (ETEQ) included questions about treatment elements covered by the individual CPAI items. Previous analysis has identified a number of potentially important treatment elements that were not included in the CPAI (Gendreau & Andrews, 1996) and these elements were included in the ETEQ (“motivation for change”, “small treatment groups”, “productive therapy sessions”, “favorable conditions”, “belief in effectiveness of correctional programs”,

“adherence to treatment intervention”). Additionally, treatment elements pertaining to the Core Correctional Practices (CCP, Andrews & Carvell, 1998) described in the introductory part were also included in this study (i.e., “human service”, “personal characteristics of staff”, “believe in their skills to run the programs effectively”, effective reinforcement “, “effective modeling”, “effective disapproval”, “problem solving techniques”, “authority”, “structured learning procedure”).

Participants were asked to state whether they consider these treatment elements to be important for program effectiveness (i.e., reducing offender’s recidivism) on a 4-point scale (ranging from 1 – an element is not important for program effectiveness to 4 element is very important for program effectiveness). At the end of the survey participants were asked to describe the treatment elements they consider to be important for treatment effectiveness that were omitted in the questionnaire.

The same survey method was used for both studies. While face validity indicators were established by surveying first year psychology students, the content validity data were obtained on the sample of treatment staff and psychologists working within the federal or provincial correctional systems and professionals from academic/research settings.

Prior to the beginning of the research, a pilot study was conducted on a sample of 15 randomly selected subjects. Psychologists ($n=5$), treatment staff ($n=5$) and first year psychology students ($n=5$) reviewed and completed the questionnaire. They were asked to comment on a) clarity of instructions b) problems in answering the questions and c) to note any potential problems in responding. Revisions of the original questionnaire were made based on their feedback and the final version consisted of 84 items divided into 6 parts corresponding to the CPAI sections. Appendix H contains the complete version of the ETEQ.

Sample

The final sample consisted of 50 first year psychology students, 50 psychologists and 34 treatment staff. This selection was based on the assumption that while the students did not have prior knowledge of the CPAI and treatment effectiveness literature, psychologists and treatment staff were familiar with the literature. There was no specific inclusion criteria for the student group. A sign up sheet was posted on the experimental board at Carleton University where students attending Psychology 100 had an opportunity to sign up for research participation. Students received two experimental credits for the completion of the questionnaire.

The Ministry of Correctional Services and the Correctional Service of Canada provided a list of psychologists and treatment staff working in the institution/community. The inclusion criteria for both samples were: 1) minimum one year of experience in program delivery and 2) working in institutional or community programs at federal or provincial level. While program delivery within the federal system was conducted mainly by program delivery officers, the provincial system had a model in which program delivery was conducted by other clinical staff (e.g., social workers). As a result, the treatment staff group consisted of social workers and program delivery officers. Fifty-three psychologists were contacted and agreed to participate and all but three sent back completed questionnaires (94% response rate). In the sample of treatment staff, forty five treatment staff agreed to participate, and 76% of them completed and returned the questionnaires. In summary, fifty psychologists and thirty four treatment staff were included in the final sample.

An additional sample of 25 professionals working in academic/research setting were contacted from the list of psychologists registered with the Canadian Psychological

Association (Criminal Justice Section). Only three subjects returned the questionnaire and therefore this sample was not included in the analysis. Such a low response rate among the professionals that conduct research and contribute greatly to the improvement of the knowledge in the correctional field is difficult to explain.

Hypotheses

Hypothesis 1: It is expected that the CPAI will have good face validity. In other words, it is hypothesized that students will demonstrate high levels of agreement on the importance of the treatment elements included in the CPAI for the reduction of recidivism.

Hypothesis 2: It is expected that both psychologist and treatment staff working in correctional field and professionals from academic/research settings will demonstrate high levels of agreement on the importance of treatment elements included in the CPAI for the reduction of recidivism, thus providing an indication of good content validity.

Results

Sample/Demographics

The student sample was predominantly female (66%) with the mean age of 21.70 years ($SD=5.58$). Within the sample of psychologist more than half were female (52%), with the mean age of 42.06 (SD 11.27) and working in corrections an average 11.8 years (SD 6.97). Treatment staff consisted of 76% females, the mean age for the sample was 36.03 ($SD=7.03$) and the sample worked in a correctional setting an average of 8.73 years ($SD=5.14$). As expected by the sample selection, all psychologists had graduate degrees, while 94% of the treatment staff had undergraduate degrees.

Analyses of the responses pertaining to the belief in the effectiveness of correctional treatment programs (e.g., Question: Do you believe that correctional programs can decrease

the criminal behavior of its participants) revealed that 90% of the total sample believed that correctional treatment programs have an impact on reoffending, while only 3% of the sample believed that correctional programs have no impact on recidivism rates. The remainder (7%) reported that they were not sure about the impact of treatment programs on recidivism. There were no significant differences among the three groups in their answers (Chi square =7.61, n.s.).

Thirty five percent of all subjects reported that they were either “familiar” or “very familiar” with the literature on treatment effectiveness while the rest reportedly was “not familiar” (38%) or “somewhat familiar” (25%) with the literature. A Chi square comparison among the three groups revealed a significant difference (Chi square =100.52; $p < .01$). The number of subjects who were “not familiar” with the treatment effectiveness literature was significantly higher in the student group ($n=44$; 88%) than in the psychologist ($n=5$; 10%) or treatment staff ($n=2$; 6%) groups. The number of psychologists who were “very familiar” and “familiar” with treatment effectiveness was significantly higher in comparison to the treatment staff group.

The “Psychology of Criminal Conduct” (Andrews & Bonta, 1994, 1998) was “not known” to 38% and “somewhat known” to 27% of the total sample, while the rest (35%) reported being either “familiar” or “very familiar” with the “Psychology of Criminal Conduct”. There were significant differences among groups (Chi Square =91.87, $p < .01$) with the number of students not familiar with the Psychology of Criminal Conduct significantly higher ($n=42$; 84%) in comparison to the psychologists ($n=0$; 0%) and treatment staff ($n=8$; 24%). Additionally, the number of psychologists familiar with the Psychology of Criminal Conduct was significantly higher in comparison to the treatment staff.

Only fourteen percent of the total sample indicated that they were “very familiar” or “familiar” with the CPAI. The majority (56%) indicated that they were “not familiar” or “somewhat familiar” (28%) with the CPAI. When results were compared across different groups statistically significant differences were found (Chi square =42.70, $p<.01$). A significantly higher number of students ($n=44$; 88%) were “not familiar” with the CPAI in comparison to treatment staff ($n=13$; 38%) and psychologists ($n=17$; 34%). Again, the group of psychologists formed a distinct category with a significantly higher number that were “very familiar” with the CPAI in comparison to the treatment group.

In summary, a statistical analysis of the results indicated that there was a significant difference among the groups; more specifically the number of students unfamiliar with the literature on treatment effectiveness, the Psychology of Criminal Conduct, and the CPAI items was higher in comparison to the group of psychologists and treatment staff. These results were expected as the choice of the student sample for the face validity study was based on the assumption that they will not be knowledgeable in treatment effectiveness/CPAI field. These results were also an indirect confirmation of the adequate selection of the student sample.

Significant differences were also found for the sample of psychologists. The number of psychologists that were familiar with literature on treatment effectiveness, psychology of criminal conduct and the CPAI was significantly higher in comparison to other groups. These results were expected as the selection of psychologists for the content validity study was based on an assumption that they were “experts” in the area.

Intraclass Correlation

Indicators of face and content validity were calculated using intraclass correlation. Intraclass correlation (ICC, Shrout & Fleiss, 1979) is defined as a statistical measure that assesses the level of agreement between raters. As noted in the previous sections, the goal of this study was to determine whether there is a high level of agreement about the importance of the CPAI items for the treatment effectiveness among students (Face validity study) and psychologist and treatment staff (Content validity study). ICC is a measure that will provide data about the interrater agreement and is a method of choice when there is a number of raters. As the main interest was the level of agreement among the raters, a two way random effect model was used. Combining multiple ratings generally produces more reliable measurement, and the ICC score was calculated for averaged ratings. A high intraclass correlation coefficient is usually associated with small within subject variance and higher agreement among raters (Bartko, 1976). Table 13 summarizes the intraclass correlations within the total sample and within the three subsamples.

Table 13:
Intraclass Correlation: Two Way Random Effect Model (Absolute Agreement)

Groups	N	ICC	95% C.I.	
			Lower	Upper
Students	50	.89**	.85	.92
Treatment Staff	34	.89**	.84	.93
Psychologists	50	.92**	.88	.94
Total sample	134	.96**	.95	.98

**p<.01

The results indicated that the intraclass correlations were high and statistically significant ($p < .01$) within all groups of raters (i.e., students, treatment staff and psychologists) and also within the total sample. The high, significant value of the intraclass correlation indicates that the raters have demonstrated a high level of agreement in rating how important the CPAI items are for program effectiveness. These are quite robust results as high level of agreement was obtained for both professionals and the students. Additionally, analyses were conducted to determine the level of agreement across different CPAI sections and these values were high and statistically significant across the different sections (e.g., Program Implementation .95; Preservice .91; Program Characteristics .98 and Staff .93).

The interpretation of the magnitude of the ICC is not clearly defined as there is no widely accepted standard indicating what represents adequate agreement. Some authors suggest using the following criteria in the interpretation: ICC less than .20 indicates slight agreement; .21 to .40 fair agreement; .41 to .60 moderate agreement; .61 to .80 substantial agreement and .80 or more almost perfect agreement. As all of obtained ICC in this study were above .80 it could be concluded that the level of agreement among subjects was high.

Both of the general hypotheses have been confirmed. A high level of agreement was demonstrated within the group of students that represented the general population. This indicated that the CPAI has satisfactory face validity and was perceived by lay people as containing items relevant for treatment effectiveness. Additionally, high levels of agreement obtained within groups of psychologists and treatment staff suggested satisfactory content validity of the CPAI. Even though intraclass correlation provided the information about the level of agreement among raters it did not indicate the degree to which the raters found the

items to be important for treatment effectiveness. Therefore, additional analyses were conducted to establish whether the items were seen as important for treatment effectiveness.

Mean Rating Values

As an indication of how different the ratings were among the three groups, the mean ratings across the different CPAI sections were compared (see Table 14).

Table 14:
Mean Ratings for CPAI Sections

CPAI sections	Students (<u>n</u> =50)		Treatment Staff (<u>n</u> =34)		Psychologist (<u>n</u> =50)	
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>
Program Implementation	3.16	.39	3.19	.50	3.14	.43
Preservice Assessment	3.31	.43	3.35	.65	3.40	.43
Program Characteristics	3.16	.37	3.01	.58	3.06	.40
Staff Characteristics	3.39	.35	3.34	.51	3.43	.37
Program Evaluation	3.32	.42	3.20	.85	3.32	.40
Other	2.82	.54	2.52	.69	2.42	.60
Treatment Score	3.24	.37	3.18	.56	3.23	.37
Total Score	3.19	.31	3.09	.53	3.13	.34

The statistical analysis of the results comparing mean values on different sections across different groups using the one way ANOVA revealed that there is no statistically significant difference among the three groups on treatment, total CPAI scores and on any of the CPAI sections with exception of the section Other. Statistically significant results for this section, $F(2, 133)=5.75, p<.01$, were found and Scheffe comparison between groups found that the mean values for this section were significantly lower for the treatment staff and

psychologists in comparison to student's rating of this section. Nevertheless, the mean values for this subcomponent (program delivery $\underline{M}$ =2.42; psychologists $\underline{M}$ =2.45, students $\underline{M}$ = 2.82) were still high indicating that its items were also considered to be important. Additionally, test comparisons indicated that mean values for section Other were significantly lower in comparison to other CPAI sections within all three samples. These results were not surprising as this section contained broader treatment issues linked to the community: for example, support for the program, funding. Even though these elements could influence program effectiveness, their impact on treatment outcome is likely to be less important in comparison to some of program related characteristics (e.g., targets, treatment intervention).

The analysis of the results pertaining to the individual treatment items added to the ETEQ revealed that the mean values ranged from 3.12 for “motivation for treatment” to 3.54 for “productive therapy sessions”. Descriptive statistics for the mean rating of individual items were presented in Appendix I. The individual items corresponding to the CCP had even higher mean values that ranged from 3.47 for “structured learning” to 3.69 “staff committed to human service”. Overall, the high mean values within the total sample for the group of 6 additional items was 3.37 and the group of the CCP based items (3.59) which indicates that these elements are considered as important for treatment effectiveness by all three groups. The results suggest that there is a belief among psychologists, treatment staff and students that these elements are important for the treatment effectiveness and that they deserve future exploration especially in the predictive validity study. As indicated previously the majority of these elements were included in new and revised version (CPAI-2000).

In summary, the analysis of mean values for CPAI sections revealed that all raters considered them as important for treatment effectiveness. Section Other was scored slightly lower than the rest of CPAI sections, however it is still considered as “somewhat important”.

Discussion

The validity of the CPAI was assessed by surveying students (face validity study), psychologists and treatment staff (content validity) about whether they consider that the CPAI is covering treatment elements relevant to treatment effectiveness (i.e., reducing recidivism). There was no statistical difference in the mean ratings of the three samples and the level of concordance was high within the samples and across all samples. These results indicate that the face and content validity assessed by the level of agreement within groups of students, psychologists and treatment staff, expressed by an ICC was high. The general hypotheses were confirmed.

More specifically, the students demonstrated high levels of agreement in rating the importance of the CPAI items for treatment effectiveness, providing data for the adequate face validity of this inventory. These results indicated that the test appears to be valid and covers elements that are likely linked to reduction in recidivism. It should be noted that over 80% of the student sample reported to be unfamiliar with the literature on treatment effectiveness, the CPAI or Psychology of Criminal Conduct (Andrews & Bonta, 1998). These results confirmed that the sample selection was adequate for assessing the face validity of the CPAI as the students were unfamiliar with the concepts that could have potentially influence their responses.

Second, psychologists and treatment staff working within correctional settings were selected as it was expected that they will be familiar with the treatment effectiveness

literature. Their knowledge of treatment effectiveness literature and experience delivering the correctional programs should make them “experts” in the field, therefore, they could be considered as an adequate sample to be surveyed for the content validity study. The results confirmed that these two groups were indeed familiar with treatment effectiveness literature. Both groups demonstrated high levels of agreement on the importance of treatment elements included in the CPAI for the reduction of recidivism and the analyses of mean scores have indicated that they consider all of the CPAI sections to be important for treatment effectiveness. Initially, it was planned that an academic/research sample, based on their theoretical/research knowledge, will provide information as to where the CPAI fits in term of evidence based practice and potentially identify additional elements linked to treatment outcome. This sample had such a low response rate that it was excluded from the analysis.

When the mean rating values for the CPAI sections were analyzed the only significant difference was found for the section Other. Psychologists rated this section as lower in importance in comparison to the student sample, however the items within this section were still rated as important. It is impressive that the mean ratings for different CPAI sections were so high across three samples. Elements pertaining to CCP and group of six elements were perceived by all three groups as important for the treatment effectiveness and they should be explored further.

At the end of the questionnaire, the participants were asked to note any additional potentially important elements that might have been omitted in the ETEQ. The analyses of their responses revealed that majority of the sample (94%) believed that the ETEQ has included all relevant treatment elements and there is no need for inclusion of any additional elements. A few suggestions made by the student sample included: a) offender participation

should be solely on voluntary basis as any type of coercion will not work on offenders, b) offenders should demonstrate some improvement during the program in order to remain in the program c) extensive follow up in the community and adequate, timely discharge planning, and d) participation of offenders in program design and delivery.

Treatment staff indicated that all relevant elements were included and they suggested that a) program deliverers should not be scrutinized so much, b) the skill level, experience and suitability should be more important than university degree, c) good rapport with the clients should be established, d) the group format should be accompanied by individual sessions and e) correctional officers should be involved in treatment delivery.

Finally, the sample of psychologists did not provide additional input as they believed that the majority of the relevant treatment elements were included. The only suggestion that this sample provided was that the treatment staff should be knowledgeable about the theory underlying particular treatment intervention.

The analyses of the responses of the final open ended questions confirmed that in general participants believed that the ETEQ (and consequently the CPAI) represent an adequate representation of treatment effectiveness literature. In summary, qualitative analysis of the literature on treatment effectiveness conducted in the introduction and quantitative analysis conducted in this study, suggest that both the face and content validity of this instrument seem to be satisfactory.

Study 3: Predictive Validity Study

Goals

Recent treatment effectiveness literature (Andrews et al, 1990; Andrews & Dowden, 1999; Andrews, Dowden, & Gendreau, 1999; Cleland, 1997; Dowden 1998) has identified a number of treatment elements empirically linked to the reduction in recidivism. The CPAI (Gendreau & Andrews, 1996), an instrument designed to assess the extent to which correctional programs meet principles of effective correctional interventions, summarized many of these treatment elements. Although the link between the CPAI items and treatment effectiveness (i.e., reduced recidivism) makes logical sense, there are no empirical indicators as to the strength of the relationship between the CPAI scores and treatment outcome.

A strong and significant correlation between the CPAI scores and treatment outcome would suggest that, in order to be more effective, policy makers and program designers should include treatment elements identified by the CPAI in the development and design phase of a program. Unfortunately, data on the predictive validity of the CPAI are limited and the absence of firm predictive validity indicators questions its usefulness in program evaluation. Recently conducted research by Latessa et al. (1999) obtained preliminary predictive validity data, however the sample was very small. Their results suggested that the higher scores on the CPAI were associated with more effective programs (i.e., programs that reduce reoffending rates). The present study will extend the research on the predictive validity of the instrument.

Past research on predictive validity suggests two methods that could be used to obtain the data on predictive validity of the CPAI. First, the CPAI could be used to score correctional programs whose participants will be followed up in a longitudinal controlled outcome study of recidivism. An individual researcher working from an academic setting would have limited

access to the information needed to conduct such a large scale study. In addition, the length of time required to complete such a study also represents a serious limitation of this research design. Second, the data on predictive validity could be obtained by using the CPAI to evaluate ongoing correctional treatment programs that have outcome data available. However, the number of program locations that have conducted controlled outcome studies with documented reoffending rates of its participants is limited. These limitations have led to the development of an alternative research design, using meta-analytical research for establishing predictive validity characteristics of the CPAI.

Method

Contribution of the Present Study

Results of meta-analytical studies conducted within the last decade (e.g., Andrews et al. 1990, Andrews & Dowden, 1999; Andrews, Dowden & Gendreau, 1999; Dowden, 1998, Lipsey 1992) offered strong empirical evidence about the effectiveness of correctional treatment programs. These meta-analytic studies expressed the strength of the treatment effectiveness by the effect size. The effect size is a numerical expression representing the difference in reoffending rates between the treatment and control group. Therefore, effect size represents a measure of treatment outcome defined as recidivism reduction.

The on-site evaluation of the programs included in previous meta-analyses, using the CPAI, cannot be conducted as the programs were delivered long ago and some of them no longer exist. However, the studies included in the previous meta-analyses contained descriptions, albeit limited, of program interventions for both treatment and control groups that could be used for scoring the CPAI items. Each of the treatment and control conditions were scored using the CPAI. The obtained difference in the CPAI scores between treatment

and control groups were linked to corresponding effect sizes, thereby providing some rudimentary data on the predictive validity of this inventory.

The present study extended previous meta analyses (e.g., Andrews et al. 1990, Dowden 1998) in several ways. First, it provided estimates of the predictive validity of the CPAI. Second, by using the CPAI to score the control groups it revealed the extent to which control groups incorporated effective treatment elements. Third, the CPAI was scored in an innovative way based on the information from the articles.

New CPAI Coding Procedures

The original CPAI coding manual was designed in such a way that the scoring was based on information gathered from various sources: the interviews with treatment staff/director, observing program delivery, and reviewing printed material /collateral information. The interviewer had an opportunity to ask additional questions and obtain detailed information, and the scoring of the CPAI items was not a problem as long as the person interviewed had knowledge about the program and details of program delivery.

The current research developed a new approach to CPAI scoring in which the program related information provided in the articles on the treatment effectiveness were used to score individual CPAI items. The articles included in the previous meta-analyses provided limited description of the treatment conditions and a pilot study was conducted to identify potential problems in scoring. The goal of the pilot project was to determine whether the articles contained sufficient information to score the CPAI using the manual developed by Gendreau and Andrews (1996; 6th edition).

Ten randomly selected articles included in the meta-analysis done by Dowden (1998) were scored using the original CPAI manual (Gendreau & Andrews, 1996). Limited

information provided in the articles contributed to the difficulties in scoring the CPAI items. The analysis revealed that approximately 40% of the CPAI items could be scored. The Program Implementation, Other, and Staff sections were identified as the most difficult to score due to lack of information in the articles pertaining to these treatment elements. Furthermore, in order to score the items from the Preservice Assessment and Program Characteristics sections a strict criterion has to be met. The lack of detailed information about the programs prevented scoring many of the items specific for these sections.

Overall, the findings from the pilot study indicated that the researcher might not be able to score some items and/or sections. Two possible solutions to the scoring problems were considered. First, items that were problematic to score across the majority of the articles could be excluded and estimates of predictive validity calculated for the rest of the items. The second possible solution was to identify scorable items and to construct a new, shorter version of the CPAI. The main purpose of this research was to establish psychometric properties of the original CPAI, and the full original version of the instrument was assessed despite scoring limitations. The scoring manual was reformulated and this revision took into account the uniqueness of scoring the CPAI based on the limited program information reported in the articles. There were three types of modification to the original CPAI scoring: a) the amount of information required to score an item as present was not so strict and the elements could be scored as present if some, not necessarily all, indicators were present b) the items were reformulated in such a way to take into account that CPAI is scored based on the articles and not on the actual program data and c) items that required detailed specific information that was not available were collapsed into newly created variables. The original meta-analyses had a number of variables that correspond to the CPAI items and these variables were included in

the CPAI revised manual (Appendix J contains new revised CPAI manual). In summary, the modification of the items was based on the assumption that when some limited information were available, elements could be scored as present even though the strict criteria for scoring was not met.

CPAI Coding Manual: Modifications

The Program Implementation section maintained many of the items from the original CPAI manual. The two items, "pilot study" and "literature review" were modified as they required information not usually reported in the articles. For example, the original CPAI scoring manual would code the item "pilot project" as present if the pilot project lasted at least one month. However, as the information on the duration of pilot project was usually not reported, the revised scoring manual indicated that this item should be scored as present if pilot project was conducted regardless of its duration. Rather than looking for a comprehensive treatment effectiveness literature review, as required by the original scoring manual, the revised manual scored the item "literature review" as present if any literature on treatment effectiveness was reviewed. Some of the articles included in the present study were published years ago when literature on treatment effectiveness was limited at best.

The Pre-Service Assessment section generally required detailed information and therefore, the items pertaining to risk, need and responsivity were reformulated to adapt to the limited information provided in the articles. Given that the most frequent risk indicator represented in the articles was previous involvement with the criminal system, the programs delivered to participants with previous criminal history were scored as having the risk principle addressed. While the original manual focussed mainly on specific responsivity characteristics, the revised manual took into account general responsivity principle. The

scoring followed Dowden's (1998) approach in which programs addressed general responsivity when they employed behavioral intervention.

There were a number of modifications in the Program Characteristics section. The original manual asked for the specific percentage of criminogenic targets present in order to score the item "program target" as present. The revised manual allowed this item to be scored if the number of criminogenic needs targeted exceeded the number of noncriminogenic needs. Further, new items had to be created to replace original items when a distinction could not be made as to whether both items were present or only one. For example, the revised scoring manual replaced two items from the original manual ("theory described" and "effective stimuli used") that could not be distinguished in the articles, with a new item that combined these original items. For the same reason, two relapse prevention items ("rehearse" and "practice") had to be collapsed into a new variable.

The Staff Characteristics section contained a number of items that, in the original scoring manual required a specific percentage of characteristics in order to score an items as present. For example, in order to score "education" as present, it was required that 75% of the staff has a B.A. degree and 10% has a M.A. degree. As the articles did not provide such specific information, the questions were reformulated and the items were scored as present if the majority of staff had at least a B.A. degree.

The Evaluation section remained the same as in the original manual. The items within this section required very specific information and were clearly defined.

The Other section contained items that were usually not reported in the articles. Therefore, the original items were kept as it was unlikely that their reformulation will increase the number of scorable items.

Finally, the treatment elements identified in the introductory sections as potentially important could not be scored based on the information contained in the articles. Therefore, these treatment elements were not included in Study Three.

In order to establish interrater reliability indicators ten studies were provided to the second rater who was a Ph.D. student in Forensic Psychology. The interrater reliability indicators were calculated using Pearson Correlation Coefficient for the individual CPAI sections and for treatment and total scores for both treatment and control group. Results indicated that interrater correlations were .88 for treatment and total CPAI scores for scoring of the treatment group, while the level of agreement was lower for the individual CPAI sections: Implementation (.74), Pre Service Assessment (.85), Program Characteristics (.85), Staff (.78), Evaluation (.67). The level of agreement for scoring of the control group was lower, i.e. Treatment (.79) and Total (.77) CPAI score, Preservice Assessment (.74). The interrater reliability estimates for scoring of control groups could not be computed for the rest of the CPAI sections due to presence of at least one constant variable (e.g., all values for section other were 0 for one rater). Limited information in the articles enabled the scoring of many items and contributed to the lower interrater reliability indicators for the control group.

Scoring the CPAI: Missing Data

Notwithstanding modifications of the CPAI scoring manual described in the previous section, a considerable number of data was still missing, as the articles did not provide sufficient information to score individual CPAI items. Most notably, two of the CPAI sections Other and Program Implementation contained a higher number of missing data in comparison to other sections. The Program Implementation section contained almost 80% missing data, mainly related to the characteristics of the program director. Furthermore, the section Other,

containing items pertaining to the background/organizational elements of the program, was not scorable at all. The information needed to score this section was too detailed to be found in the articles on treatment effectiveness.

A high percentage of missing data within the Staff section is not surprising, as treatment effectiveness literature has paid a limited attention to staff related items. The Program Evaluation section contained a number of the items pertaining to the research evaluation of the treatment (e.g., “control group”, “follow up period”). As previously noted, these items were part of article selection criteria and were present in all articles.

Finally, the Pre-Service Assessment and Program Characteristics sections had the lowest percentage of missing data. The possible explanation is that a) articles provided more information pertaining to treatment elements from these sections (e.g., “treatment targets”, “intervention used”, “risk/need/responsivity assessment”) and b) number of items within these sections were modified and these revisions increased number of items that were scored.

Scoring of the control group proved to be very challenging, as the description of the control conditions was very limited in most of the studies. Control group conditions were diverse with a number of studies reporting on control conditions with no treatment or minimal treatment (30%) and only limited number of control conditions with alternative treatment (15%). The studies with alternative treatment control groups also provided limited information. The majority of the control group conditions were marked as treatment as usual (e.g., probation, parole, institutionalization, school) without detailed explanation. The analysis of missing data for the control group conditions, revealed that the Program Implementation and Other sections had the highest and the Preservice Assessment and Program Characteristics sections the lowest number of missing data.

Prior to statistical analysis, all missing data were coded as “0”. This rule was implemented when there was no indication that the treatment element was present or when data was not available to score the item. Consequently, the percentage of the items present within each sections was calculated from the total number of items. This coding approach might underestimate the percentage of the elements present. Interpretation of the results should take into account this methodological weakness and the conclusion will have to be drawn taking into account these limitations.

Sample

The sample of studies was selected from the set of the articles included in Dowden (1998) meta-analytic study. Dowden’s (1998) sample consisted of four different sets of studies: Whitehead & Lab, (1989), Andrews et al., (1990), Andrews et al., (1996), and a sample added by Dowden (1998), resulting in a total sample of 225 articles that yielded 374 effect sizes. The studies were included if they met the following selection criteria: a) offering a correctional treatment program to adult or juvenile offenders in which there was some form of interventions offered that was compared to a control group not receiving the same intervention, b) the intervention used was described c) the number of participants was known d) a measure of recidivism was used and e) data on recidivism in a follow up period were available.

In addition, in the previous part it was emphasized that the CPAI was designed to evaluate the potential of a treatment program to reduce recidivism rates. Treatment programs were defined as structured intervention with a goal of reducing re-offending for its participants. Some of the intervention that were included in previous meta-analyses contained justice intervention that could not be considered structured interventions and as such could not

be scored on the CPAI. Therefore, the studies that included justice intervention were excluded from the final sample. A few unpublished manuscripts were also excluded from the sample as the principal researcher was not able to obtain copies. The final sample included in this study consisted of 173 articles (published and unpublished) which yielded 266 effect sizes.

Hypotheses

Hypothesis 1: It is expected that larger differences between treatment and control groups on total scores of the CPAI will be associated with higher effect sizes. Preliminary research findings of Latessa et al. (1999) suggest that higher scores on the CPAI were associated with more successful programs.

Hypothesis 2: It is hypothesized that larger differences between treatment and control groups on treatment scores of the CPAI will be associated with higher effect sizes. The treatment scores cover the elements that the empirical literature has identified as most important for treatment effectiveness.

Hypothesis 3: It is hypothesized that a larger difference between treatment and control groups on two CPAI sections, Pre-service Assessment and Program Characteristics, will be associated with higher effect sizes compared to the rest of CPAI sections. The results of previous meta-analyses (Andrews et al. 1990, Andrews, 1996; Dowden, 1998) have found that items within these subcomponents (e.g., risk, need, and responsivity) are significantly associated with effective treatment reflected by the effect size.

Results

Sample Descriptives

The final sample included in the present study consisted of 173 articles on treatment effectiveness that yielded 266 effect sizes. The analysis of the study characteristics revealed

that 50% of the articles were based on a male population. Female offenders were represented in only 8% of the studies, while the samples for the rest of the programs were mixed. Studies that addressed treatment effectiveness of young offenders (61%) were predominant.

Almost half of the studies (48%) were based on community delivered programs while institutional/residential delivered program were present in 36% of the studies. The remaining studies focussed on system and nonsystem diversion programs and special programs. The most represented referral source for the program participation was the criminal justice system with program participation that was either mandatory (67%) or voluntary (20%). The rest of program participants were referred by various sources (e.g., friends, family, schools, job, and mental health).

The analysis of studies pertaining to the level of information about the programs evaluated revealed that only 25% of the programs contained detailed description of the programs while 44% of the sample had just a general statement about the programs. The remaining studies used only labels or phrases about treatment conditions.

Every treatment conditions was compared to the corresponding control group. Studies reported on diverse control group conditions with 30% of the studies having control group that received nothing, was wait listed or had minimal treatment exposure. The rest of studies reported the following control groups conditions: treated as usual (12%), treated as usual - school (3%), treated as usual -probation (15%) and treated as usual institutionalization (18%). Finally, 15% of the control groups consisted of alternative treatment conditions. The remaining studies had unspecified (other, placebo) control conditions.

Internal consistency estimates were calculated for both treatment and control group. The values of Chronbach alpha were the same for treatment.81 (mean interitem correlation

.06) and .81 for control group (mean interitem correlation .10). Treatment scores had alpha values of .71 (mean interitem correlation .07) and .70 (mean interitem correlation .09). The obtained internal consistency estimates for the meta-analytic study were lower in comparison to the values obtained when the CPAI was scored based on the interview (total score, $\alpha = .93$; treatment score $\alpha = .85$) and the questionnaire (total score, $\alpha = .88$; treatment score $\alpha = .83$). The most likely explanation is that considerable number of missing information in the articles led to deflated alpha values in the meta analytic study.

Mean Effect Sizes

Although, the meta-analytic literature has a number of different techniques used for computing effect sizes (Rosenthal, 1991; Wolf, 1986), this study as a continuation of the previous meta-analytic studies (i.e., Andrews et al. 1990; Dowden, 1998), adopted phi coefficient, the effect size measure used in previous studies.

The overall mean effect size for the sample was .12 ($SD = .19$) with a 95% confidence interval of $+.09$ to $+.14$. This mean value for the effect size is slightly higher than the mean effect size (.08) reported by Dowden (1998). As the previous meta-analytic studies (Andrews et al, 1990; Dowden, 1998) revealed that justice interventions yielded lower effect sizes compared to the rest of program categories, the difference is likely a result of the exclusion of the justice system intervention from the present study.

The results revealed a considerable variability of the effect sizes within the study sample ranging from $-.43$ to $.82$. Dowden (1998) suggested that some variability within the sample could be accounted for by the type of intervention used. The present study also resulted in different mean effect size values for non behavioral (.07), radical behavioral (.17), cognitive behavioral (.21), and social learning (.26) approaches. The analysis of variance

revealed that the outcome was significantly different depending on the type of intervention used $F(1, 265) = 12.05, p < .01; \eta^2 .32$. Scheffe comparisons identified that the only difference that reach statistical significance was that the social learning approach had significantly higher mean effect size compared to nonbehavioral programs. It seemed that the type of the intervention used can explained part of the variability within this study.

CPAI Sections – Descriptives

Information from the articles on program effectiveness describing treatment and corresponding control conditions were scored using the revised CPAI manual. The percentage of the elements present within CPAI sections, treatment and total score was calculated out of the total number of scorable items. The Table 15 summarizes the means and standard deviation for the treatment and control group.

Table 15:
CPAI Descriptives for Treatment and Control Groups

CPAI Sections (N=266)	Treatment		Control		t-test	Sig.
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>		
Program Implementation	.03	.05	.01	.02	9.08	.01
Preservice Assessment	.30	.13	.14	.11	23.79	.01
Program Characteristics	.15	.11	.02	.05	18.45	.01
Staff Characteristics	.14	.16	.03	.09	12.30	.01
Program Evaluation	.59	.09	.52	.05	13.24	.01
Other	.00	.00	.00	.00		
Total Score	.20	.07	.10	.04	23.49	.01
Treatment Score	.21	.10	.07	.06	24.22	.01

A review of Table 15 reveals low CPAI mean values across treatment and control conditions. There was considerable variability across different CPAI sections with some

sections having none of the elements scored (e.g., Other) and higher mean values for some section (Evaluation). The higher mean values for Program Evaluation (.59 for treatment and .52 for control conditions) in comparison to the rest of sections were likely a result of the items from this section (e.g., “recidivism rates”, “follow up period”) that were used as a selection criteria for studies included in previous meta-analyses. As previously noted, the section Other could not be scored as it consisted of items requiring miscellaneous program information, which was underreported in the literature. The slightly higher values were obtained for the Preservice Assessment (.30 treatment group; .14 control group) while the Program Characteristics and Staff had on average 15% of the elements present for treatment and close to 3% for control group.

Overall, the mean values for treatment (.20) and total (.21) scores within the treatment sample indicated that approximately 20% of the elements were present. Within the control group the values had the similar pattern with treatment and total score being .10 and .07 respectively, revealing relatively low numbers of the items present. (see Appendix K for variability statistics).

T tests were computed in order to determine whether the differences between means were significant. The first set of analyses indicated that all of the comparisons between treatment and control group were statistically significant and that the difference can not be explained by chance alone. Mean values for the treatment group were significantly higher than mean values for the control group across all CPAI sections, treatment and total CPAI scores. Additional analyses were conducted to determine whether CPAI sections were different within treatment and control groups. Difference among all sections was significantly

different with the exception of the difference between Staff and Program characteristics section, that were scored in a similar way for both treatment and control group.

In summary, mean CPAI values varied significantly across different sections and between treatment and control groups. The low obtained values for the CPAI scores across different sections and for total and treatment score cannot be explained solely by low number of treatment elements that these programs included. The lack of information in the articles which led to scoring the number of the items as absent, contributed to lower mean values for these programs. This information likely reflected reported quality of program, rather than its real quality.

Summary Levels across CPAI Scores

According to the CPAI manual, the programs that had less than 50% elements present were coded as “unsatisfactory”. Additionally, the programs were considered as “satisfactory” when 50-69% elements were present and “very satisfactory” when they incorporated over 70% of the elements. These summary levels were used for both the individual CPAI sections and treatment and total CPAI scores.

The analysis revealed that the majority of the CPAI scores fall in the unsatisfactory category for both treatment and control group. The results for the treatment group indicated that all scores for the Program Implementation and Other section and the majority of scores for the Staff (98%), Program Characteristics (99%) and Preservice Assessment (94%) sections were in unsatisfactory range. The only section in which the majority (80%) of the program was in satisfactory category was the Program Evaluation section. The rest of the programs was placed in very satisfactory (19%) and unsatisfactory category (1%). Finally, all of CPAI total scores and 99% of CPAI treatment scores were in unsatisfactory range.

The results for the control group followed a similar pattern of the results. All scores for the Program Characteristics, Implementation, and Other sections and the majority of the scores for the Pre Service Assessment (98%) and Staff Characteristics (99%) sections were in unsatisfactory range. The Program Evaluation section was the only section with 97% of scores being in satisfactory range. Finally, all of the CPAI total and treatment scores were within unsatisfactory range.

Variability of the sample expressed by both standard deviation and range indicated considerable variability across different programs. While some of the mean CPAI scores were high (e.g., .88) other scores were very low (e.g., .00). Notwithstanding the variability indicators within control and treatment group, the majority of CPAI scores fell in unsatisfactory category. Still, this category with the values from 0-49% had a considerable variability as the majority of mean values were within this category.

Low mean values of the CPAI scores can be potentially explained by two factors. First, a limited amount of information prevented adequate scoring. Some of the programs might have been scored differently had some additional information about treatment condition been available. Scoring the absent and not applicable items as “0” contributed to the lower mean values. Second, many of the delivered programs were from the time when treatment effectiveness literature was scarce and the knowledge about what constitutes effective treatment was fragmented. This probably led to underreporting of some potentially significant elements. Therefore, CPAI results reflect the quality of the program as reported by articles and not necessary the real quality of the program.

Predictive Validity Estimates

As noted previously, the predictive validity study was conducted in three phases. First, effect sizes, expressed as phi coefficient, were calculated for each of the articles. For articles that contained more than one treatment-control comparison, effect sizes were calculated for each of the comparison sets. Effect size reflects the difference in recidivism rates of treatment and control group.

Second, each of treatment and control conditions described in the articles was scored using the revised CPAI manual and the percentage of treatment elements present across different CPAI sections was calculated for a) treatment group b) control group and c) difference between treatment and control group.

Finally, correlations between a) treatment group and effect size b) control group and effect size and c) difference between treatment and control group and effect size were calculated. Table 16 summarizes these correlations.

Inspection of the table revealed that the correlations between effect sizes and CPAI scores for the treatment groups were statistically significant. The weakest, albeit statistically significant correlation was found for the Program Implementation (.15) and the strongest for Program Characteristics (.46). The correlations were even higher for total (.50) and treatment scores (.52). These results indicated that a higher number of the CPAI items present was linked to higher values of effect size. If the effect size was considered as an indicator of treatment effectiveness then it could be concluded that higher CPAI scores would possibly lead to a more effective program.

The results from previous meta-analyses have indicated that when principles of risk, need and responsivity were taken into account and when the delivered intervention was

Table 16:
Bivariate Correlations of CPAI Scores with Effect Sizes

CPAI Sections (N=266)	Treatment r	Control r	Difference r
Program Implementation	.15*	.15*	.10
Pre service Assessment	.44**	.13*	.41**
Program Characteristics	.46**	.07	.43**
Staff Characteristics	.32**	.17**	.24**
Program Evaluation	.29**	.07	.27**
Other			
Treatment Score	.52**	.13*	.49**
Total Score	.50**	.16**	.46**

* $p < .05$

** $p < .01$

cognitive behavioral, the programs yielded higher effect sizes (Andrews et al. 1990; Dowden, 1998). CPAI sections Preservice Assessment and Program Characteristics consisted of these treatment elements, therefore the magnitude of the correlation between these two subcomponents and the effect sizes (.44 for Pre-Service and .46 for Program Characteristics) supported the findings from the previous meta-analyses.

Program Implementation, Evaluation and Staff scores demonstrated lower, but still statistically significant correlations with the effect size. These results indicated that these sections contribute to treatment effectiveness, however, their contribution was not as high as for the CPAI treatment sections. The high correlations between CPAI total scores with the effect size indirectly indicated that a program will likely be more effective if more of the CPAI items were included in a program. In addition, the high value of the correlation between the treatment CPAI score and effect size, suggest that the most significant contributors to

effective programs were treatment sections (Preservice Assessment and Program Characteristics).

When the correlations between control group scores and corresponding effect sizes were analyzed, the values were lower in comparison to the treatment group, however, all correlations with exception of the Program Implementation and Evaluation section were statistically significant. The significant positive correlation between CPAI treatment (.13) and total scores (.16) and effect size was found for the control group conditions. That is, the more of the CPAI items that were present within the control group the more likely the control condition will be effective. The lower obtained correlations within the control group could be explained by a lower number of elements present within control group conditions. Interestingly, the results remained consistent, with more items from the CPAI scores present in the group resulting in more effective intervention even for control groups.

Given that the effect size is an indicator of the difference in recidivism rates between treatment and control group, estimates of predictive validity of the CPAI scores were obtained when effect size values were correlated with the difference in CPAI scores between treatment and control group. The same pattern of results emerged with all correlations reaching statistical significance (with the exception of Program Implementation). The highest values among CPAI sections were obtained for Pre service assessment (.41) and Program Characteristics (.43).

Strong positive statistically significant correlations were obtained for both treatment (.49) and total (.46) CPAI scores with effect size. These results confirmed two general hypotheses, indicating that the larger difference between treatment and control group on treatment and total CPAI score, will lead to the higher effect sizes values. It means, when the

treatment group included more of the CPAI items than the control group, it is likely that the treatment group will have less reoffending in comparison to the control group. These results represent preliminary indicators of good predictive validity.

The third hypothesis which emphasized the significance of the treatment sections (Preservice Assessment and Program Characteristics) for program effectiveness has also been confirmed. Results show that the larger difference between Preservice Assessment and Program Characteristics scores for treatment and control groups were associated with higher effect sizes compared to the rest of subcomponents. These results support the previous finding which indicated the importance of these two CPAI section for treatment effectiveness.

Overall, all CPAI sections (with exception of Program Implementation) were significantly correlated with the effect size. The difference between Program Implementation scores for treatment and control group was the lowest in comparison to all other CPAI sections. Additional analysis revealed that the percentage of missing items within this section was the highest and the mean values lowest in comparison to other sections. The same was applicable for the Other section, as none of its items could be scored, and therefore the correlation could not be calculated. Nevertheless, it cannot be concluded that these sections are not important for treatment effectiveness and should be excluded. Previous literature review indicated that these sections will likely have a positive but not necessarily high relationship with treatment outcome as they consist of the background/organizational elements that are distantly related to treatment outcome.

Intercorrelations among the CPAI scores were statistically significant within treatment group. The analysis of their magnitude revealed that the highest correlation with total scores was obtained for Preservice Assessment .78 and Program Characteristics. 84. Consequently,

the link between treatment score and total score reached very high values of .94. The rest of sections had moderate to high correlation with total score. The same pattern of results was obtained for the correlations among CPAI sections within the control group and when the differences between treatment and control group were taken into account. Most specifically, the treatment scores had a very high correlation with the total score within control group (.94) and difference between treatment and control group (.92).

Regression Analyses

Regression analysis was conducted to establish whether CPAI section scores predict treatment outcome, i.e. effect size. When individual CPAI sections were entered in regression analysis, each section was identified as significant predictor with the exception of section Other that was excluded from the analyses as the items from this section could not be scored. However, the percentage of variance in effect size that could be attributed to the independent sections was low for Program Implementation (2%), Evaluation (9%) and Staff (10%) and higher for Preservice (19%) and Program Characteristics (21%) sections.

When all of CPAI sections were entered in multiple regressions, the overall R increased to .52 and the percentage of variance in the effect sizes contributed to all five sections increased to 27% indicating a significant relationship between these sections and effect size, $F(1,265) = 19.54, p < .01$.

Additional regression analyses were conducted to determine the contribution of individual CPAI sections to the effect sizes when statistical controls were introduced. When Preservice Assessment and Program Characteristics scores were entered in the regression analyses first, the overall R was .52 and the rest of CPAI sections did not contribute significantly to the effect size beyond the contribution of these sections (Overall R was .52

and the percentage of the variance explained remained the same when the new variables were entered). When Program Characteristic and Preservice Assessment were partialled out in the regression equation, the Staff, Implementation and Evaluation sections failed to maintain significant relationships with effect sizes. None of their partial correlations were statistically significant. However, when the rest of the sections were entered into regression analyses, Preservice Assessment and Program Characteristics scores remained significant contributors to the effect size.

Individual CPAI Items

Even though the main goal of this study was to link CPAI scores to the effect size, additional analyses pertaining to the individual CPAI items were conducted. The Pre Service Assessment and Program Characteristics sections were identified by both regression and correlation analyses as significant contributors to effect size. Previous meta-analyses (Andrews et al, 1990; Dowden 1998) have identified some of the treatment elements included in these sections as the most important for treatment effectiveness, more specifically risk, need and responsivity characteristics. The analyses of these elements within the present study revealed the same pattern of the results with higher mean effect sizes values in comparison to Dowden's (1998) results. The higher values were likely a result of the exclusion of the justice intervention from the sample.

Studies that took risk factors into account had higher mean effect sizes (.13, $n=213$) than studies that did not consider risk factors (.05, $n=53$). The statistical significance of this difference, $F(1,265)=8.70$ $p<.01$, $\eta^2=.18$, revealed the importance of risk factors for treatment effectiveness expressed as effect size.

A strong positive effect was found when the need principle was taken into account (.19, $n = 153$) while the failure to account for the need principle led to a considerably lower (.02, $n = 113$) mean effect size. An analysis of variance indicated that the need principle significantly contributed to the treatment outcome $F(1,265)=63.05$ $p<.01$, $\eta^2 .44$.

The effect size for the programs that had responsivity characteristics taken into account (.20, $n = 80$) was higher than when responsivity factors (.07, $n = 186$) were not taken into account. The difference in the magnitude of effect size was statistically significant, $F(1,265)=25.55$ $p<.01$, $\eta^2 .30$.

Additional analyses were conducted for two of the main elements of program characteristics (i.e., “types of intervention used” and “targets”). The analysis of variance revealed that programs targeting criminogenic needs (.19, $n = 162$) yielded significantly higher effect size than the programs targeting non criminogenic needs (.01, $n = 104$) $F(1,265)=77.54$ $p<.01$, $\eta^2 .48$.

Statistically significant difference $F(1, 265) = 31.43$; $p<.01$; $\eta^2 .33$) was found between programs that used behavioral interventions (.21, $n = 75$) and programs with nonbehavioral interventions (.07, $n = 191$).

Finally, the rest of individual treatment elements and their contribution to the effect sizes were also explored. Table 17 summarizes the mean values for the variables found to be significantly related to the effect sizes and the results of ANOVA.

These results identified a number of the individual CPAI items significantly linked with effect sizes. Majority of the items that significantly contributed to effect sizes were from two treatment sections, Preservice Assessment and Program Characteristics. Previously analyzed results noted when risk, need and responsivity factors were assessed, they were significantly

Table 17:
Individual CPAI Items: Mean Effect Sizes, F Tests and Eta

CPAI item	Element present		<u>F</u>	<u>sig</u>	eta
	No	Yes			
Program Implementation					
Program Director initiated	.11(263)	.34(3)	.44	.05	.13
Literature Review	.11(237)	.19(29)	4.73	.05	.13
Preservice Assessment					
Appropriate Clients	.02 (37)	.14(229)	25.17	.01	.30
Assigned based on risk	.06 (65)	.13(201)	5.35	.05	.14
Assigned based on need	.08(203)	.22 (63)	30.37	.01	.32
Assigned based on responsivity	.11 (254)	.32 (12)	15.69	.01	.24
Program Characteristics					
Printed Manual	.08(197)	.20 (69)	17.96	.01	.25
Matching Staff/Program	.11(262)	.46 (4)	14.83	.01	.23
Matching Staff/Client	.11 (258)	.34 (8)	12.48	.01	.21
Disapproval	.11 (259)	.29 (7)	6.53	.05	.15
Relapse Prevention Rehearse	.11(250)	.21 (16)	4.07	.05	.12
Relapse Prevention Plan/Practice	.09(233)	.28 (33)	30.74	.01	.32
Brokerage	.11(240)	.20 (26)	6.42	.05	.15
Significant Others	.10(219)	.19 (47)	8.43	.01	.18
Staff Characteristics					
Staff Experience	.11 (249)	.24 (17)	7.84	.01	.17
Relationship factors	.10(254)	.35 (12)	21.66	.01	.27
Skill factors	.09(226)	.25 (40)	26.75	.01	.30
Supervised staff	.09(213)	.19 (53)	10.46	.01	.19
Trained Therapists	.06(133)	.16(133)	17.56	.01	.25

Table 17:
Individual CPAI Items: Mean Effect Sizes, F Tests and Eta (Continued)

CPAI item	Element present		F	sig	eta
	No	Yes			
Program Evaluation					
Quality Assurance	.11 (259)	.35 (7)	12.06	.01	.21
Evaluator Involved	.07 (194)	.22 (72)	38.02	.01	.35

linked to the effect sizes. Additionally, when these elements were taken into account when assigning offenders to the programs, the link with treatment outcome was significant as evident from Table 17. Presence of printed manuals and matching staff/program and staff/clients yielded higher effect sizes than when these elements were not taken into account. However, it should be noted that the number of studies addressing these elements was very low. The new variable, that consisted of two original CPAI items (theory of punishment and effective stimuli), were also linked to the effect size. Finally all of the items pertaining to relapse prevention were significantly correlated with effect size.

A number of individual items from the Staff Characteristic section also contributed significantly to the effect size, most notably “relationship factors”, and “skill factors”, “trained therapist”, “supervised” staff, and “staff experience”. The only items from other CPAI sections that contributed significantly to the effect sizes were “literature review”, “quality assurance” and “involved evaluator”. (Appendix L summarizes eta values for all individual items).

Treatment Level

The CPAI treatment scores were recoded according to the percentage of items present: programs with scores between .00 and .13 were coded as “1” (very low), between .14

and .23 as “2” (low), between .24 and .33 as “3” (moderate) between .34 and .58 were coded “4” (high) treatment score levels. The analysis of frequencies for this newly created variable named Treatment Level, revealed that 30% of the programs scored in very low range, 32% in low range while the rest of the programs scored in moderate (26%) and very high(12%) range. Additional analyses were conducted to determine the impact of Treatment Level on recidivism (effect size). The different levels yielded different mean effect sizes with the weak mean effect size. 01 ($n=80$) obtained for “very low” treatment level. The mean effect sizes were higher for “moderate” (.11, $n=86$), “high” (.20, $n=68$) and “very high” (.25, $n=32$) treatment levels. The analysis of variance revealed that the outcome was significantly different depending of the different treatment score level $F(3,265)=29.39, p<.01; \eta^2=.50$). Scheffe comparison indicated that treatment score levels were significantly different with the exception of the difference between the “high” and “very high” level which did not reach statistical significance.

Further analyses explored the relationship between the Treatment Level and New Appropriate Treatment, a composite variable discussed in the introduction. The simple correlation between the Treatment Level and New Appropriate Treatment was high and statistically significant ($r=.79, p<.01$). In order to determine whether the Treatment Level contributes to the variability in the effect size beyond the contribution of New Appropriate Treatment, both variables were entered in a multiple regression analysis. When New Appropriate Treatment was entered into the regression analysis the overall r was .49 and the percentage of variance in the effect size that could be contributed to the New Appropriate Treatment was 24%. Upon the entry of Treatment Level the overall R increased to .53 and the

percentage of variance in the effect size explained by these two variables increased to 28%, indicating that the Treatment Level had incremental value.

The relationship between the Treatment Level and the effect size was further analyzed under variety of control conditions that can potentially influence the treatment outcome. The goal of these additional analyses was to establish whether the results would hold under different treatment, setting/organizational and case characteristics. Internal, external and theoretical validity considerations were also taken into account. Table 18 summarizes values of eta and the mean effect sizes across different levels of control conditions.

As the Table 18 shows, the Treatment Level remained significantly correlated with the effect size at different levels of control variables. With respect to treatment variables, Treatment Level was significantly correlated with the effect size for both appropriate (.33) and inappropriate treatment (.28), for the programs using behavioral (.31) and non behavioral interventions (.47) and also for the programs maintaining (.46) and not maintaining (.59) the integrity of treatment implementation. When the contribution of Treatment Level to the effect size was examined for setting and organizational factors, the Treatment Level was significantly correlated with the effect size for non restorative justice programs (.51). However, the Treatment Level was highly (.77) but not significantly correlated with the effect size for the restorative justice programs. Non significance is likely a result of the small number of restorative justice interventions ($n=7$). The Treatment Level significantly influenced the magnitude of the effect size for community (.60) and institutional based programs (.44); for the programs delivered by criminal justice (.49) and other (.57) agencies. When case characteristics were analyzed, the Treatment Level contributed significantly to the

Table 18:
Mean Effect Sizes by CPAI Treatment Level and Control Variables

Control Variable	CPAI Treatment Level				Eta
	Very Low	Low	Moderate	High	
Treatment Characteristics					
Appropriate Tx	.03(6)	.14(41)	.23(60)	.26(32)	.33**
Inappropriate Tx	-.01(74)	.07(45)	.00(8)		.28**
Behavioral	-.09(2)	.15(6)	.21(35)	.25(29)	.31*
Non behavioral	-.01(78)	.11(80)	.20(33)	.35(3)	.47**
Integrity Yes	.06(12)	.11(14)	.10(15)	.30(6)	.46*
Integrity No	-.02(53)	.11(53)	.25(40)	.33(14)	.59**
Setting Organization Factors					
Restorative	.10(2)	.16(3)	.17(2)	.35(1)	.77
Non restorative	-.01(78)	.11(83)	.21(66)	.26(31)	.51**
Community	.00(61)	.11(55)	.27(34)	.32(21)	.60**
Institution	-.05(19)	.10(31)	.14(34)	.14(11)	.44**
Criminal justice	.00(34)	.12(30)	.09(24)	.23(11)	.49**
Non criminal justice	-.01(46)	.10(56)	.27(44)	.27(21)	.57**
Case Characteristics					
Violent	-.02(26)	.10(25)	.22(28)	.23(17)	.54**
Non violent	.00(54)	.11(61)	.20(40)	.29(15)	.50**
Young	.01(40)	.10(48)	.20(24)	.26(12)	.52**
Adult	-.03(40)	.12(38)	.20(44)	.26(20)	.51**
Female	.03(11)	.13(9)	.25(12)	.55(2)	.60**
Male	-.01(69)	.11(77)	.19(56)	.24(30)	.51**

Table 18:
Mean Effect Sizes by CPAI Treatment Level and Control Variables (Continued)

Control Variable	CPAI Treatment Level				Eta
	Very Low	Low	Moderate	High	
Ethnic	.00(64)	.11(64)	.22(51)	.25(23)	.53**
Non ethnic	-.04(16)	.12(22)	.15(17)	.27(9)	.50**
Internal Validity Considerations					
Attrition Yes	-.01(32)	.12(33)	.20(45)	.30(15)	.49**
Attrition No	-.01(33)	.10(34)	.25(10)	.37(5)	.62**
Random	-.02(37)	.14(29)	.23(29)	.25(20)	.58**
Non random	.00(43)	.09(57)	.18(39)	.28(12)	.46**
Internal evaluator	-.08(4)	.15(18)	.27(32)	.31(19)	.50**
External evaluator	.00(76)	.09(68)	.15(36)	.19(13)	.40**
External Validity Consideration					
Small sample	-.08(16)	.14(22)	.27(31)	.36(13)	.59**
Big sample	.02(49)	.10(45)	.13(24)	.25(7)	.45**
Follow up (1year+)	.00(49)	.13(43)	.17(34)	.36(13)	.59**
Follow up (<1year)	-.05(16)	.07(24)	.27(21)	.25(7)	.54**
Follow up (2 years+)	.00(32)	.13(27)	.14(18)	.30(6)	.54**
Follow up (<2 years)	-.02(33)	.09(40)	.24(37)	.33(14)	.55**
Theoretical Consideration					
Theoretical	.00(41)	.12(51)	.19(52)	.27(28)	.48**
Atheortiecal	-.02(39)	.09(35)	.24(16)	.19(4)	.54**
Tx model -specific	.00(30)	.11(40)	.21(56)	.25(29)	.48**
Tx model - vague	.00(50)	.10(46)	.16(12)	.35(3)	.48**

effect size for violent (.54) and non violent (.50), young(.52) and adult (.51), female (.60) and male (.51), ethnic (.53) and non ethnic (.50) samples.

Additional analyses also addressed internal validity considerations that can potentially influence treatment outcome, most notably attrition problems, random assignment to treatment condition and the presence of internal/external evaluator. The Treatment Level significantly contributed to the effect size under these consideration: across the studies with (.49) and without (.45) attrition problems, and studies with random (.58) and non random (.46) assignment of subjects to treatment conditions. The same pattern was evident when the program evaluator was internal (.50) or external (.40).

The significant contribution of Treatment Level to the effect sizes persisted when external validity factors were taken into account: for small (.59) and large (.45) samples, for the studies with follow up period under (.59) or over 1 year (.54), or under (.54) or over 2 years (.55). Furthermore, analyses that tested the relationship under different theoretical considerations had the same results: Treatment Level was significantly associated with the mean effect sizes across programs with theoretical (.48) or atheoretical (.54) approach, and for programs using specific model (.48) or vague model (.48). These robust results are evident across number of control conditions, and suggest that the Treatment Level enhances the treatment outcome.

Additional analyses were conducted in respect to the treatment intervention targeting criminogenic and noncriminogenic needs. Table 19 summarizes the values of eta and mean effect sizes for criminogenic and non criminogenic program targets.

The inspection of the Table 19 shows that the Treatment Level contributed significantly to the effect sizes across treatment interventions which targeted individual

Table 19:
Mean Effect Sizes by CPAI Treatment Level and Treatment Targets

Targets	CPAI Treatment Score				Eta
	Very Low	Low	Medium	High	
Criminogenic Targets					
Antisocial feelings	.02(3)	.14(12)	.20(21)	.26(21)	.28*
Family intervention	.00(13)	.08(9)	.24(16)	.32(11)	.59**
Reduce antisocial peers	.04.(1)	.13(8)	.13(2)	.24(8)	.45*
Promote anticriminal peers	.07(3)	.11(14)	.27(12)	.28(7)	.43*
Substance abuse	.06(8)	.10(19)	.18(7)	.32(2)	.48**
Vocation employment	.02(10)	.11(11)	.32(11)	.34(6)	.65**
Academic	.04(15)	.12(23)	.25(18)	.32(10)	.52**
Anger/Self control	.03(7)	.16(20)	.21(32)	.26(22)	.30*
Noncriminogenic Targets					
Self esteem	-.14(5)	.06(3)	.00(2)	.24(3)	.71*
Emotional problems	.02(34)	.38(26)	.18(16)	.25(2)	.40**
Conventional ambition	.00(13)	.10(7)	.17(3)	.56(1)	.78**
Physical Training	-.04(14)	.04(10)	.12(4)	.35(8)	.65**

criminogenic and noncriminogenic needs. Significant correlations between the Treatment Level and the effect size was found for both criminogenic (antisocial feelings,.28; family interventions,.59; antisocial peer association,.45; anticriminal peers,. 43; substance abuse,.48; anger/self control,.30; vocation,.65; academic,.52) and non criminogenic targets (self esteem,.71; emotional problems,.40; conventional,.78; and physical training,.65).

In summary, Treatment Level was a significant contributor to the variability in effect sizes and its effect remained significant under different levels of control conditions, across different treatment, case and setting and organizational characteristics, under internal,

external, and theoretical validity consideration, when programs were targeting criminogenic or non criminogenic needs.

Discussion

The development of the CPAI (Gendreau & Andrews, 1996) has resulted in improvements of the design and delivery of correctional programs. Recent research has focused on its usefulness in predicting treatment effectiveness. The main goal of the present study was to determine the importance of CPAI items in the delivery of effective correctional programs and to establish whether CPAI scores were linked to treatment outcome.

The results of the current study provided support for the validity of the CPAI items in the delivery of effective correctional service. They demonstrated that CPAI scores were significantly related to the measure of treatment effectiveness, i.e. effect size. The relationship was established across different sections and within treatment, control groups and the difference between treatment and control group. The general hypotheses have been confirmed. More specifically, the larger differences between treatment and control group on treatment and total CPAI scores were linked with higher effect sizes. In addition, the larger difference between treatment and control group on Preservice Assessment and Program Characteristics were associated with higher effect sizes in comparison to other CPAI sections.

When CPAI scores were entered into a regression analysis to determine whether they will predict treatment outcome, it was consistently revealed that programs with higher CPAI scores produced more positive outcome. Even though all of the CPAI sections with exception of section Other were significantly linked to the treatment outcome, sections Preservice Assessment and Program Characteristics were the main contributor to treatment outcome.

When these sections were partialled out, the contribution of the rest of the CPAI sections to treatment outcome was not significant.

These results strongly suggest that programs with a higher number of CPAI treatment elements included will likely have lower reoffending rates in comparison to the control group, especially if sections Preservice Assessment and Program Characteristics were adequately addressed. This finding indicates that the CPAI is a valid inventory for the assessment of potential of correctional treatment programs to reduce reoffending.

When CPAI treatment scores were recoded according to the percentage of the items present in low, very low, moderate and high treatment level categories, the correlation between treatment scores and the treatment outcome was even more evident. Treatment level was a major source of variability in the effect size under a variety of statistical and methodological control conditions. The positive effect of Treatment Level was evident at different levels of control variables and Treatment Level was positively associated with treatment outcome (reduced reoffending) in appropriate and inappropriate treatment intervention, in behavioral and nonbehavioral programs, in programs targeting criminogenic and noncriminogenic needs, in programs with high and low treatment integrity, in community and institutional programs, in programs delivered by criminal justice and other agencies, in theoretical and atheortiecal programs, in program with a specific and vague theoretical model, with violent and non violent sample, with men and women, with ethnic minority and ethnic majority, with younger and older offenders, in small and large samples, in studies with different length of follow-up, in studies with and without attrition problems, in random and non random designs, and in studies with internal or external evaluator.

The methodological limitations of the study should be noted. First, the coding method used in this study represented a modification of the original CPAI scoring. The new approach in scoring the CPAI was developed in order to be able to score the CPAI based on the information provided in the articles. The percentage of missing data would be even higher had we relied solely on the original coding manual. Second, the scoring of missing items as 0 (element not present) could be misleading. Some of the items might have been present had the additional information been available. It is likely that the obtained relationship would have remained significant, possibly even higher, had more details on the programs were available. Third, the lack of direct treatment outcome measure and scoring of the CPAI based on program description from the articles questions the generalizability of the results. However, these results were taken as estimates of predictive validity rather than direct indicators of the validity itself. Finally, there is a certain circularity of the establishing predictive validity estimates by using articles on treatment effectiveness. CPAI is an instrument that summarized elements identified by meta-analytic studies. The scoring of the articles based on the same set of studies, could questions the adequacy of this approach. However, the previous meta-analytic studies did not provide data on the specific items, but on categories related to treatment elements. Additionally, many of the CPAI elements were not identified by meta-analytic studies but by different empirical research and clinical experience. Even though there is a possibility that some of the items are overlapping, the majority of them remained outside of the scope of previously conducted meta-analyses.

Despite the limitations noted above, it is clear that there is a link between CPAI scores and effect size. This dissertation does not represent a definite demonstration of the relationship between CPAI and recidivism. Additional research using a design within a

controlled evaluation study directly linking the CPAI scores with recidivism rates might indicate whether the CPAI could be used with confidence in predicting recidivism. As previously noted, the results of Latessa et al. study (1999), even though conducted on a very small sample of programs, are very promising. The results of the present study, conducted by examining the treatment effectiveness literature and using effect size as an outcome measure, support these findings. Overall, the present data suggest that the programs including more CPAI items will have a greater impact on reducing recidivism within the treatment group relative to control group, therefore the predictive validity estimates of the CPAI support the use of the CPAI in the design and evaluation of correctional treatment programs.

Conclusions and Implications for Future Research

Not so long ago, the treatment effectiveness literature focused on exploring whether correctional programs work, without trying to understand what makes these programs effective. In the recent period, research of the “black box” of correctional programs led to the identification of a number of individual treatment elements linked to the reduction of recidivism and treatment effectiveness.

Following tradition in explaining what works instead of whether the programs work, Gendreau and Andrews (1996) developed the CPAI by summarizing the treatment elements empirically linked to treatment effectiveness. It makes logical sense that this instrument based on recent findings about what makes a program effective will be a good predictor of recidivism. However, the characteristics of this instrument still need to be explored. This dissertation attempted to answer how close the CPAI has come to being a solid, psychometrically sound instrument that could be used with confidence in the design and the evaluation of correctional treatment programs.

In the first part of the dissertation the interrater reliability was assessed by comparing two independent ratings of twenty federal and provincially delivered program using the 6th edition of the CPAI (Gendreau & Andrews, 1996). Scoring based on the interviews conducted with program directors were compared to the scoring of questionnaire based CPAI. Results indicated that the CPAI can be scored objectively by both methods. High interrater reliability coefficients indicate that we should have confidence in scoring the CPAI, as the raters agreed in over 90% (questionnaire based CPAI) and 94% (interview based CPAI) of the cases whether the treatment elements are present or absent. Regardless of the level of experience in scoring the CPAI or the scoring method used, the CPAI scoring manual has an objective

scoring criteria that leads to reliable results. Internal consistency estimates indicated that the CPAI is a homogenous scale.

The second part of the research addressed content and face validity of the CPAI. These psychometric properties were explored by surveying psychologists and treatment staff working in corrections (content validity) and first year psychology students (face validity) to determine whether the treatment elements included in the CPAI were considered important for the treatment effectiveness. The results revealed a high level of agreement within and across the samples indicating a presence of satisfactory face and content validity characteristics. Additionally, the literature review of treatment effectiveness conducted in the introductory part indicated that the CPAI included the majority of the items found to be linked with treatment effectiveness, demonstrating that the CPAI items have strong empirical support.

Finally, this dissertation provided preliminary estimates of predictive validity of the CPAI. The magnitude of the relationship between CPAI scores and reduction in recidivism expressed as effect size is impressive. These results indicated that the higher number of CPAI items was linked to the treatment effectiveness criterion (i.e., effect size) and that this correlation was statistically significant across different section, treatment and total CPAI scores. Two of the CPAI sections Preservice Assessment and Program Characteristics were identified as sections with the highest contribution to the effect size. These results confirmed previous finding that identified treatment elements linked to risk, need, responsivity and type of intervention employed to the treatment outcome, emphasizing that these elements should be addressed in a program design in order for a program to be effective in reducing recidivism rates. The CPAI treatment scores were a major source of variability in effect sizes and

Treatment level was incremental to the New Appropriate Treatment, a variable which previous meta-analysis (i.e., Dowden, 1998) have identified as having the strongest link to the treatment outcome. The positive effects of CPAI treatment level were evident across a variety of statistical, methodological, case characteristics, and setting organizational and treatment control conditions.

This dissertation has provided correctional literature with encouraging indicators of the psychometric properties of CPAI. Even though the CPAI shows promise as a reliable and valid instrument in the evaluation of correctional treatment programs, the data on psychometric characteristics are still in a preliminary phase. Notwithstanding the publication of the CPAI-2000, the majority of the items from CPAI (1996; 6th edition) are included in new version. Furthermore, these items represent treatment sections found to be strongly correlated with treatment outcome. The new items included in the CPAI 2000 should be included in future studies. The empirical evidence presented in this dissertation suggest strongly that the new version will likely have satisfactory psychometric characteristics.

A few suggestions emerged from the present study for future research on treatment effectiveness. First, the articles on treatment effectiveness should provide detailed description of the treatment interventions employed in order to advance theoretical and empirical knowledge of “what works” in reducing recidivism. The difficulties in determining the presence of treatment elements based on the information provided in the articles led to obtaining estimates rather than direct indicators of predictive validity. Until the studies on treatment effectiveness start including detailed descriptions of the interventions used, the questions of treatment effectiveness will not be explained adequately and with confidence.

Second, program designers should be encouraged to incorporate the CPAI elements in the design phase of the program. It is not sufficient to take only elements closely linked to treatment outcome (e.g., risk, need, responsivity) but also background/organizational variables. Despite the findings that indicated that the most significant contributors to the treatment outcomes were the treatment CPAI sections, previous metanalytic results have suggested that staff and program implementation elements are also important. The significant relationship of these sections with the effect size obtained in this dissertation supported their importance. Despite the finding that they do not contribute to the variance in treatment outcome beyond the treatment CPAI sections they are likely to enhance program effectiveness. This dissertation also pointed out that these sections were underreported in comparison to other CPAI sections and were persistently neglected across different studies.

Third, there is a need to have an evaluation of the programs as an ongoing process. Therefore, the purpose of the CPAI is not only to be included in the design of the program or at the final evaluation stage but also to be used to monitor program quality. The problematic/missing elements could be identified during the program delivery and consequently modified. There is a strong need to have an evaluation component incorporated in the program and have continuous gathering of recidivism data. Strict evaluation design and studies that would investigate the link between program elements and outcome will lead to the detailed knowledge about the treatment elements associated with the recidivism reduction. Such studies will provide empirical evidence on predictive validity of the CPAI and consequently to the delivery of correctional programs with impressive recidivism reduction. Studies such as the current one should indicate a need to have policy and funding changes that would open a new era in which we will not attempt to “prove” that treatment works but rather strive towards

finding even more effective elements that would lead to greater recidivism reduction. The future certainly belongs to the evidence based human service delivery that persistently leads to improved treatment effectiveness with correctional clients.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Adam, R., & Vetter, H.J. (1982). Social structure and psychodrama outcome: A ten-year follow-up. Journal of Offender Counseling, Services, and Rehabilitation, 6, 111-119.

*Adamson, L., & Dunham, H.W. (1956). Clinical treatment of male delinquents: A case study in effort and result. American Sociological Review, 21, 312-320.

*Alexander, J.F., Barton C., Schiavo, R.S., & Parsons, B.V. (1976). Systems-behavioral intervention with delinquents: Therapist characteristics, family behavior, and outcome. Journal of Consulting and Clinical Psychology, 44, 556-664.

*Anderson, D.B. (1981) The relationship between correctional education and parole success. Journal of Offender Counseling, Services and Rehabilitation, 5, 13-25.

*Andrews, D.A. (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.A. (1995a). Assessing program elements for risk reduction: The Correctional Program Assessment Inventory: Paper presented at the Research to Results Conference of the International Community Corrections Association, October, Ottawa, Ontario.

Andrews, D.A. (1995b). The psychology of criminal conduct and effective treatment. In J. McGuire (Ed.) What works: Reducing reoffending-Guidelines from Research and Practice (pp. 35-62). Chichester, England: John Wiley & Sons Ltd.

Andrews, D.A. (1996). Behavioral, cognitive behavioral and social learning contributions to criminological theory. Paper presented at the American Society of Criminology annual meeting Chicago, Illinois, November, 1996.

Andrews, D.A. (2002). Principles of effective correctional practice. In L. Motiuk & R.Serin (Eds.), Compendium 2000 on Effective Correctional Programming. (pp.9-17) Correctional Service of Canada.

Andrews, D.A., & Bonta, J. (1994). The psychology of criminal conduct. Cincinnati, OH: Anderson Publishing Company.

Andrews, D.A., & Bonta, J. (1998). Psychology of Criminal Conduct (2nd ed). Cincinnati OH: Anderson Publishing Company.

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

Andrews, D.A., & Carvell, C. (1998). Core correctional treatment - Core correctional supervision and counseling: Theory, research, assessment and practice. Ottawa, Carleton University.

Andrews, D.A., & Dowden, D.A. (1999). Managing correctional treatment for reduced recidivism: A meta-analytic review of program integrity. Unpublished manuscript. Carleton University, Ottawa.

Andrews, D.A., Dowden, C., & Gendreau, P. (1999). Clinically relevant and psychologically informed approaches to reduced re-offending: A meta-analytic study of human service, risk, need, responsivity and other concerns in justice contexts. Unpublished manuscript. Carleton University/University of New Brunswick.

Andrews, D.A., & Friesen, W. (1987). Assessments of antirriminal plans and the prediction of criminal future: A research note. Criminal Justice and Behavior, 14, 33-37.

Andrews, D.A., & Kiessling, J.J. (1980). Program structure and effective correctional practices: A summary for CaVic research. In R.R. Ross & P. Gendreau (Eds.), Effective correctional treatment (pp.439-463). Toronto: Butterworth.

*Andrews, D.A., Kiessling, J.J., Robinson, D., & Mickus, S. (1986). The risk principle of case classification: An outcome evaluation with young adult probationers. Canadian Journal of Criminology, 28, 377-396.

Andrews, D.A., Zinger, I., Hoge, R.D., Bonta, J., Gendreau, P., & Cullen, F.T. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta analysis. Criminology, 28, 369-404.

Annis, H.M. (1986). A relapse prevention model for treatment of alcoholics. In W.R. Miller & N. Heather (Eds.), Treating addictive behaviors (pp. 407-421). New York: Plenum.

*Annis, H.N., & Chan, D. (1983). The differential treatment model: Empirical evidence from a personality typology of adult offenders. Criminal Justice and Behavior, 10, 159-173.

Antonowicz, D.H., & Ross, R.R. (1994). Essential components of successful rehabilitation programs for offenders. International Journal of Offender Therapy and Comparative Criminology, 38(2), 97-104.

*Bachara, G.H. & Zaba, J.N. (1978). Learning disabilities and juvenile delinquency. Journal of Learning Disabilities 11, 242-246.

Baily, W.C. (1966). Correctional outcome: An evaluation of 100 reports. Journal of Criminal Law: Criminology and Police Science, 57, 153-160.

*Baird, S.C., Heinz, R.H., & Bemus, B.J. (1979) Project Report 14: A two year follow up. Wisconsin: Department of Health and Social Services, Case Classification/Staff Development Project, Bureau of Community Corrections.

*Bank, L., Marlowe, J.H., Reid, J.B., Patterson, G.R., & Weinrott, M.R. (1991). A comparative evaluation of parent-training interventions for families of chronic delinquents. Journal of Abnormal Child Psychology, 19, 15-33.

*Barbash, J.T. (1963). A study of psychological therapy and post-release adjustment. American Journal of Corrections, Jan-Feb, 26-29.

*Barkwell, L.J. (1976). Differential treatment of juveniles on probation. An evaluative study. Canadian Journal of Criminology and Corrections, 18, 363-378.

*Barton, C., Alexander, J.F., Waldron, H., Turner, C.W., & Warburton, J. (1985). Generalizing treatment effects of functional family therapy: Three replications. The American Journal of Family Therapy 13, 16-26.

*Barton, C., & Butts, J.A. (1990). Viable options: Intensive supervision programs for juvenile delinquents. Crime and Delinquency, 36, 238-250.

*Baron, R., Feeney, F., & Thornton, W. (1973). Preventing delinquency through diversion: The Sacramento County 601 Diversion Project. Federal Probation, 37, 13-18.

Baxter, D.J., Marion, A., & Gogen, B. (1995). Predicting treatment response in correctional settings. Forum on Corrections Research, 7, 38-41.

*Beck, J.L. (1981) Employment, community treatment placement and recidivism: A study of released federal offenders. Federal Probation, 45, 3-8.

Bender, D., Bliesener, T., & Losel, F. (1996). Deviance or resilience? A longitudinal study of adolescents in residential care. In G. Davies, S. Lloyd-Bostock, M. McMurren, & C.

Wilson (Eds.). Psychology, Law, and Criminal Justice (pp. 409-423) Berlin: Walter de Gruyter.

*Berman, J.J. (1978). An experiment in parole supervision. Evaluation Quarterly, 2, 71-90.

Berman, J.S., & Norton, N.C. (1985). Does professional training makes a therapist more effective. Psychological Bulletin, 98, 401-407.

*Birkenmayer, A.C. & Polonski, M. (1976). The community adjustment of male training school recidivists: II. The D.A.R.E. experience. Toronto: Ministry of Correctional Services.

*Bleick, C.R- & Abrams, A.I. (1987). The Transcendental Meditation Program and criminal recidivism in California. Journal of Criminal Justice, 15, 211-230.

*Bogert, A- J. & French, A.P. (I 978). Successful short-term family therapy with incarcerated adolescents. Journal of Juvenile and Family Courts, Feb, 3-8.

Bonta, J. (1995). The responsivity principle and offender rehabilitation. Forum on Corrections Research, 7, 34-37.

Bonta, J.L., Boyle, J., Motiuk, L.L., & Sonnichsen, P. (1985). Restitution in correctional half-way houses: Victim satisfaction, attitudes, and recidivism. Canadian Journal of Criminology, 25, 386-397.

Bonta, J., & Gendreau, P. (1992). Coping with prison. In P.Nuedfeld & P.E. Tetlock (Eds.), Psychology and social policy (pp.343-354). New York: Hemisphere.

Bonta, J., Law, M., & Hanson, R.K. (1998). The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis. Psychological Bulletin, 123, 123-142.

*Borduin, C.M., Mann, B.J., Cone, L.T., Henggeler, S.W., Fucci, B.R., Blaske, D.M., & Williams, R. A. (1995). Multisystemic treatment of serious juvenile offenders. Long-term prevention of criminality and violence. Journal of Consulting and Clinical Psychology, 63, 569-578.

*Boudouris, J. & Turnbull, B.W. (1985). Shock probation in Iowa. Journal of Offender Counseling Services and Rehabilitation, 9, 53-67.

*Brier, N. (1994). Treatment targeted for adjudicated youth with learning disabilities: Effects on recidivism. Journal of Learning Disabilities, 27, 215-222.

Brill, R., & Reitsma, M. (1980). Comparison of the effects of matching treaters to delinquents clients using conceptual level and I-level theories. Differential View, 9, 37-47.

Brown, J.M., & Miller, W.R. (1993). Impact of motivational interviewing on participation and outcome in residential alcoholic treatment. Psychology of Addictive Behaviors, 7(4), 211- 218.

*Bry, B.H. (1982). Reducing the incidence of adolescent problems through preventative intervention: One- and five-year follow-up. American Journal of Community Psychology, 10, 265-276.

*Buckner, J.C., & Chesney-Lind, M. (1983). Dramatic cures for juvenile crime: An evaluation of a prisoner run delinquency prevention program. Criminal Justice and Behavior, 10, 227-247.

Buehler, R.E., Patterson, G.R., & Furniss, J.M. (1966). The reinforcement of behavior in institutional settings. Behavioral Research and Therapy, 4, 157-167.

Bush, J. (1995). Teaching self-risk management to violent offenders. In J. McGuire (Ed.), What works: Effective methods to reduce re-offending (pp.139-154). Chichester, England: John Wiley & Sons Ltd.

*Byles, J.A- (1981). Evaluation of an attendance center program for male juvenile probationers. Canadian Journal of Criminology, 23, 343-355.

*Byles, J.A. & Maurice, A. (1979). The Juvenile Services Project: An experiment in delinquency control. Canadian Journal of Criminology, 21, 155-165.

*Castellano, T.C. & Soderstrom, I.R. (1992). Therapeutic wilderness programs and juvenile recidivism: A program evaluation. Journal of Offender Rehabilitation, 17, 19-46.

*Clarke, R.V.G. & Cornish, D.B. (1978). The effectiveness of residential treatment for delinquents. In L.A. Hersov, M. Berger, and D. Shaffer (eds.), Aggression and Antisocial Behavior in Childhood and Adolescence. Oxford: Pergamon.

Cleland, C.M. (1997). A meta-analytic approach to the link between needs-targeted treatment and reductions in criminal offending. Unpublished doctoral dissertation.

*Cochran, A.W. (1965). Is education of value to the parolee? Journal of Correctional Education, 22-25.

*Collingwood, T.R., Douds, A., & Williams, H. Juvenile diversion: The Dallas police department youth service program, Federal Probation, 23-27

Cohler, B.J. (1987). Adversity, resilience, and the study of lives. In E.J. Anthony & B.J. Cohler (Eds.), The invulnerable child (pp. 363-404).

*Craft, M., Stephenson, G. & Granger, C. (1966). A controlled trial of authoritarian and self governing regimes with adolescent psychopaths. American Journal of Orthopsychiatry, 34, 543-554.

Cullen, F.T., & Gendreau, P. (1989). The effectiveness of correctional rehabilitation In L. Goodstein and D.L. MacKenzie (Eds.), The American Prison: Issues in Research Policy (pp. 23-24). New York: Plenum.

*Davidson, W. S. & Johnson, C.D. (draft). Family intervention with juvenile offenders: An experimental comparison. Unpublished manuscript. Michigan State University.

*Davidson, W.S., Redner, R., Blakely, C., Mitchell, C., & Emshoff, J. (1987). Diversion of juvenile offenders: An experimental comparison. Journal of Consulting and Clinical Psychology, 55, 68-75.

*Davidson W. S., & Robinson, M. (1975). Community psychology and behavior modification: A community-based program for the prevention of delinquency. Corrective and Social Psychology, 21, 1-12.

*Davidson, W. S. & Wolfed, T.R- (1977). Evaluation of a community-based behavior modification program for prevention of delinquency. Community Mental Health Journal, 13, 296-306.

Davies, G.E., & Leitenberg, H. (1987). Adolescent sex offenders. Psychological Bulletin, 101, 417-427.

*Deng, X. (1997). The deterrent effects of initial sanction on first-time apprehended shoplifters. National Journal of Offender Therapy and Comparative Criminology, 41, 284-297.

Dobson, K.S., & Shaw, B.F. (1988). The use of treatment manuals in cognitive therapy: Experience and issues. Journal of Consulting and Clinical Psychology, 56, 673-680.

Dowden, C. (1998). A meta-analytic examination of the risk, need and responsivity principles and their importance within the rehabilitation debate. Unpublished master's thesis. Carleton University, Ottawa.

*Druckman, J.M. (1979). A family-oriented policy and treatment program for female juvenile status offenders. Journal of Marriage and the Family, 41, 627-636.

*Duguid, S. (1981) Prison education and criminal choice: The context of decision-making. Canadian Journal of Criminology, 23, 421-438.

*Dunford, F.W., Huizinga, D., & Elliot, D. S. (1990). The role of arrest in domestic assault: The Omaha Police Experiment. Criminology, 28, 183-206.

Durlak, J. (1981). Evaluating comparative studies of paraprofessional and professional helpers. Psychological Bulletin, 81, 998-1008.

*Dutton, D.G. (1986). The outcome of court-mandated treatment for wife assault: A quasi-experimental evaluation. Violence and Victims, 1, 163-175.

Eisberg, M., & Fabelo, T. (1996). Evaluation of the Texas correctional substance abuse treatment initiative: the impact of policy research. Crime and Delinquency, 42(2), 296-309.

*Elder, E., Skok, R-L. & McLaughlin, T.F. (1990). Inside Education. Corrective and Social Psychiatry and Journal of Behavior Technology Methods and Therapy, 36, 70-74.

*Elrod, H.P. & Minor, K.I. (1992). Second wave evaluation of a multi-faceted intervention for juvenile court probationers. International Journal of Offender Therapy and Comparative Criminology, 36, 247-262.

*Emshoff, J.G., & Blakely, C.G. (1983). The diversion of delinquent Youth: Family focused intervention. Children and Youth Service Review, 5, 343-356

*Fairhead, S. (1981). Persistent petty offenders. London: Her Majesty's Stationary Office.

Fagan, J. (1990). Treatment and reintegration of violent juvenile offenders: experimental results. Justice Quarterly, 7(2), 233-263.

Farrington, D.P., & West, D.J. (1993). Criminal/personal/life histories of chronic offenders. Risk and protective factors and early identification. Criminal Behavior and Mental Health, 3(4), 492-523.

*Field, G. (1989). The Comer-stone Program: A client outcome study. Federal Probation, 51, 50-55.

Garet, C.J. (1985). Effects of residential treatment of adjudicated delinquents: A meta-analysis. Journal of Research in Crime and Delinquency, 22, 287-308.

Garmezy, N., Masten, A.S., & Tellefen, A. (1984). The study of stress and competence in children: block for developmental psychopathology. Child Development, 55, 97-111.

Gendreau, P. (1996). Offender rehabilitation: What we know and what needs to be done. Criminal Justice and Behavior, 23(1), 144-161.

Gendreau, P., & Andrews, D.A. (1990). The Correctional Program Evaluation Inventory. University of New Brunswick, Saint John, New Brunswick/ Carleton University, Ottawa, Ontario.

Gendreau, P., & Andrews, D.A. (1996). The Correctional Program Assessment Inventory (6th edition). University of New Brunswick, Saint John, New Brunswick/ Carleton University, Ottawa, Ontario.

Gendreau, P., Cullen, F., & Bonta, J. (1994). Intensive rehabilitation supervision: The next generation in community corrections. Federal Probation, 58, 72-78.

Gendreau, P., Goggin, C., & Annis, H. (1990). Survey of existing substance abuse programs. Forum on Corrections Research, 2, 608.

Gendreau, P., Goggin, C., Cullen, F., & Andrews, D. (2002). The effects of community sanctions and incarceration on recidivism. In L. Motiuk and R. Serin (Eds) Compendium 2000 on Effective Correctional Programming (pp. 18-23). Correctional Service of Canada.

Gendreau, P., Goggin, C., & Smith, P. (2002). Evaluating knowledge: How Meta-Analysis can serve needs of Correctional clinicians and policy makers. In L. Motiuk and R. Serin (Eds) Compendium 2000 on Effective Correctional Programming (pp. 202-207). Correctional Service of Canada.

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

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

Gendreau, P., & Ross, R. (1987). Revivification of rehabilitation: evidence from 1980s. Justice Quarterly, 4, 349-408.

Genevie, L., Margolies, E., & Muhlin, G.L. (1986). How effective is correctional intervention? Social Policy, 17, 52-57.

Gensheimer, L.H., Mayer, J.P., Gottschalk, R., & Davidson, W.S. (1986). Diverting youth from the juvenile justice: A meta-analysis of intervention efficacy. In Apter, S.J., &

Goldstein, A.P. (Eds.), Youth violence: Programs and Prospects (pp. 39-58). New York: Pergamon Press.

*Gilbert, G.R- (1977). Alternate routes: A diversion project in the juvenile justice system. Evaluation Quarterly, 1, 301-318.

Gillis, C., Getkate, M., Robinson, D., & Porporino, F. (1995). Correctional work supervisor leadership and credibility: Their influence on offender work motivation. Forum on Corrections Research, 7, 15-17.

Gillis, C., & Grant, B.A. (1996). The relation between offender motivation, day parole outcome and postrelease recidivism. Research Report. Ottawa: Correctional Service of Canada.

Glass, G.V., McGraw, B., & Smith, M.L. (1981). Meta-analysis of social research. Beverly Hills: Sage.

Gordon, D.A., Andrews, D.A., Hill, J., & Kurkoswski, K. (1992). Therapeutic integrity and the effectiveness of family therapy: A meta-analysis. Unpublished manuscript.

Gordon, D.A., & Arbuthnot, J. (1988). The use of paraprofessionals to deliver home based family therapy to juvenile delinquents. Criminal Justice and Behavior, 15(3), 364-378.

*Gordon, D.A., & Edwards, J. (draft). Evaluation of a state training program in family therapy for juvenile probation officers. Unpublished manuscript, Ohio State University.

*Gottfredson, G.D. (1987). Peer group interventions to reduce the risk of delinquent behavior: A selective review and a new evaluation. Criminology, 25, 671-714.

Gottschalk, R., Davidson, W.S., Gensheimer, L.K., & Mayer, J.P. (1987a). Behavioral approaches with juvenile offenders. A meta-analysis of long-term treatment efficacy. In E.K.

Morris & C.J. Braukmann (Eds.), Behavioral Approaches to Crime and Delinquency (pp. 399-423). New York: Plenum Press.

Gottschalk, R., Davidson, W.S., Gensheimer, L.K., & Mayer, J.P. (1987b). Community based intervention. In H.C. Quay (Ed.). Handbook of juvenile delinquency (pp. 266-289). New York, John Willey & Sons.

*Grant, J.D. (1965). Delinquency treatment in an institutional setting. In H.C. Quay (ed.), Juvenile Delinquency: Research and Theory. Princeton, N.J.: Van Nostrand.

*Grant, J.D., & Grant, M.Q. (1959). A group dynamics approach to the treatment of nonconformists in the navy. Annals of the American Academy of Political and Social Science, 322, 126-135.

*Greenwood, P.W., & Turner, S. (1993). Evaluation of the Paint Creek Youth Center: A serious program for serious delinquents. Criminology, 31, 263-279.

Grey, G. (1997). Coercion and Treatment Effectiveness. Unpublished Master's Thesis. University of New Brunswick, Saint John.

* Gruher, M. (1979). Family counseling and the status offender. Juvenile and Family Court Journal, 20, 23-27.

*Hackler, I.C., & Hagan, J.L. (1975). Work and teaching machines as delinquency prevention tools: A four-year follow-up. Social Service Review, 49, 92-106.

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

*Hanson, R.K., Steffy, R.A., & Gauthier, R. (1993). Long-term recidivism of child molesters. Journal of Consulting and Clinical Psychology, 61, 646-652.

Hare, R.D. (1996). Psychopathy: A clinical construct whose time has come. Criminal Justice and Behavior, 23, 25-54.

Harris, P.W. (1988). The interpersonal maturity level classification system: I-Level. Criminal Justice and Behavior, 15(1), 58-77.

Heather, N., Rollnic, S., Bell, A., & Richmond, R. (1996). Effects of brief counselling among male heavy drinkers identified in several hospital wards. Drug and Alcohol Review, 15(1), 29-38.

Henggeler, S.W., Melton, G.B., Brondino, M.J., Scherer, D.G., & Hanley, J.H. (1997). Multisystemic therapy: An effective approach to incarcerating serious juvenile offenders and their families: The role of treatment fidelity in successful dissemination. Journal of Consulting and Clinical Psychology, 65, 821-833.

Henggeler, S.W., Melton, G.B., & Smith, L.A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. Journal of Consulting and Clinical Psychology, 60(6), 953-961.

Henning, K.R., & Frueh, C.B. (1996). Cognitive-behavioral treatment of incarcerated offenders: An evaluation of the Vermont Department of Correction's Cognitive Self-Change Program. Criminal Justice and Behavior, 23(4), 523-541.

Hill, J.K., Andrews, D.A., & Hoge, R.D. (1991). Meta-analysis of treatment programs for young offenders: The effect of clinically relevant treatment on recidivism, with controls introduced for various methodological variables. Canadian Journal of Program Evaluation, 6, 97- 109.

Hodge, J.E. (1991). Addiction to crime. Issues in Criminological and Legal Psychology, 2(17), 92-96.

Hoge, R.D., Andrews, D.A., & Leschied, A.W. (1996). An investigation of risk and protective factors in a sample of youthful offenders. Journal of Child Psychology and Psychiatry and Applied Disciplines, 37(4), 419-424.

Hoge, R.D., Leschied, A., & Andrews, D.A. (1993). An investigation of young offender services in the province of Ontario. Toronto, Ministry of Community and Social Services.

Hollin, C.R. (1993). Advances in the psychological treatment of criminal behavior. Criminal Behavior and Mental Health, 3(3), 142-153.

Hollin, C.R. (1995). The meaning and implications of "Program Integrity". In J. McGuire (Ed.), What works: Reducing reoffending-Guidelines from Research and Practice (pp. 195-208). Chichester, England: John Wiley & Sons Ltd.

Holsinger, A.M. (1999). Evaluation of Community Correctional Facilities. Unpublished Doctoral Thesis. University of Cincinnati, Cincinnati, OH.

*Howitt, P. S., & Moore, E.A. (1991). The efficacy of intensive early intervention: An evaluation of the Oakland County Probate Court Early Offender Program. Juvenile and Family Court Journal, 42, 25-36.

*Hughes, O.V. (1993). Anger management program outcomes. Forum on Corrections Research, 3, 5-9.

Izzo, R., & Ross, R. (1990). A meta-analysis of rehabilitation programs for juvenile delinquents: A brief report. Criminal Justice and Behavior, 17, 134-142.

*Jackson, P.C. (1983). Some effects of parole supervision on recidivism. British Journal of Criminology, 22, 17-34.

*Jeffrey, R., & Woolpert, S. (1974). Work furlough as an alternative to incarceration: An assessment of its effects on recidivism and social cost. The Journal of Criminal Law and Criminology, 65, 405-415.

Jesness, C.F. (1971-72). Comparative effectiveness of two institutional treatment programs for delinquents. Child Care Quarterly, 1(2), 119-130.

*Jesness, C.F. (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.F. (1972). Comparative effectiveness of two institutional treatment programs for delinquents. Child Care Quarterly, 1, 19-130.

*Jesness, C.F. (1975). Comparative effectiveness of behavior modification and transactional analysis programs for delinquents. Journal of Consulting and Clinical Psychology, 43, 758-779.

Johnson, S.V. (1981). Staff drift: A problem in program integrity. Criminal Justice and Behavior, 8(2), 223-232.

*Johnson, T.F. (1977). The results of family therapy with juvenile offenders. Juvenile Justice, 28, 29-33.

*Johnson, B.D., & Goldberg, R.T. (1983). Vocational and social rehabilitation of delinquents. Journal of Offender Counseling Services, and Rehabilitation, 5, 43-60.

*Jolin, A. & Stipak, B. (1992). Drug treatment and electronically monitored home confinement: An evaluation of a community-based sentencing option. Crime and Delinquency, 38, 158-170.

*Kahn, T.J., & Chambers, I.J. (1991). Assessing reoffense risk with juvenile sexual offenders. Child Welfare, 10, 333-345.

Kaplan, M.R., & Saccuzzo, D.P. (1993). Psychological testing: Principles, applications, and issues. Pacific Grove, California: Brooks/Cole Publishing Company.

Kassebaum, G., Ward, D., & Wilner, D. (1971). Prison treatment and parole survival: an empirical assessment. New York: John Wiley.

Kazdin, A.E. (1986). Comparative outcome studies of psychotherapy: Methodological issues and strategies. Journal of Consulting and Clinical Psychology, 54, 95-105.

*Kelley, T.M., Havva, A.K., & Blak, R.A. (1979). The effectiveness of college student companion therapists with predelinquent youths. Journal of Police Science and Administration, 7, 186-195.

*Kelly, F.J., & Baer, D.L. (1971). Physical challenge as a treatment for delinquency. Crime and Delinquency, 17, 437-445.

Kennedy, S., & Serin, R. (1995). Treatment responsivity: Contributing to effective correctional programming. Forum on Corrections Research, 7, 46-52.

*Kernodle Jr., I.R., Joyce, C.C., & Farmer, R.J. (1997). Changing the behavior of DWI offenders. Journal of Offender Rehabilitation, 22, 113-128.

*Kingsnorth, R.F. (1991). The Gunther Special: Deterrence and the DLTI offender. Criminal Justice and Behavior, 18, 251-266.

Kirby, B.C. (1954). Measuring effects of treatment of criminals and delinquents. Sociology and Social Research, 38, 368-374.

*Kirigan, K.A., Braukman, J.D., Atwater, J.D., & Montrose, M.W. (1982). An evaluation of Teaching Family (Achievement Place) Group Homes for juvenile offenders. Journal of Behavior Analysis, 15, 1- 16.

*Klein, M.W. (1986). Labeling and theory and delinquency policy: An experimental test. Criminal Justice and Behavior, 13, 47-79.

*Klein, N.C., Alexander, F., & Parsons, B.V. (1977). Impact of family system intervention on recidivism and sibling delinquency: A model of primary prevention and program evaluation. Journal of Consulting and Clinical Psychology, 3, 469-474.

*Kratcoski, P.C., & Kratcoski, L.D. (1982). The Phoenix Program: An educational alternative for delinquent youths. Juvenile and Family Court Journal, 33, 17-23.

*Lab, S., Shields, S., & Schondel C. (1993). Research Note: An evaluation of juvenile- sexual offender treatment. Crime and Delinquency, 39, 543-553.

*Langworthy, R., & Latessa, E.J. (1993). Treatment of chronic drunk drivers: The Turning Point Project. Journal of Criminal Justice, 21, 265-276.

Larsen, J., Hudson, S.M., & Ward, T. (1995). Evaluation of attributional change in a relapse prevention program for child molesters. Behavioral Change, 12(3), 127-138.

Latessa, E.J., Holsinger, A.M., Jones, D., Fulton, B., Johnson, S., & Kadleck, C. (1999). Evaluation of the Ohio Department of Youth Services' Community Correctional Facilities (Report). Ohio: University of Cincinnati.

Latessa, E.J., Holsinger, A.M., Jones, D., Fulton, B., Stichman, A., & Moon, M.M. (1999). An evaluation of selected juvenile justice programs in Ohio using the Correctional Program Assessment Inventory (Report). Ohio: University of Cincinnati.

Lederman, M. V. (1986). Toward the assessment and evaluation of consultation services: A consultation process and product inventory. Unpublished master's thesis, Carleton University, Ottawa, Ontario, Canada.

Lee, R., & Olejnik, S. (1981). Professional outreach counseling can help the juvenile probationers: a two year follow-up study. The Personal and Guidance Journal, 60, 445-449.

*Leeman, L.W., Gibbs, J.C., & Fuller, D. (1993). Evaluation of a multi-component group treatment program for juvenile delinquents. International Journal of Offender Therapy- and Comparative Criminology, 27, 281-291.

Leschied, A.W. (2002). Implementation of effective correctional programs. In L. Motiuk and R. Serin (Eds) Compendium 2000 on Effective Correctional Programming (pp. 41-46). Correctional Service of Canada.

Leschied, A.D.W., Jaffe, P.G., & Stone, G.L. (1985). Differential response of juvenile offenders to two detention environments as a function of conceptual level. Canadian Journal of Criminology, 27(4), 467-476.

*Lewis, R-V. (1983). Seared straight-California style: Evaluation of the San Quentin Program. Criminal Justice and Behavior, 10, 209-226.

*Linden, R., Perry, L., Ayers, D., & Parlett, T.A. (1984). An evaluation of a prison education program. Canadian Journal of Criminology, 26, 65-73.

*Lindforss, L., & Magnusson, D. (1997). Solution-focused therapy in prison. Contemporary Family Therapy, 19(1), 89-103.

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, Nev.

Lipsey, M. (1992). Juvenile delinquency treatment: A meta-analytic inquiry into the variability of effects. In T.D. Cook, H. Cooper, D.S., Cordray, H. Hartman, L.V. Hedges, R.J.

Light, T.A. Louis, & F. Mosteller (Eds.), Meta-analysis for explanation: A casebook (pp. 83-127). New York: Russell Sage Foundation.

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- Guidelines from Research and Practice (pp. 63-78). Chichester, England: John Wiley & Sons Ltd.

*Lipsey, M.W., Cordray, D. S., & Berger, D.E. (1981). Evaluation of a juvenile diversion program using multiple lines evidence, Evaluation Review, 5, 283-306.

Lipton, D.S., Martinson, R., & Wilks, J. (1975). The effectiveness of correctional treatment: A survey of treatment evaluation studies. New York: Plenum.

Lipton, D.S., Pearson, F.S., Cleland, C., & Yee, D. (1997). Synthesizing correctional treatment outcome: Preliminary CDATE findings. Paper presented at the 5th Annual National Institute of Justice Conference on Research and Evaluation in Criminal Justice, July, Washington, DC.

*Little, G.L., Robinson, K.D., & Burnette, K.D. (1990). Treating drunk drivers with Moral Reconation Therapy: A two-year recidivism study. Psychological Reports, 66, 1379-1387.

Logan, C.H. (1972). Evaluation research in crime and delinquency: A reappraisal. Journal of Criminal Law, Criminology and Police Science, 63, 378-387.

Losel, F. (1989). Evaluation research on correctional treatment in West Germany: A Meta- analysis. In H. Wegener, H., F. Losel and J. Haisch (Eds.), Criminal Behavior and the Justice System: Psychological Perspectives. New York, Springer-Verlag

Losel, F. (1993). The effectiveness of treatment in institutional and community settings. Criminal Behavior & Mental Health, 3, 416-437.

Losel, F. (1995). The efficacy of correctional treatment: A review and synthesis of meta- evaluations. In J. McGuire (Ed.) What works: Reducing reoffending-Guidelines from Research and Practice (pp. 79-111). Chichester, England: John Wiley & Sons Ltd.

Luthar, S.S. (1993). Annotation: methodological and conceptual issues in research on childhood resilience. Journal of Child Psychology and Psychiatry, 34, 441-453.

*Maguire, K.E., Flanagan, T.J., & Thornberry, T.P. (1988). Prison labor and recidivism. Journal of Quantitative Criminology, 4, 3-18.

*Marques, J.K., Day, D.M., Nelson, C., & West M.A. (1994). Effects of cognitive-behavioral treatment on sex offender recidivism: Preliminary results of a longitudinal study. Criminal Justice and Behavior, 21, 28-54.

*Marshall, W.L., & Barbaree, H.E. (1988). The long-term evaluation of a cognitive-behavioral treatment program for child molesters. Behavior Research and Therapy, 26, 499-511.

Martin, S.E., Sechrest, L.B., & Redner, R.(1981). New directions in the rehabilitation of criminal offenders. National Academy Press, Washington, D.C.

Martinson, R. (1974). What works?-Questions and answers about prison reform. The Public Interest, 10, 22-54.

*Maskin, M.E. (1976). The differential impact of work vs. communication-oriented juvenile correction programs upon recidivism rates in delinquent males. Journal of Clinical Psychology, 32, 432-433.

Masten, A.S., Best, K.M., & Garmezy, N. (1990). Resilience and development: Contributions from the study of children who overcome adversity. Development and Psychopathology, 2, 435-444.

Mayer, J.P., Gensheimer, L.K., Davidson, W.S., Gottschalk, R. (1986). A meta-analysis of impact in the natural environment. In Apter, S.J., & Goldstein, A.P. (Eds.), Youth violence: Programs and Prospects (pp. 24-38). New York: Pergamon Press

*McCord, J., & McCord, W. (1959). A follow-up report on the Cambridge-Somerville Youth study. The Annals of the American Academy of Political and Social Science, 32, 89-98.

McMurran, M., & Hollin, C.R. (1993). Young offenders and alcohol related crime: A practitioner's guidebook. Chichester: John Wiley and Sons.

*McPherson, S.J., McDonald, L.E., & Ryder, C.W. (1983). Intensive counseling with families of juvenile offenders. Juvenile and Family Court Journal, 34, 27-33.

*Michaels, K.W., & Green, K.H. (1979). A child welfare agency project: Therapy for families of status offenders. Child Welfare, 51, 216-220.

Miller, W.R.(1992). Motivational interviewing: Research practices and puzzles. Addictive Behaviors, 21(6), 835-842.

*Minor, K.I., & Elrod, H.P. (1990). The effect of a multi-faceted intervention on the offense activities of juvenile probationers. Journal of Offender Counseling, Services and Rehabilitation, 15, 87-108.

*Mitchel, C.M. (1983). The dissemination of a social intervention: Process and effectiveness of two types of paraprofessional change agents. American Journal of Community Psychology, 11, 723-739.

Moncher, F.J., & Prinz, R.J. (1991). Treatment fidelity in outcome studies. Clinical Psychology Review, 11, 247-266.

*Moore, R.H. (1987). Effectiveness of citizen volunteers functioning as counselors for high-risk young male offenders. Psychological Reports, 61, 823-830.

Nevo, B. (1985). Face validity revisited. Journal of Educational Measurement, 22(4), 278- 293.

*Newland, G. (1987). Prevention of delinquency and coping with status offenders by parent training casework. Report presented to The Academy of Criminal Justice Sciences Annual Convention, St. Louis, Missouri, and March 1987.

* Odell, B.N. (1974). Accelerating entry into the opportunity structure: A sociologically- based treatment for delinquent youth. Sociology and Social Research 58, 312-317.

*O'Donnell, C.&, Lydgate, R- & Fo, W.S.O. (1979). The buddy system: Review and follow-up. Child Behavior Therapy, 1, 161-169.

*O'Neil, M. (1990). Correctional higher education: Reduced recidivism? Journal of Correctional Education, 41(1), 28-31.

*Ostrum, T.M., Steele, C.V., Resenblood, L.K., & Mirels, H.L. (1971). Modification of delinquent behavior. Journal of Applied Social Psychology, 1, 118-136.

Palmer, T. (1975). Martinson revisited. Journal of Research in Crime and Delinquency, 12, 133-152.

*Palmer, T., & Lewis R-V. A differentiated approach to juvenile diversion. Journal of Research and Delinquency, 17, 209-227.

*Pearson, F. S. (1988). Evaluation of New Jersey's Intensive Supervision Program. Crime and Delinquency, 34, 437-448.

*Persons, R. (1967). Relationship between psychotherapy with institutionalized boys and subsequent adjustments. Journal of Consulting Psychology, 31, 137-141.

Porporino, F.J., & Robinson, D. (1992). The correctional benefits of education: A follow- up of Canadian Federal Offenders. Journal of Correctional Education, 43(2), 92-98.

*Prendergast, M.L., Wellisch, J., & Wong, M.M. (1996). Residential treatment for women parolees following prison-based drug treatment: Treatment experiences, needs and services, outcomes. The Prison Journal, 76, 253-274.

Quay, H.C. (1977). The three faces of evaluation: What can be expected to work. Criminal Justice and Behavior, 4(4), 341-354.

Quay, H.C. (1987). Institutional treatment. In H.C. Quay (Ed.). Handbook of juvenile delinquency (pp. 244-265). New York: John Willey & Sons.

*Quay, H.C., & Love, C. T. (1977). The effect of a juvenile diversion program on rearrests. Criminal Justice and Behavior, 1, 377-396.

*Raynor, P., & Vanstone, M. (1996). Reasoning and rehabilitation in Britain: The results of the straight thinking on probation (STOP) program. International Journal of Offender therapy and Comparative Criminology, 40(4), 272-284.

*Redfering, D.L. (1973). Durability of effects of group counseling with institutionalized females. Journal of Abnormal Psychology, 82, 85-86.

*Regoli, R., Wilderman, E., & Pogrebin, M. (1985). Using an alternative evaluation measure for assessing juvenile diversion. Children and Youth Services Review, 1, 21-38.,

Reitsma-Street, M. (1984). Differential treatment of young offenders: A review of the conceptual level matching model. Canadian Journal of Criminology, 26(2), 199-212.

Reitsma-Street, M., & Leschied, A.W. (1988). The conceptual-level matching model in corrections. Criminal Justice and Behavior, 15(1), 92-108.

Reitsma-Street, M., & Street, H. (1981). Analysis of a program for low conceptual level delinquent youth. Differential View, 11, 19-31.

Rezmovic, E.L. (1984). Assessing treatment implementation amid the slings and arrows of reality. Evaluation Review, 2, 187-204.

Rice, M.E., Harris, G.T., & Cormier, C.A. (1992). An evaluation of a maximum security therapeutic community for psychopaths and other mentally disordered offenders. Law and Human Behavior, 16, 399-412.

*Rice, M.E., Harris, G.T., & Quinsey, V.L. (1991). Evaluating treatment programs for child molesters. Research Report, 8, 1-34.

*Robertson, G., & Gunn, J. (1987). A ten-year follow-up of men discharged from Grendon Prison. British Journal of Psychiatry, 151, 674-678.

*Robin, G.D. (1969). Anti-poverty programs and delinquency. The Journal of Criminal Law, Criminology and Police Science, 60, 323-33 1.

Rollnic, S., & Heather, W., & Bell, A. (1992). Negotiating behavior change in medical setting: The development of brief motivational interviewing. Journal of Mental Health, 1(1), 25- 37.

Rollnic, S., Heather, N., Gold, R., & Hall, W. (1992). Development of a short “readiness to change” questionnaire for use in brief opportunistic intervention among excessive drinkers. British Journal of Addictions, 87(5), 743-754.

*Romero, J.J., & Williams, L.M. (1983). Group psychotherapy and intensive probation supervision with sex offenders: A comparative study. Federal Probation, 47, 36-42.

*Romig, D.A. (1976). Length of institutionalization, treatment program completion, and recidivism among delinquent adolescent males. Criminal Justice Review, 1, 115-119.

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

*Rosenthal, J.A., & Glass G.V. (1986). Impacts of alternatives to out-of-home placement: A quasi-experimental study. Children and Youth Services Review, 8, 305-321.

Ross, R.R., & Fabiano, E. (1985). Time to think: A cognitive model of delinquency prevention and offender rehabilitation. Johnson City: Institute of Social Sciences and Arts.

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

*Ross, R.R., & McKay, B.H. A study of institutional treatment programs. International Journal of Offender Therapy and Comparative Criminology, 165-173.

*Royse, D., & Buck, S.A. (1991). Evaluating a diversion program for first-time shoplifters, Journal of Offender Rehabilitation, 17, 147-158.

Rutter, M. (1985). Resilience in the face of adversity. Protective factors and resistance to psychiatric disorder. British Journal of Psychiatry, 147, 598-611.

Rutter, M. (1987). Psychosocial resilience and protective mechanisms. American Journal of Orthopsychiatry, 57, 316-331.

Rutter, M. (1990). Psychological resilience and protective mechanisms. In J. Rolf, A.S. Masten, D. Cicchetti, K.H. Nuechterlein, & S. Weintraub (Eds.), Risk and protective

factors in the development of psychopathology (pp. 181-214). New York: Cambridge University Press.

*Saden, S.J. (1962). Correctional research. Journal of Correction Education, 13, 22-26.

*Sandhu, H.S., Dodder, R.A., & Mathur, M. (1993). House arrest: success and failure rates in residential and nonresidential community-based programs. Journal of Offender Rehabilitation, 19, 131-144.

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

Saylor, W.G., & Gaes, G.G. (1992). PREP study links UNICOR work experience with successful post-release outcome. Office of Research and Evaluation. Department of Justice. Federal Bureau of Prisons.

*Schneider, A.L., & Schneider, P.R. (1985). The impact of restitution on recidivism of juvenile offenders: An experiment in Clayton County, Georgia. Criminal Justice Review, 10, 1-10.

*Schnur, A.C. (1948). The educational treatment of prisoners and recidivism. American Journal of Sociology, 14, 142-147.

*Schumacher, R.E., Anderson, D.B., & Anderson, S.L. (1990). Vocational and academic indicators of parole success. Journal of Correctional Education, 41, 8-13.

*Scott Jr., R.F., Hawkins Jr., R-D., & Farnsworth, M. (1994). Operation Kick-It. Texas prisoners rehabilitate themselves by dissuading others. Journal of Offender Rehabilitation, 20, 207-215.

*Shenman, L.W. & Berk, R.A. (1984). The specific deterrent effects of arrest for domestic assault. American Sociological Review, 49, 261-272.

*Shore, M.F., & Massimo, J.L. (1979). Fifteen years after treatment: A follow-up study of comprehensive vocation-oriented psychotherapy. American Journal of Orthopsychiatry, 49, 240-245.

*Shorts, I.D. (1986). Delinquency by association. British Journal of Criminology, 26, 156- 163.

*Siegal, H.A. (1985). The intervention approach to drunk driver rehabilitation- Part II. Evaluation. The International Journal of the Addictions, 20, 675-689.

*Siegal, H.A., & Cole, P.A. (1993). Enhancing criminal justice based treatment through the application of the intervention approach. Journal of Drug Issues, 23, 131-141.

Simourd, L., & Andrews, D.A. (1994). Correlates of delinquency: A look at gender differences. Forum on Corrections Research, 6, 26-31.

*Sloane, H.N., & Ralph, J.L. (1973). A behavior modification program in Nevada. International Journal of Offender Therapy and Comparative Criminology, 17, 290-296.

*Sontheimer, H., & Goodstein, L. (1993). An evaluation of juvenile intensive aftercare probation: Aftercare versus system response effects. Justice Quarterly, 10, 1993.

*Sorenson, J.L. (1978). Outcome evaluation of a referral system for juvenile offenders. American Journal of Community Psychology, 6, 381-388.

*Sowles, R.C., & Gill, J. (1970). Institutional and community adjustment of delinquents following counseling. Journal of Consulting and Clinical Psychology, 34, 398-402.

*Spergel, I.A., Reamer, F.G., & Lynch J.P. (1981). Deinstitutionalization of status offenders: Individual outcomes and system effects. Journal of Research in Crime and Delinquency, 18, 4-33.

Stewart, L., & Millson, W.A. (1995). Offender motivation for treatment as a responsivity factor. Forum on Corrections Research, 7, 5-7.

*Stewart, M.J., Vockel, E.L., & Ray, R.E. (1986). Decreasing court appearances of juvenile status offenders. Social Casework, 67, 74-79.

Stouthamer-Loeber, M., Loeber, R., Farrington, D. P., Zhang, Q., van Kammen, W., & Maguin, E. (1993). The double edge of protective and risk factors for delinquency: interrelations and developmental patterns. Development and Psychopathology, 5, 683-701.

*Stringfield, N. (1977). The impact of family counseling in resocializing adolescent offenders within a positive peer treatment milieu. Offender Rehabilitation, 1, 349-360.

*Swar, R-B., Jayaratne, S., & Tripod, T. (1976). Changing adolescent deviant behavior through reprogramming of the behavior-of parents and teachers: An experimental evaluation. Canadian Journal of Behavioral Science, 8, 132-143.

Thornton, D.M. (1987). Treatment effects on recidivism: A reappraisal of the 'nothing works' doctrine. In B.J. McGurk, D.M. Thornton, & M. Williams (Eds.). Applying Psychology to Imprisonment (pp. 181-189). London: HMSO.

*Turner, S., & Petersilia, J. (1992). Focusing on high-risk parolees: An experiment to reduce commitment to the Texas Department of Corrections. Journal of Research in Crime and Delinquency, 29, 34-61.

*Vingilis, E., Adlaf, E., & Chung, L. (1982). The Oshawa Impaired drivers Programme: An evaluation of a rehabilitation programme. Canadian Journal of Criminology, 22, 93-102.

*Vito, G.F. (1989). The Kentucky Substance Abuse Program: A private program to treat probationers and parolees. Federal Probation, 53, 65-72.

Von Voorhis, P. (1985). Restitution outcome and probationers' assessments of restitution. Criminal Justice and Behavior, 12(3), 259-287.

*Wade, T.C., Morton, T.L., Lind, J.E., & Ferris, N.R. (1977). A family crisis intervention approach to diversion from the juvenile justice system. Juvenile Justice Journal, 28, 43-51.

Walitzer, K.S., Dermen, K.H., & Connors, G.J. (1999). Strategies for preparing clients for treatment: A review. Behavior Modification, 23(1), 129-151.

Walsh, A. (1985). An evaluation of the effects of adult basic education on rearrest rates among probationers. Journal of Offender Counselling, Service and Rehabilitation, 9(4), 69-76.

Watson, R.J., & Stermac, L.E. (1994). Cognitive group counselling for sexual offender. International Journal of Offender Therapy and Comparative Criminology, 38(3), 259-270.

*Wexler, H.K., Falkon, G.P., & Lipton, D.S. (1990). Outcome evaluation of a prison therapeutic community for substance abuse treatment. Criminal Justice and Behavior, 17(1), 71-92.

Whitehead, J.T., & Lab, S.P. (1989). A meta-analysis of juvenile correctional treatment. Journal of Research in Crime and Delinquency, 26(3), 276-295.

*Whitaker, I.M., & Severy, L.J. (1984). Service accountability and recidivism for diverted youth: A client- and service-comparison analysis. Criminal Justice and Behavior, 12, 47-74.

*Williams, L. (1984). A police diversion alternative for juvenile offenders. Police Chief (Feb), 54-56.

*Willman, M.T., & Snorturn, J.R. (1982). A police program for employment and youth gang members. International Journal of Offender Therapy and Comparative Criminology, 26, 207-214.

*Winterdyk, J., & Roesch, R. (1982). A Wilderness Experimental Program as an alternative for probationers. Canadian Journal of Criminology, 23, 39-49. -9-Y

*Witte, A.D. (1977). Work release in North Carolina - A program that works! Law and Contemporary Problems, 41, 230-251.

Wolf, F.M. (1986). Meta-Analysis: Quantitative methods for research synthesis. Newbury Park, CA: Sage.

*Wormith, J. S. (1984). Attitude and behavior change of correctional clientele. Criminology, 22(4), 595-618.

Yeaton, W.H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. Journal of Consulting and Clinical Psychology, 49, 156-167.

*Young, M.C., Gartner, J., O'Connor, T., Larson, D., & Wright, K. (1995). Journal of Offender Rehabilitation, 22, 97-118.

*Zeisel, H. (1982). Disagreement over the evaluation of a controlled experiment. American Journal of Sociology, 88, 378-389.

*Zink T.M. (1970). A study of the effect of prison education on societal adjustment.

Journal of Correctional Education, 22, 18-20.

Appendix A

Correctional Program Assessment Inventory (CPAI) - Scoring Guide

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

February 1996: Gendreau & D.A. Andrews (with minor corrections by D.A. Andrews)

Items noted with * are scored. Items are scored either 0 or 1

A. PROGRAM DEMOGRAPHICS

1. Name of program:	Name specific program.
2. Name of contact person:	Name, title, address, phone, fax # of program setting.
3. Years in operation:	Record consecutive years of program operation to present.
4. Setting:	Record whether the program is housed in a community residential centre, or institution/prison, a probation office or a combination of these. If an institution/prison setting note whether it is a maximum, medium or minimum setting. Briefly describe structure (e.g., cells/rooms dorms, academic and vocational classrooms, gym, dining room, lounge, etc.).
5. Number of offenders/
probationers/residents	Record present #
6. Number of staff:	Full time staff are those who work at least 35 hours a week unless otherwise specified by program.
7. Program budget:	The budget amount in dollars applies only to the program itself unless program and setting are synonymous.
8. Government/private:	A government program is a directly funded part of the government agency/department that provides all resources to the program. A private sector program receives a time-limited grant/contract from various sources, e.g., government, business and charitable organizations.
9. Program philosophy:	Check only if program provided a documented rationale/philosophy/mission statement. Append the relevant document.

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

B. PROGRAM IMPLEMENTATION

Program Director is defined as the individual currently responsible for the overall management of the program. Items #10-14 refer to Program Director.

*10. Program initiation:	Program director either was the sole originator of the program or was instrumental in helping set up the program.
*11. Qualifications:	Is professionally trained (university degree) in one of the helping professions (e.g., education, nursing, psychology, social work, with course specialization in correctional/forensic/legal area).
*12. Previous experience:	Full time experience with an offender treatment program of any description for at least 3 years.
*13. Select/train staff:	Directly involved in hiring and providing training to staff.
*14. Conduct program:	Involved in conducting some aspects of the program that involves a) direct service delivery to clients, and b) direct supervision of staff.
*15. Literature review:	Literature search conducted to identify relevant program materials to design program. Search must consist of some of the major criminological and psychological journals and key texts. Look for a comprehensive attempt. Reference to a book or two, a government report or a couple of journal articles does not suffice. Record information in recording notes.
*16. Pilot:	Before formal program began, pilot program undertaken to sort out program logistics and content. Pilot lasted at least one month.
*17. Need:	Documented assessment of need for program. Append information, if available.
*18. Value congruency	Values and goals of program consistent with existing values in institution/community. If not, note concerns in recording notes.
*19. Cost-effective:	Perceived by administration and line staff, and/or the community to be less costly than other alternatives. If not, note concerns in recording notes.

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

- *20. Sustainable: Funding considered adequate for the task (operational and personnel). If not, note concerns in recording notes.

C. CLIENT PRE-SERVICE ASSESSMENT

21. Presenting problem: Examples are substance abuse, emotional problems, anti-social values, sexual offending, etc.
- *22. Appropriateness: The type of client presently received by the program is appropriate as defined by programmers. If not, note concerns in recording notes.
- *23. Exclusions: There is a rational clinical or legal basis for the exclusion of certain types of clients from program participation. Note what types in recording notes.
- *24. Risk: A reasonable survey of risk factors for recidivism (e.g., age, number of convictions) from the perspective of the programmers. This criteria also applies to #26, #28.
- *25. Method: This item is scored if the risk factors are measured by a recognized psychometric scale/ test (e.g., LSI-R). File records and clinical notes are acceptable if the relevant information is quantified.
- *26. Need: As with #25 - needs are dynamic attributes of offenders and their circumstances that are linked with recidivism (e.g., anti-social attitudes)
- *27. Method: As with # 25.
- *28. Responsivity: The responsivity of offenders to different styles and modes of service. How characteristics of offenders may interact with style and mode of service. Examples: low conceptual level offenders respond better to higher levels of program structure, higher anxiety offenders respond poorly to confrontation, highly verbal approaches would be inappropriate with those of low verbal ability.
- *29. Method: As with # 25.

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

- *30. Risk level defined: Clients personal characteristics are summarized as to level of risk - either qualitatively (high-medium-low) or an actual score such as LSI-R of 24+ = high risk or Wisconsin of 2= low risk.
- *31. Need factors defined: As with #30.
- *32. Responsivity factors defined: As with #30.

D. PROGRAM CHARACTERISTICS

- *33. Program targets: The program must target criminogenic behavior and attitudes; only then will offenders criminal behavior be reduced. Please note that if the program targets 3 or less behaviors, all must be in the set noted below. If 4 behaviours then 3, and if 5 or more, 80% of program targets must be in the acceptable set noted below:
- a) Change attitudes, orientations and values favorable to law violations and anticriminal role models
 - b) Change antisocial feelings
 - c) Reduce antisocial peer associations
 - d) Reduce problems associated with alcohol/drug abuse
 - e) Reduce anger/hostility level
 - f) Replace the skills of lying, stealing and aggression with prosocial alternatives
 - g) Increase self-control, self management and problem solving skills
 - h) Encourage constructive use of leisure time
 - i) Improve skills in interpersonal conflict resolution
 - j) Promote more positive attitudes/increase performance re: school/work
 - k) Promote family affection/communication
 - l) Promote family monitoring and supervision
 - m) Improve family problem solving
 - n) Resolve deviant sexual arousal

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

- o) Provide low-pressure, sheltered environment for mentally disordered offenders
- p) Relapse prevention: ensuring the client is able to recognize high risk situations that lead to low-breaking and has a concrete plan to deal with these situations
- q) Alleviate the personal and circumstantial barriers to service (client motivation, background stressors).

*34. Treatment:

In recording responses as to type of treatment, please note the name of the treatment, e.g., social skills training, group therapy, milieu therapy, family therapy, etc. Enquire about the details of the treatment (obtain a program outline or manual) and then check this item if any of the following criteria are met.

Acceptable responses re: type of treatment are noted below. These strategies have been demonstrated to be effective in reducing recidivism.

- a) Radical behavioral - classical conditioning (e.g. desensitization), operant conditioning schedules of reinforcement (e.g., fixed interval variable ratio), involving prompting, shaping, fading and stimulus control strategies that are directed to specific criminogenic behaviors. Token economies and contingency management strategies are typical examples of operant programs.
- b) Social learning - modeling and behavioral rehearsal techniques that engender self-efficacy.
- c) Cognitive behavioral- cognitive therapy, cognitive skills training, problem solving therapy, rational -emotive therapy, self- control procedures, self-instructional training, stress inoculation training are some of the common techniques in this area
- d) Programs that, while denying they are behavioral in nature, state criminogenic need (see #34) was targeted and a structured intervention was employed to change the need being targeted. Investigate this item carefully.

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

	Record necessary information in recording notes.
Notes:	<p>Those strategies that are <u>ineffective</u> in reducing recidivism are non-directive, psychoanalytic, group milieu therapies and chemotherapy's. "Punishing smarter" programs (e.g., boot camps, drug testing, electronic monitoring, shock incarceration) are also ineffective.</p> <p>For a point to be recorded for # 34 there must be a match between what is defined as an acceptable target behaviour and the treatment on the basis of the categories noted in #33. Please note that if no point is recorded for #33, no point can be given to #34.</p>
*35. Location:	In a prison, program participants are separated from rest of population unless the entire institution is involved in the program. In the community, whereabouts of clients and peer associations are closely monitored by staff or other means.
*36. Manual:	Program has a detailed program/activities manual. Append information.
*37. Involvement:	Clients spend at least 40% of time per week in therapeutic tasks which can also include work and leisure activities.
*38. Intensity/risk:	The higher risk clients receive the highest intensity/duration of service.
*39. Treatment & Client:	Offenders are assigned to programs that match up best with their interests, style of learning etc. See #28 for examples.
*40. Staff & program:	Staff assigned to a program their skills match up with best (e.g., staff who function best with higher degree of structure work in a contingency management program).
*41. Staff & client	Staff are assigned to clients they can work with effectively (e.g. high conceptual level staff assigned to like clients)
*42. Client input:	Input into some programmatic structures and can modify some features of rules and regulations of

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

- the program. Must be approved by programme supervisor.
- *43. Ratio: Rewards should outnumber punishers by at least 4:1. Punishment: score items #45 -47 only if it can be clearly established that the punishment used in the program are clearly intended to suppress anti-social behavior after termination of the program.
- *44. Theory: Theory supporting the type of punishment used is described (e.g., negative law of effect, competing response theory).
- *45. Stimuli: Effective punishing stimuli are electric shock, drug induced aversion, mild aversive stimuli, covert sensitization, overcorrection, time out, and response cost.
- *46. Procedure: Punishing stimuli are administered in the following manner: escape is impossible, maximum intensity, earliest point in deviant response, after every occurrence of deviant behavior, immediate, not spread out, alternative prosocial behaviours provided after punishment administered, vary punishers. Last item is optional.
- *47. Negative effects: There is assessment as to whether the punishment produce emotional reactions (fear, interfere with new learning, disrupts social relationships), produces avoidance/aggression toward punishers, increases future use of punishment by offender, produce response substitution, and lack generalization. For this item to be scored 4 of the 5 criteria should be met.
- *48. Completion criteria; Criteria clearly outlined as to when program terminates for each client. Termination should be defined by progress in acquiring pro-social behaviors, beliefs while in program or engaging in behaviour that seriously jeopardized the safety of staff and other clients.
- Note: #49 -54 pertain to relapse prevention strategies which are commonly applied to substance

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

- abusers and sex offenders. Please note #52-54 are infrequently found in prison based settings.
- *49. Monitor: Client trained to observe and anticipate problem situations
- *50. Rehearse: Client plans and rehearse alternative prosocial responses.
- *51. Practice: Client practices new prosocial behaviours in increasingly difficult situations. Improved competencies rewarded.
- *52. Brokerage Client referred to other services that have established formal links with the present program under review and that offer service relevant to offenders needs. There must be documented confirmation the services have been received.
- *53. Support/therapists: Significant others (e.g., family, friends) trained to provide support.
- *54. Booster session: Client returns to relearn/reinforce skills originally taught in formal treatment phase of program.

E. STAFF CHARACTERISTICS/PRACTICES

All items refer to full/part time staff (excluding program director).

- *55. Education: 75% staff have an undergraduate degree 10% of staff have an advanced degree.
- *56. Area of study: 75% of staff have training in criminal justice, education, nursing, psychology, social work or specialized fields (e.g., addictions).
- *57. Experience: 75% of staff have worked in treatment programs with offenders for at least 1 year.
- *58. Personal qualities: Staff are hired on at least 5 of the following factors: clarity, conceptual level, empathy, fairness, life experiences, non-confrontational but firm, problem solving, spontaneous. Record evidence in notes. Also, should be a check on persons background.
- *59. Stability: 50% of staff have remained on job for at least 2 years.

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

*60. Assessment/clinical supervision	Staff are assessed yearly on clinical skills related to service delivery. Also, staff receive regular clinical supervision. Documentation required.
*61. Training:	Staff receive formal training in theory and practice of intervention employed. Staff receive further exposure (e.g., workshops, courses, on-job training, conferences) relevant to program material. One instance per staff per year. Documentation required.
*62. Program input:	Staff are able to modify program structure. Modifications must be approved by program supervisor/review board. Record examples in notes.

F. EVALUATION

*63. Quality assurance:	A management audit system mechanism in place that consists of a) within program checks, file review, problem oriented records, etc., that monitor treatment progress, b) clinical supervision and c) client feedback.
*64. Consumer satisfaction:	Clients surveyed each year as to satisfaction with services
*65. Within program:	Periodic, objective, standardized assessment of clients on program target behaviours. Must be tangible evidence in files.
*66. Follow-up:	Client re-arrest, reconviction, or reincarceration data gathered at 6 months or more after leaving program.
*67. Methodological quality:	As above, with at least one evaluation in last 5 years, comparing treatment outcome with a risk control comparison group of some sort.
*68. Unpublished report:	A document containing introduction, method, results, discussion on file that details effectiveness of program.
*69. Peer review:	As above, in an edited journal.

G. OTHER

*70. Client records:	Client records are kept in a confidential file. Records must have social history, record of

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

	presenting problem, assessment data, program progress notes, etc. Review actual files.
*71. Ethical guidelines:	Documentation of the ethics of intervention (e.g., least intrusive intervention, etc.).
*72. Program change:	No noticeable changes in program components in last 2 years that jeopardized smooth functioning of program.
*73. Program funding:	As above in #72.
*74. Program community support:	As above in #72
*75. Advisory board:	A group of individuals (e.g., Board of Directors, or one person such as a consultant) officially designated to advise program.

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

H. SCORING

Record a % score for each subsection and for the total. A% score is used as some N/As may apply to some questions and the points available in each section vary.

Category		Pts. available	N/A	# Pts. Scored	%
A.	Program demographics	no scoring			
B.	Program implementation	11			
C.	Client pre-service assessment	11			
D.	Program characteristics	22			
E.	Staff characteristics	8			
F.	Evaluation	7			
G.	Other	6			
Total		65			
	Very satisfactory	70% +			
	Satisfactory	50-69%			
	Unsatisfactory	below 50%			

I. EVALUATOR RECORDING NOTES

Use this section to record problematic and contextual aspects of your evaluation and any qualifications you may wish to make re: confidence in your results and ratings.

J. EVALUATOR

Name and position titles of evaluator

CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI)
SCORING GUIDE

Additons by A. Nesovic: Interrater reliablity study..

*76. Motivation:	Motivation for treatment is assessed prior to the beginning of the program. Score this item as present if an attempt is made to assess whether the offenders are motivated for treatment and one of the following methods was used: interview, survey, or any motivational scale (e.g., URICA)
*77. Motivation assignment:	Score this item as present if motivation was taken into account when assigning offenders to particular program. Please note: if no point was given for previous item (motivation) no point can be given for this item.
*78. Adherence	Treatment staff adhere to planned treatment intervention. Score this item as present if there is an objective method of assessment: e.g., videotaping, monitoring sessions
*79. Productivity of therapy sessions:	Scored as present if there is an objective method of assessment: e.g., questionnaire, interview with participants
*80. Staff beliefs:	Staff beliefs in effectiveness of correctional program were taken into account when hiring treatment staff.
*81. Favorable conditions	There are favorable conditions for program delivery. Score if there are indications that agency/organization is supportive of the programs and/or provide resources and/or has good line of communication and/or good manager/staff relations.
*82. Ratio/Size of the treatment group:	Score as present if a) size of the group is between 4-8 and b) maximum 1:4 staff ratio

Appendix B

Correctional Program Assessment Inventory Questionnaire (CPAIQ)

PROGRAM CHARACTERISTICS

1. Please note the name of your program and its location:

2. Your name:

3. Your title:

4. Address and telephone number:

5. How many years has this program been operating?

6. Please describe the program setting and its security level (e.g., community, maximum security prison, etc.):

7. What percentage of the program participants are offenders, males and females?

Offenders _____%

Males _____%

Females _____%

8. What is the number of full and part time staff in the program:

Full time _____

Part time _____

9. What is the program budget (in dollars)?

10. Is the program funded by the government? Yes No

If no, please specify the source of funding:

11. Is a documented program's philosophy available? Yes No

If yes, please provide details of the program's philosophy:

PROGRAM IMPLEMENTATION

12. Did the program director¹ participate in designing the program?

Yes

No

13. Was the program director professionally trained?

Yes No

If yes, please describe his/her qualifications.

14. Had the program director been involved in conducting a treatment program in the past?

Yes

No

If yes, please specify the length of his/her experience in years.

15. Was the program director involved in:

Hiring staff

Yes

No

Training staff

Yes

No

Supervision of staff

Yes

No

Service delivery to clients

Yes

No

1. In case that the program does not have a program director the questions #12-15 apply to the individual responsible for the overall management of the program

16. Before the program began, was a literature review conducted to assist in the program design?

Yes

No

If yes, please provide a brief list of the references used (e.g., books, journal articles, etc.).

17. Was a pilot program conducted prior to the beginning of the formal program in order to sort out program logistics and content?

Yes

No

If yes, please indicate the duration of the pilot program in months

18. Was there an assessment of the need for the program in the institution or community?

Yes

No

If yes, please specify how it was done.

19. Are the values and goals of the program consistent with existing values in institution/community?

Yes

No

If not, please note the concerns.

20. Do administration/line staff/community perceive the program as cost effective? Yes
No

If not, please note the concerns.

21. Is the program funding considered adequate for the task (operational and personnel)?

Yes

No

If not, please note the concerns.

CLIENT PRE-SERVICE ASSESSMENT

22. Please outline the methods and procedures (e.g., questionnaires, clinical ratings, interviews, etc.) by which the program assesses the clients.

23. Please indicate whether and why the program provided is appropriate for the clients?

24. Is there a rational exclusion criteria of certain clients from the program? Yes No
If yes, please note the exclusion criteria.

25. Were risk factors for recidivism assessed prior to the beginning of the program?

Yes No

If yes, please specify what method of assessment is used?

26. Are the results of the client's risk assessment summarized in a way which provides an index of the client's risk level?

Yes No

27. Were need factors (dynamic attributes of offenders and their circumstances linked to criminal behavior) assessed prior to the beginning of the program?

Yes No

If yes, please specify what method of assessment is used.

28. Are the results of the client's need assessment summarized in a way which provides an index of the client's need factors?

Yes No

29. Were the changes in risk/need factors reassessed periodically?

Yes No

If yes, please specify at which interval risk/need factors were reassessed.

30. Were responsivity factors (offender's personal characteristics and styles of interaction) assessed prior to the beginning of the program?

Yes No

If yes, please specify what method of assessment is used.

31. Are the results of the client's responsivity assessment summarized in a way which provides an index of the client's responsivity factors?

Yes No

32. Was motivation for treatment among clients assessed prior to the beginning of the program?

Yes No

If yes, please specify what method of assessment is used.

33. Was motivation for treatment among clients reassessed periodically?

Yes No

If yes, please specify at which interval it was reassessed.

PROGRAM CHARACTERISTICS

34. Please specify what types of behavioral /psychological change the program targets.

35. Please specify what types of treatment are provided to the clients (e.g. psychodynamic, behavioral, cognitive, etc.).

36. If the program is delivered within institution, are the clients separated from the rest of prison population?
Not applicable Yes No
37. If the program is delivered within community, are the clients closely monitored?
Not applicable Yes No
38. Does the program have a detailed program/activities manual? Yes No
39. Please note how much time per day/week the clients spend in program related activities.

40. Please specify the client/staff ratio.

41. Are there favorable conditions for the program delivery within the agency/
institution/community?
Yes No
42. Is there an assessment in place of how productive therapy sessions are?
Yes No

43. Do the intensity and duration of treatment vary according to the risk level of clients?

Yes No

If yes, please indicate how this is done.

44. Does the program match the type of treatment with the characteristics of the clients?

Yes No

If yes, please indicate how this is done.

45. Does the program match the type of treatment with the skills of the treatment providers?

Yes No

If yes, please indicate how this is done.

46. Does the program match the personal characteristics and skills of the treatment providers with the individual characteristics of the clients? Yes No

If yes, please indicate how this is done.

47. Are the clients able to modify the functioning of the program? Yes No

If yes, please give example of changes made to the program based on client input and note whether these changes were approved by the program supervisor.

48. Please describe the incentives/rewards and discentives/punishers used within the program.

49. Please, specify what is approximate ratio of rewards to punishers.

50. Please describe the theory that underlies the type of punishment used in the program.

51. Please describe punishing stimuli administered within the program.

52. Please describe how these punishing stimuli are administered.

53. Is there an assessment of the negative effects of the punishment (e.g. fear, interference with new learning).

Yes No

If yes, please note how these effects were assessed.

54. Is program completion criteria defined?

Yes No

If yes, please note completion criteria.

55. Are the clients trained to observe and anticipate problem situations?

Not applicable Yes No

If yes, please describe the training they receive.

56. Are the clients trained to rehearse alternative behavior?

Not applicable Yes No

If yes, please describe the training they receive.

57. Are the clients trained to practice new behavior in increasingly difficult situations?

Not applicable Yes No

If yes, please describe the training they receive.

58. After completing the program, are the clients referred to the services relevant for their needs?

Not applicable Yes No

If yes, please describe the services they are referred to.

59. Are significant others (e.g., family, friends) trained to provide support?

Not applicable Yes No

If yes, please describe the training they receive.

60. Do the clients come back for “booster sessions” after release?

Not applicable Yes No

If yes, please describe these sessions.

STAFF CHARACTERISTICS

61. Please indicate the gender, education and major area of professional training of your staff by entering the number of staff for each of the categories below.

Gender	Number of Staff
Males	_____
Females	_____

Education	Number of Staff
High School	_____
Community College	_____
University Certificate	_____
BA	_____
BSc	_____
MA/MSc	_____
PhD	_____
Other _____	_____

Area of Study	Number of Staff
Clergy/Theology	_____
Criminology	_____
Education	_____
Medicine	_____
Nursing	_____
Psychology	_____
Social Work	_____
Sociology	_____
Other _____	_____

62. Please estimate the average years of experience that your staff have in working with offenders (Please circle the appropriate number).

1. Less than 6 months
2. Six months to two years
3. Two to five years
4. Five to ten years
5. More than ten years

63. Were staff personal characteristics considered in the hiring of staff?

Yes No

If yes, please list the personal characteristics considered in the hiring of staff.

64. What has been the turnover among staff during the last three years? Please indicate the number of staff who have resigned or were dismissed during this period.

65. Are staff clinical skills assessed periodically by supervisors?

Yes No

If yes, please indicate the frequency and nature of the assessment.

66. Do staff receive regular clinical supervision?

Yes No

67. Do program staff adhere to planned program interventions?

Yes No

If yes, please describe how it is monitored?

68. Do program staff participate in on-going training programs, workshops or conferences?

Yes No

If yes, please indicate how many of these events staff have attended during the last two years?

69. Do staff have input into the design and functioning of the program?

Yes No

If yes, please give few examples.

70. Are staff able to modify the functioning of the program?

Yes No

If yes, please give example of changes made to the program based on staff input and note whether these changes were approved by the program supervisor.

71. Please indicate whether the following factors were considered when selecting program staff:

- a) Relationship/Interaction factors (e.g. warm, genuine, with sense of humor, enthusiastic, self-confident, empathic, respectful, flexible, engaging, non-blaming, reflective, mature, intelligent)

Yes No

- b) Belief Factors (e.g. committed to human service and believes that offenders can change)

Yes No

- c) Skill factors (e.g. directive, solution focused, use prosocial modeling, effective reinforcement, effective disapproval, problem solving practices)

Yes No

72. Please indicate whether program staff are trained (pre-service or in-service) and/or clinically supervised on the following practices:

- a) Staff use effective reinforcement (e.g. immediately tell client when they approve of the type of behavior/speech exhibited and they explain why, encourage client to think about why it is desirable behavior)

trained Not applicable Yes No

supervised Not applicable Yes No

- b) Staff use effective modeling (e.g. staff demonstrate desired behavior in concrete and vivid ways, the client participating in the program is rewarded for demonstrating desired behavior, the model is the source of reinforcement rather than being punishing or negative).

trained Not applicable Yes No

supervised Not applicable Yes No

- c) Staff use effective disapproval (e.g. tell clients why they do not like exhibited behavior and encourage clients to think why it is undesirable behavior, the staff provide appropriate anticriminal modeling and when the client exhibits appropriate behavior the staff approve of the change).

trained Not applicable Yes No

supervised Not applicable Yes No

- d) Staff use problem solving techniques (e.g. identify problems in behavior and its consequences, generate alternative solutions to the situations, and encourage clients to evaluate alternative options).

trained	Not applicable	Yes	No
supervised	Not applicable	Yes	No
- e) Staff use structured learning procedures (e.g. defines skill to be learned, model the skill for the client, give the client the opportunity to practice the skill in hypothetical situations, provide feedback to the client)
trained	Not applicable	Yes	No
supervised	Not applicable	Yes	No
- f) Staff use authority effectively (e.g. they are direct and specific concerning their demands, specify the choices and their consequences, give encouraging messages, support words with actions, "firm but fair").
trained	Not applicable	Yes	No
supervised	Not applicable	Yes	No
- g) Staff use cognitive restructuring (e.g. encourage the clients to describe a situation, thoughts and feelings, to recognize risky thinking, and to practice alternative thoughts).
trained	Not applicable	Yes	No
supervised	Not applicable	Yes	No

73. Please specify any other practices that the staff were trained and/or supervised on.

EVALUATION

74. Do supervisors provide Quality Assurance assessments? Yes No

If yes, please describe the assessment conducted.

75. Were clients surveyed as to their satisfaction with the program? Yes No

If yes, please note how frequently they are surveyed?

76. Please indicate which client behaviors are assessed during the program as well as how they are assessed (e.g., objective tests, observations, interviews) and at what intervals the assessment is conducted?

77. Are follow-up recidivism data for the clients gathered? Yes No

If yes, please specify the duration of follow-up period:

78. Has any formal evaluation of the program been carried out? Yes No

If yes, please describe program evaluation (e.g., procedure, when it was done):

79. Is there any unpublished report that details effectiveness of the program? Yes No

If yes, please describe what that report consists of:

80. Has any evaluation of the program been published in an edited journal? Yes No
If yes, please specify the name of the journal(s):

OTHER

81. Are client records kept in a confidential file? Yes No
If yes, please describe the content of the file.

82. Is the ethics of intervention documented? Yes No
If yes, please specify how it is documented.

83. Please indicate if there has been any noticeable change during the last two years
regarding program components?
If yes, please describe its effect on the program.

84. Please indicate if there has been any noticeable change during the last two years regarding program funding?

If yes, please describe its effect on the program.

85. Please indicate if there has been any noticeable change during the last two years regarding community support for the program?

If yes, please describe its effect on the program.

86. Has a program advisory committee been established to provide advice to the program.

Yes No

If yes, what are the professional affiliation of committee members.

87. Having completed this questionnaire, please indicate whether you feel there are any essential points which have been omitted or require elaboration.

Appendix C

List of the Participating Program Sites

Federal programs-institutional

Bath Institution	Anger management
Bath Institution	Sex offenders program
Joceville Institution	OSAPP
Pittsburgh Institution	Cognitive skills
Millhaven Institution	Violence prevention program
Grand Valley Institution	Survival of trauma

Federal programs - community

Toronto Parole Office	Choices
Hamilton Parole Office	Boosters program-Cognitive Skills
Ottawa Parole Office	Cognitive skills
John Howard Society of Ottawa	Counterpoints

Provincial programs - institutional

Rideau CC	10 week cognitive/behavioral treatment program
Guelph CC	Anger and aggression program
Vanier Center	Antisubstance abuse program
Ottawa Carleton DC	Psychological rehabilitation
Burtch CC	Tough Talk
Toronto West DC	Anger management for man

Provincial programs community

Elisabeth Fry Society	Majory Amos
St. Leonard London	Anger Management
Catholic Services Ottawa	New Directions
John Howard Society Peel	Bail verification

Appendix D

Interrater Reliability Indicators - Two Coded Categories

CPAI Sections	Items	Kappa	Pearson r
Program implementation	11	.89**	.94**
Preservice Assessment	15	.89**	.98**
Program Characteristics	24	.84**	.90**
Staff characteristics	9	.74**	.61**
Evaluation	7	.75**	.82**
Other	6	.95**	.95**
Treatment Score	33	.87**	.95**
Total score	72	.86**	.94**

** $p < .01$

Appendix E

Kappa Coefficients for Individual Program Sites.

Program	Kappa Coefficient	
	Written	Interview
1	.93**	
2	.91**	.834**
3	.86**	.841**
4	.78**	
5	.88**	.94**
6	.92**	.77**
7	.73**	
8	.72**	.80**
9	.82**	
10	.80**	.98**
11	.92**	.82**
12	.77**	
13	.93**	.82**
14	.96**	
15	.82**	.85**
16	.68**	.88**
17	.73**	.76**
18	.67**	.93**
19	.86**	.85**
20	.76**	

**p<.01

Appendix F

Ranges within the CPAI Sections (CPAI-Q)

CPAI Sections	Rater 1		Rater 2	
	Min	Max	Min	Max
Program Implementation	54.50	100.00	54.70	100.00
Pre service Assessment	6.00	100.00	.00	100.00
Program Characteristics	18.70	89.40	23.50	89.40
Staff characteristics	37.50	100.00	37.50	100.00
Program Evaluation	.00	100.00	.00	.00
Other	.00	100.00	.00	.00
Treatment Score	12.90	94.10	12.50	94.50
Total Score	29.30	87.30	25.00	85.70

Appendix G

Ranges within the CPAI Sections (Interview)

CPAI Sections	Rater 1		Rater 2	
	Min	Max	Min	Max
Program Implementation	45.45	90.90	45.45	90.90
Preservice Assessment	13.33	86.66	13.33	86.66
Program Characteristics	25.00	68.42	27.27	73.68
Staff Characteristics	33.30	88.88	37.50	100.00
Program Evaluation	.00	100.00	.00	100.00
Other	.00	100.00	.00	100.00
Treatment Score	21.21	72.72	26.47	78.78
Total score	28.81	77.72	35.00	76.62

Appendix H

Effective Treatment Elements Questionnaire (ETEQ)

Instructions:

Please indicate for each of the following treatment elements how important you consider them to be for the effectiveness of a correctional program. Please write 1, 2, 3, or 4 in the blank next to each statement using the following scale:

1. treatment element is not important
2. treatment element is somewhat important
3. treatment element is important
4. treatment element is very important

Example:

1. _____ Program director selects/trains staff

If you consider that the presence of this treatment element in a correctional program is **not important** for its effectiveness, you should write 1 in the blank. If you think that it is **somewhat important**, you should write 2. If you consider the treatment element **important**, you should write 3. If you think that this element is **very important** for program effectiveness, you should write 4.

Not important	Somewhat important	Important	Very important
1	2	3	4

PROGRAM IMPLEMENTATION

1. _____ Program director¹ is involved in designing the program.
2. _____ Program director is professionally trained (e.g., in psychology, social work).
3. _____ Program director has worked with an offender treatment program for at least 3 years
4. _____ Program director selects/trains staff.
5. _____ Program director conducts some aspects of the program (e.g. participate in program delivery and directly supervise staff)
6. _____ Literature review is conducted to identify relevant program materials.
7. _____ Pilot program is conducted to determine practical aspects of the program.
8. _____ There is documented assessment of need for the program.
9. _____ Values and goals of the program are consistent with existing values in institution/ community.
10. _____ Program is perceived by administration and program delivery staff as cost effective.
11. _____ Program funding is considered adequate.

CLIENTS² PRESERVICE ASSESSMENT

12. _____ Treatment provided is appropriate for the type of offenders participating in the program.
13. _____ Exclusion criteria of certain types of offenders is based on a rational clinical or legal basis.
14. _____ Risk factors³ for reoffending (e.g., age, number of convictions) are assessed.
15. _____ Risk factors are assessed by a recognized psychometric scale/test or standardized interview that can be quantified.
16. _____ Risk level is defined (e.g., high, low).
17. _____ Criminogenic needs⁴(e.g., anti-social attitudes, anti-social peers) are assessed.

1. Individual responsible for the overall management of the program
 2. Offender
 3. Characteristics of offenders that are associated with reoffending.
 4. Dynamic characteristics of offenders that are associated with reoffending

Not important	Somewhat important	Important	Very important
1	2	3	4

18. _____ Criminogenic needs are assessed by a recognized psychometric scale/test or standardized interview that can be quantified.
19. _____ Criminogenic needs are identified.
20. _____ The changes in risk/needs were reassessed periodically.
21. _____ Offender's personal characteristics and styles of interactions (e.g., intelligence, verbal ability, level of anxiety) are assessed.
22. _____ Offender's personal characteristics and styles of interaction are assessed by a recognized psychometric scale/test or standardized interview that can be quantified.
23. _____ Responsivity factors are specified.

PROGRAM CHARACTERISTICS

24. _____ Program targets criminogenic behavior and attitudes.
25. _____ Behavioral/cognitive treatments are provided to the offenders.
26. _____ Program participants are separated from the rest of prison population.
27. _____ Program has a detailed program/activities manual.
28. _____ Offenders spend the majority of time each week in program related tasks.
29. _____ Intensity and duration of program vary according to the risk level of the offenders.
30. _____ Programs match the type of treatment with the characteristics of the individual offenders.
31. _____ Programs match the personal and professional skills of the staff with the type of treatment they provide.
32. _____ Programs match the personal and professional skills of the staff with the type of offenders and the nature of his/her problem.
33. _____ Motivation for change among offenders is assessed.
34. _____ Offenders have input into structure and rules of the program.
35. _____ Rewards used in the program outnumber punishers used in the program.
36. _____ The theory supporting type of punishment used is described.
37. _____ Effective punishing stimuli are used (e.g., time out).
38. _____ Punishing stimuli are administered in the specific manner (e.g., maximum intensity, after every occurrence of undesired/inadequate behavior).
39. _____ Emotional reactions to punishing stimuli are assessed.

Not important	Somewhat
important	Important	Very
important
1	2	3	4
40. _____ Program completion criteria is clearly outlined.
41. _____ Offenders are trained to observe and anticipate problem situation.
42. _____ Offenders plan and rehearse alternative prosocial responses.
43. _____ Offenders practice new prosocial behaviors in increasingly difficult situations.
44. _____ After program completion, offenders are referred to other services that offer programs relevant for their needs.
45. _____ Significant others (e.g., family, friends) are trained to provide support.
46. _____ After program completion, offenders are brought back into program for “booster” sessions.
47. _____ Therapist¹ speaks on behalf of offender at home/school/community.
48. _____ Therapist follows up on offender’s referrals.
49. _____ Treatment groups are small with high offender-staff ratio.
50. _____ Favorable conditions for program delivery exist within agency/institution/ community.
51. _____ Therapy sessions are productive.

STAFF CHARACTERISTICS

52. _____ The majority of the program staff have at least an undergraduate degree.
53. _____ The majority of the program staff are trained in human service, health, and mental health.
54. _____ The majority of the program staff are trained in a correction-related field (e.g., social work, psychology, criminology, criminal justice).
55. _____ The majority of the program staff have worked in treatment programs with offenders for at least 1 year.
56. _____ Personal qualities of program staff were taken into consideration when hiring staff
57. _____ Turnover of program staff is low.
58. _____ Program staff receive clinical supervision and their clinical skills are assessed regularly.
59. _____ Program staff are formally trained in theory and practices of treatment interventions used.

1. treatment provider

- | | Not important
1 | Somewhat
important
2 | Important
3 | Very
important
4 |
|-----------|---|----------------------------|----------------|------------------------|
| 60. _____ | Program staff can modify program with approval of program supervisor. | | | |
| 61. _____ | Program staff believe that they have skills to run the program effectively. | | | |
| 62. _____ | Program staff believe that a correctional program can reduce reoffending rates | | | |
| 63. _____ | Program staff adhere to planned program interventions. | | | |
| 64. _____ | Staff are warm, genuine, with sense of humor, enthusiastic, self-confident, empathic, respectful, flexible, engaging, non-blaming, reflective, mature, intelligent. | | | |
| 65. _____ | Staff are committed to human service and believes that offenders can change | | | |
| 66. _____ | Staff use effective reinforcement (e.g., immediately tell offender when they approve of the type of behavior/speech exhibited and they explain why, encourage client to think about why it is desirable behavior). | | | |
| 67. _____ | Staff use effective modeling (e.g. staff demonstrates desired behavior in concrete and vivid ways, the offender participating in the program is rewarded for demonstrating desired behavior, the model is source of reinforcement rather than being punishing or negative). | | | |
| 68. _____ | Staff use effective disapproval (e.g., tell offenders why they do not like exhibited behavior and encourage offenders to think why it is undesirable behavior, staff provide appropriate anticriminal modeling and when client exhibits appropriate behavior staff approves of change). | | | |
| 69. _____ | Staff use problem solving techniques (e.g., identify problems in behavior and its consequences, generate alternative solutions to the situations, and encourage clients to evaluate alternative options). | | | |
| 70. _____ | Staff use authority effectively (e.g., they are direct and specific concerning their demands, specify the choices with their consequences, give encouraging messages support words with actions, "firm but fair"). | | | |
| 71. _____ | Staff use structured learning procedures (e.g., defines skill to be learned, model the skill for the client, give the client opportunity to practice the skill in hypothetical situations, provide feedback to the client) | | | |

EVALUATION OF PROGRAM

72. _____ Supervisors conduct quality assurance assessments (e.g., file reviews, clinical supervision, offenders feedback).
73. _____ Offenders are surveyed on their satisfaction with service provided.
74. _____ There is objective, periodic, and standardized assessment of offenders on target behaviors.

Not important	Somewhat important	Important	Very important
1	2	3	4

- 75. _____ Offender reconviction data are gathered.
- 76. _____ At least one evaluation of program is conducted.
- 77. _____ The methods, results, discussion of the program effectiveness are documented.
- 78. _____ Evaluation of the program is published in a journal.

OTHER

- 79. _____ Offender’s records are kept in a confidential file.
- 80. _____ The ethics of intervention is documented.
- 81. _____ There are no changes in program components in the last two years.
- 82. _____ There are no changes in program funding in the last two years.
- 83. _____ There are no changes in community support for program in the last two years.
- 84. _____ An advisory board oversees/advises program.

Having completed this questionnaire, please indicate whether you think there are any important treatment elements which have been omitted.

PARTICIPANT INFORMATION

It will be greatly appreciated if you can provide the following information:

Gender _____ Age _____

Educational level attained (Please circle the appropriate category).

- 1 Secondary
- 2 Community College
- 3 University- Undergraduate Studies
- 4 University- Graduate Studies
- 5 Other, please specify _____

Primary area of study (Please circle the appropriate category).

- 1 Criminal Justice
- 2 Psychology
- 3 Criminology
- 4 Sociology
- 5 Clinical/Counseling
- 6 Education
- 7 Psychiatry
- 8 Nursing
- 9 Social Work
- 10 Other, please specify _____

Working in the correctional/criminal justice field.

Yes

No

If yes, how many years? _____ years.

Where (please circle appropriate category)?

Administration

Policy/development

Program delivery

Research

Academic

Other, please specify _____

Do you believe that correctional programs can decrease criminal behavior of its participants?

Yes

No

Not sure

Please state how familiar you are with

Literature on treatment effectiveness (for example: Lipsey, Losel, Andrews, Palmer, Dowden, Cullen, Gendreau, Bonta, Hoge)

1	2	3	4
not familiar	somewhat familiar	familiar	very familiar

Psychology of criminal conduct (Andrews & Bonta)

1	2	3	4
not familiar	somewhat familiar	familiar	very familiar

Correctional Program Assessment Inventory (CPAI, Gendreau & Andrews)

1	2	3	4
not familiar	somewhat familiar	familiar	very familiar

Appendix I

ETEQ Individual Items: Descriptive Statistics

Table I1:
Descriptives for Individual ETEQ Items (Psychologists)

	<u>N</u>	<u>M</u>	<u>SD</u>
program director designing program	49	2.88	.88
program director is professionally trained	49	3.51	.68
program director experience	49	3.10	.80
program director select/train staff	49	3.00	.82
program director conducts some aspect	49	2.92	.86
literature search	50	3.66	.56
pilot program	50	3.26	.85
need for program documented	50	3.38	.73
values/goals documented	49	3.26	.70
cost effective	50	2.62	.64
program funding adequate	50	3.30	.68
treatment appropriate	50	3.80	.45
exclusion criteria	50	3.60	.57
risk factors	50	3.58	.67
risk factors assessed	50	3.30	.79
risk level defined	50	3.36	.69
criminogenic need	50	3.64	.53
criminogenic need assessed	50	3.32	.82
criminogenic need identified	50	3.62	.60
changes assessed	50	3.32	.74
responsivity factors	50	3.28	.70
responsivity assessed	50	2.82	.90
responsivity factors specified	49	3.20	.76

Table 11:
Descriptives for Individual ETEQ Items (Psychologists) (Continued)

	<u>N</u>	<u>M</u>	<u>SD</u>
target criminogenic behavior and attitudes	50	3.74	.44
cognitive/behavioral treatment	50	3.72	.57
separated from the rest	50	2.22	1.07
manual	50	3.08	.80
time spent	50	2.90	.89
intensity/ duration	50	3.26	.85
treatment offender match	50	3.38	.67
staff treatment match	50	3.52	.54
staff offenders match	50	3.26	.72
motivation for treatment	50	2.90	.79
offenders input	50	2.48	.89
rewards/punishers	48	3.21	.92
theory of punishment	47	2.62	.95
effective punishing stimuli	47	2.66	.89
administration of punishing stimuli	47	2.62	.92
emotional reaction to stimuli	47	2.85	.91
completion criteria	50	3.56	.64
anticipate problem situation	50	3.54	.71
rehearse sociable acceptable responses	50	3.64	.53
practice socially acceptable	50	3.56	.61
referred to other services	49	3.69	.51
family trained	50	2.72	.86
booster	50	3.06	.93
therapist speak	49	1.80	.91
therapist follow up	48	2.75	.91
small treatment groups	50	3.18	.72
favorable condition for delivery	49	3.45	.68

Table 11:
Descriptives for Individual ETEQ Items (Psychologists) (Continued)

	<u>N</u>	<u>M</u>	<u>SD</u>
productive therapy sessions	48	3.46	.68
education staff	50	2.86	1.01
staff training	50	3.14	.81
staff trained in corrections	50	3.22	.82
staff experience	50	3.00	.83
staff qualities	49	3.51	.61
staff turnover	50	3.30	.61
staff supervision	50	3.44	.73
staff trained in intervention used	50	3.64	.53
staff has input	49	3.24	.72
staff believe they can run program	50	3.50	.61
staff believe in effectiveness of program	50	3.58	.54
staff adhere to planned program intervention	50	3.30	.71
staff is warm genuine	50	3.66	.52
staff committed to human service	50	3.68	.51
effective reinforcement	50	3.60	.57
effective modeling	50	3.68	.51
effective disapproval	50	3.56	.58
problem solving technique	50	3.68	.47
authority used	50	3.60	.57
structured learning	50	3.56	.58
quality assurance	50	3.50	.58
offenders satisfaction	50	3.12	.69
assessment of target behavior	50	3.50	.58
reconviction data	50	3.62	.57
program evaluation	50	3.66	.52
programs methods documented	50	3.50	.58

Table I1:
 Descriptives for Individual ETEQ Items (Psychologists) (Continued)

	<u>N</u>	<u>M</u>	<u>SD</u>
journal published	49	2.39	.89
offenders records	50	3.60	.70
ethics documented	49	3.18	.91
no changes in program components	45	1.91	.82
no changes in program funding	43	2.26	1.00
changes in community support	45	2.09	.99
advisory board	49	2.33	.97

Table I2:
Descriptives for Individual ETEQ Items (Students)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
program director designing program	50	3.18	.77
program director is professionally trained	50	3.82	.39
program director experience	50	3.02	.79
program director select/train staff	50	3.08	.88
program director conducts some aspect	50	3.12	.69
literature search	50	3.12	.66
pilot program	48	3.29	.80
need for program documented	50	3.02	.74
values/goals documented	50	3.24	.72
cost effective	50	2.76	.66
program funding adequate	50	3.24	.59
treatment appropriate	50	3.80	.49
exclusion criteria	49	3.10	.74
risk factors	50	3.52	.61
risk factors assessed	50	3.14	.73
risk level defined	50	3.42	.73
criminogenic need	49	3.41	.73
criminogenic need assessed	49	3.24	.66
criminogenic need identified	49	3.57	.58
changes assessed	50	3.48	.65
responsivity factors	50	3.34	.75
responsivity assessed	50	3.06	.79
responsivity factors specified	49	3.00	.76
target criminogenic behavior and attitudes	49	3.44	.68
cognitive/behavioral treatment	50	3.58	.64

Table 12:
Descriptives for Individual ETEQ Items (Students) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
separated from the rest	50	2.54	.97
manual	50	2.76	.82
time spent	50	3.04	.70
intensity/ duration	50	3.38	.75
treatment offender match	50	3.66	.52
staff treatment match	50	3.46	.65
staff offenders match	50	3.48	.68
motivation for treatment	49	3.35	.56
offenders input	50	2.60	.88
rewards/punishers	50	2.72	.88
theory of punishment	48	2.96	.87
effective punishing stimuli	50	3.08	.78
administration of punishing stimuli	50	3.16	.79
emotional reaction to stimuli	50	3.20	.73
completion criteria	50	3.34	.63
anticipate problem situation	50	3.34	.72
rehearse sociable acceptable responses	50	3.24	.77
practice socially acceptable	50	3.28	.86
referred to other services	50	3.54	.65
family trained	50	3.32	.82
booster	50	3.08	.90
therapist speak	50	2.30	.81
therapist follow up	50	3.14	.81
small treatment groups	50	3.30	.71
favorable condition for delivery	50	2.98	.74
productive therapy sessions	50	3.60	.57
education staff	50	2.98	.82

Table 12:
Descriptives for Individual ETEQ Items (Students) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
staff training	50	3.40	.70
staff trained in corrections	50	3.58	.50
staff experience	50	3.22	.79
staff qualities	50	3.44	.61
staff turnover	49	2.98	.90
staff supervision	50	3.26	.75
staff trained in intervention used	50	3.50	.68
staff has input	50	3.24	.69
staff believe they can run program	49	3.51	.54
staff believe in effectiveness of program	50	3.50	.61
staff adhere to planned program intervention	50	3.20	.67
staff is warm genuine	50	3.60	.61
staff committed to human service	50	3.64	.60
effective reinforcement	50	3.54	.65
effective modeling	50	3.48	.61
effective disapproval	50	3.42	.67
problem solving technique	50	3.56	.58
authority used	50	3.50	.71
structured learning	50	3.38	.73
quality assurance	50	3.40	.64
offenders satisfaction	50	3.00	.90
assessment of target behavior	50	3.34	.56
reconviction data	50	3.52	.68
program evaluation	50	3.68	.51
programs methods documented	50	3.58	.50
journal published	50	2.70	.97
offenders records	50	3.56	.88

Table I2:
Descriptives for Individual ETEQ Items (Students) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
ethics documented	50	3.28	.67
no changes in program components	49	2.20	.94
no changes in program funding	49	2.33	.99
changes in community support	49	2.59	.91
advisory board	49	3.16	.75

Table I3:
Descriptives for Individual ETEQ Items (Treatment Delivery)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
prdirector desinging program	34	2.59	1.21
program director is professionally trained	34	3.26	.93
program director experience	34	3.03	.99
program director select/train staff	34	2.94	.92
program director conducts some aspect	34	2.50	1.08
literature search	34	3.32	.88
pilot program	34	3.53	.79
need for program documented	34	3.62	.60
values/goals documented	34	3.68	.59
cost effective	34	3.12	.88
program funding adequate	34	3.53	.71
treatment appropriate	33	3.79	.60
exclusion criteria	34	3.53	.86
risk factors	34	3.47	.90
risk factors assessed	34	3.00	1.02
risk level defined	34	3.24	.82
criminogenic need	34	3.68	.59
criminogenic need assessed	34	3.29	.97
criminogenic need identified	34	3.65	.60
changes assessed	34	3.44	.70
responsivity factors	34	3.26	.83
responsivity assessed	34	2.82	.94
responsivity factors specified	34	3.21	.84
target criminogenic behavior and attitudes	33	3.70	.64
cognitive/behavioral treatment	33	3.85	.44

Table I3:
Descriptives for Individual ETEQ Items (Treatment Delivery) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
separated from the rest	34	1.74	.99
manual	34	3.32	.88
time spent	34	2.82	.80
intensity/ duration	33	3.27	.84
treatment offender match	34	3.38	.82
staff treatment match	33	3.45	.67
staff offenders match	33	3.09	.80
motivation for treatment	34	3.12	.88
offenders input	34	2.85	.89
rewards/punishers	33	3.09	.95
theory of punishment	33	2.79	1.08
effective punishing stimuli	33	2.73	1.13
administration of punishing stimuli	31	2.26	1.12
emotional reaction to stimuli	31	2.61	1.17
completion criteria	34	3.65	.77
anticipate problem situation	33	3.61	.75
rehearse sociable acceptable responses	33	3.67	.74
practice socially acceptable	33	3.55	.79
referred to her services	33	3.64	.65
family trained	31	2.84	1.07
booster	32	2.97	.93
therapist speak	32	1.84	.85
therapist follow up	31	2.64	1.05
small treatment groups	32	3.50	.67
favorable condition for delivery	32	3.66	.60
productive therapy sessions	33	3.61	.70
education staff	33	2.36	1.22

Table I3:
Descriptives for Individual ETEQ Items (Treatment Delivery) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
staff training	34	2.62	1.21
staff trained in corrections	34	2.91	1.14
staff experience	34	2.88	.91
staff qualities	33	3.45	.83
staff turnover	33	3.18	.85
staff supervision	34	3.03	.99
staff trained in intervention used	34	3.74	.57
staff has input	32	2.84	1.19
staff believe they can run program	34	3.68	.73
staff believe in effectiveness of program	34	3.74	.57
staff adhere to planned program intervention	34	3.59	.70
staff is warm genuine	34	3.79	.48
staff committed to human service	34	3.79	.48
effective reinforcement	34	3.71	.52
effective modeling	34	3.68	.68
effective disapproval	34	3.62	.60
problem solving technique	34	3.68	.53
authority used	34	3.68	.64
structured learning	33	3.45	.83
quality assurance	34	3.32	.81
offenders satisfaction	33	3.52	.62
assessment of target behavior	32	3.28	.77
reconviction data	33	3.36	.82
program evaluation	33	3.55	.75
programs methods documented	32	3.50	.76
journal published	33	2.67	1.16
offenders records	32	3.59	.91

Table I3:
 Descriptives for Individual ETEQ Items (Treatment Delivery) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
ethics documented	30	3.17	.91
no changes in program components	32	2.09	1.06
no changes in program funding	31	2.45	1.09
changes in community support	31	2.39	.99
advisory board	32	2.25	1.02

Table I4:
Descriptives for Individual ETEQ Items (Total)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
prdirector desinging program	133	2.92	.96
program director is professionally trained	133	3.56	.70
program director experience	133	3.05	.85
program director select/train staff	133	3.02	.86
program director conducts some aspect	133	2.89	.89
literature search	134	3.37	.72
pilot program	132	3.34	.82
need for program documented	134	3.31	.74
values/goals documented	133	3.36	.70
cost effective	134	2.80	.73
program funding adequate	134	3.34	.66
treatment appropriate	133	3.80	.50
exclusion criteria	133	3.40	.75
risk factors	134	3.53	.71
risk factors assessed	134	3.16	.83
risk level defined	134	3.35	.74
criminogenic need	133	3.56	.63
criminogenic need assessed	133	3.30	.80
criminogenic need identified	133	3.61	.59
changes assessed	134	3.41	.70
responsivity factors	134	3.30	.75
responsivity assessed	134	2.91	.87
responsivity factors specified	132	3.13	.79
target criminogenic behavior and attitudes	132	3.62	.60
cognitive/behavioral treatment	133	3.70	.58

Table I4:
Descriptives for Individual ETEQ Items (Total) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
separated from the rest	134	2.22	1.06
manual	134	3.02	.85
time spent	134	2.93	.80
intensity/ duration	133	3.31	.81
treatment offender match	134	3.48	.67
matching: staff treatment	133	3.48	.61
matching: staff offender	133	3.30	.74
motivation for treatment	133	3.12	.76
offenders input	134	2.62	.89
rewards/punishers	131	2.99	.93
theory of punishment	128	2.79	.96
effective punishing stimuli	130	2.84	.93
administration of punishing stimuli	128	2.74	.99
emotional reaction to stimuli	128	2.93	.94
completion criteria	134	3.50	.68
anticipate problem situation	133	3.48	.72
rehearse sociable acceptable responses	133	3.50	.70
practice socially acceptable	133	3.45	.76
referred to other services	132	3.62	.60
family trained	131	2.98	.93
booster	132	3.05	.92
therapist speak	131	2.00	.89
therapist follow up	129	2.88	.93
small treatment groups	132	3.30	.71
favorable condition for delivery	131	3.33	.74
productive therapy sessions	131	3.55	.65
education staff	133	2.78	1.02

Table 14:
Descriptives for Individual ETEQ Items (Total) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
staff training	134	3.10	.94
staff trained in corrections	134	3.28	.85
staff experience	134	3.05	.84
staff qualities	132	3.47	.67
staff turnover	132	3.15	.80
staff supervision	134	3.27	.82
staff trained in intervention used	134	3.61	.60
staff has input	131	3.14	.86
staff believe they can run program	133	3.55	.62
staff believe in effectiveness of program	134	3.59	.58
staff adhere to planned program intervention	134	3.34	.70
staff is warm genuine	134	3.67	.54
staff committed to human service	134	3.69	.54
effective reinforcement	134	3.60	.59
effective modeling	134	3.60	.60
effective disapproval	134	3.52	.62
problem solving technique	134	3.63	.53
authority used	134	3.58	.64
structured learning	133	3.47	.70
quality assurance	134	3.42	.66
offenders satisfaction	133	3.17	.78
assessment of target behavior	132	3.39	.63
reconviction data	133	3.52	.68
program evaluation	133	3.64	.58
programs methods documented	132	3.53	.60
journal published	132	2.58	.99
offenders records	132	3.58	.82

Table I4:
Descriptives for Individual ETEQ Items (Total) (Continued)

	<u>N</u>	<u>Mean</u>	<u>SD</u>
ethics documented	129	3.22	.82
no changes in program components	126	2.07	.93
no changes in program funding	123	2.33	1.01
changes in community support	125	2.36	.98
advisory board	130	2.62	.99

Appendix J

Meta-analysis- CPAI Scoring Manual

PROGRAM DEMOGRAPHICS

1. Article ID _____ ID.
2. Appropriateness of intervention STUD
 - 1 Justice
 - 2 Inappropriate
 - 3 Unspecified
 - 4 Appropriate
3. From the list below check the single best summary category: (Note: Try not to use the category "other" unless absolutely necessary since it will not add much information in the analysis).

INTERV

JUSTICE INTERVENTIONS. Treatment provided by a justice agency in conjunction with informal or formal probation, parole, institutionalization or court order

PROBATION

- 1 Probation, regular (e.g. in comparison to something else)
- 2 Probation, reduced caseload, intensive supervision
- 3 Probation, restitution (money, labour, etc.)
- 4 Probation, release; waive probation
- 5 Probation, additional counseling
- 6 Probation, other enhancement
- 7 Probation, other

PAROLE

- 8 Parole, regular (e.g. in comparison to something else)
- 9 Parole, reduced caseload, intensive supervision
- 10 Parole, restitution (money, labour, etc.)
- 11 Parole, release; waive probation
- 12 Parole, additional counseling
- 13 Parole, other enhancement
- 14 Parole, Other

INSITUTIONALIZATION

- 15 Institutionalization, regular (compared to something else)
- 16 Institutionalization, community residential
- 17 Institutionalization, restitution (money, labour, etc)
- 18 Institutionalization, additional counselling
- 19 Institutionalization, other enhancement
- 20 Institutionalization, other

DETERRENCE/SCHOCK

- 21 Scared straight; prison visit/buddy system
- 22 Short term “ shock” incarceration
- 23 Other deterrence/shock
- 24 Any other justice intervention

NON-JUSTICE INTERVENTIONS. Treatments that are not provided in conjunction with justice supervision even if the clients are referred by a justice agency or the treatment is provided by a justice agency as a separate service to clients not under justice authority, e.g. counseling program for non-offenders run by the probation department

RESIDENTIAL

- 25 Teaching family home (as a total program)
- 26 Wilderness camp, etc (residential, not time limited)
- 27 Token economy; behavioral contingency system
- 28 Guided group interaction/positive peer culture
- 29 Milieu therapy
- 30 Additional counseling
- 31 Other enhancement
- 32 Other, residential

SCHOOL BASED

- 33 Special classes; continuation school
- 34 Extra tutoring
- 35 Token economy; behavioral contingency system
- 36 Special training, etc for teachers; school personnel
- 37 Additional counseling
- 38 Other enhancement
- 39 Other school based

COUNSELLING/PSYCHOTHERAPY (freestanding, i.e., non-residential and non-school based; talk/insight oriented; behaviorally oriented goes below).

- 40 individual counselling
- 41 group counselling (with other clients)
- 42 family counselling
- 43 parent/spouse counselling (without client)
- 44 casework (other than within justice system)
- 45 other, counselling/psychotherapy

BEHAVIORALLY ORIENTED other than skills oriented (distinguished by behavioral jargon, concepts, etc.) freestanding, i.e. Non-residential and non-school based.

- 46 behavioral contracting/contingency management
- 47 cognitive behavioral (desensitization, modelling, expectancy)
- 48 shaping of specific behaviors; operant conditioning
- 49 token economy, not part of residential or school program
- 50 restitution, non-justice system
- 51 other, behaviorally oriented

SKILL ORIENTED (freestanding, i.e., non-residential and non-school based)

- 52 academic education/tutoring (not school based)
- 53 moral education, training, etc.
- 54 employment-working at job
- 55 employment-training- learning job content
- 56 vocational-job finding- interview skills, career planning
- 57 interpersonal, e.g. role playing for social skills, assertion
- 58 recreational, sports, fitness
- 59 wilderness/outward bound; survival training
- 60 abstinence-drugs, alcohol (education, not counselling)
- 61 parent/ spouse training (without client)
- 62 other, skill oriented

SYSTEM ORIENTED, non-residential and not school based

- 63 advocacy on behalf of youth
- 64 personnel training (teachers, employers, justice agents)
- 65 other, system oriented

MULTIPLE SERVICES, non-residential and not school based

- 66 service brokerage (client sent to number of services)
- 67 multimodal (every client receives more than one service)
- 68 other, multiple services
- 69 any other non-justice intervention

4. How many subjects in treatment group. NTRE

5. How many subjects in control group NCON

6. How many subjects in TREATMENT/CONTROL group NPOOL

7. Type of setting SET

- 1 Non system Diversion
- 2 System Diversion
- 3 Probation/Parole/Community
- 4 Institutional/Residential
- 5 SP

8. Source of clients/subjects for treatment: (Note: The issue here is who took the initiative in identifying or choosing subjects for the treatment e.g., were they identified by clinic or by researchers using the client's records etc.) TREATGR

- 1 sought treatment voluntarily ("self-referral", walk-in)
- 2 referred/identified by family, friends
- 3 referred/identified by non-criminal justice community agency(schools, job, mental health, etc.)
- 4 referred/identified by criminal justice agency but "voluntary" (e.g. via police, pro-

bation, court, etc.)

- 5 referred identified by criminal justice agency but participation mandated (e.g. by court, terms of probation, institution). Assume it is mandatory if it is a CJS agency unless there is specific information that it is voluntary. Don't override a specific statement that it is voluntary even if you presume there is some coercion.
- 6 referred/identified by multiple sources, none predominates
- 7 solicited or arranged by researcher
- 8 other

9. What the control group receives: Note: The difference between receives nothing and treated as usual hinges on whether or not two groups have an institutional framework or experience in common. CONTROL

- 1 receives nothing; no evidence of any treatment or attention; may still be participating in activities that are incidental to the treatment strategy or client population as defined.
- 2 wait list, delayed treatment control, etc.; contact limited to application, screening pretest, posttest etc.
- 3 minimal contact; instructions, intake interview, etc. but not wait listed
- 4 treated as usual, routine care, (Note: Treated as usual refers to: control receives the usual treatment without the special enhancement that constitutes the treatment of interest; this refers to treatment occurring within a framework common to E and C groups with something added for the E group).
- 5 treated as usual, SCHOOL
- 6 treated as usual, PROBATION
- 7 treated as usual, INSTITUTIONALIZATION
- 8 other treatment as usual
- 9 attention placebo; e.g. C receives discussion, attention, or very diluted or less intensive version of treatment
- 10 treatment element placebo; C receives target treatment except for defined element presumed to be the crucial ingredient
- 11 alternate treatment; C is not really a "control", but another treatment (other than "usual" treatment) being compared with the focal treatment.

10. Is coercion a part of the referral CJREFER2

- 1 Referred by criminal justice system
- 2 Referred by criminal justice system & mandated

11. What is the general philosophical orientation/rationale/mission statement of the program? Level of program/treatment theoretical development: choose the higher end of the scale if you can give the researchers the benefit of the doubts in order to get variability on this item. This is not the same as implementation detail: the nature of the treatment concepts, irrespective of how implemented is the key. Presume it is a treatment strategy unless otherwise stated. Stay

with what is presented in the article. Don't assume that more theorizing went on than got into print. THEORY SOMETHEO 0,1 (1,2—3,4,5)

- 1 Atheoretical (black box). No rationale for treatment concept or why it is supposed to work: described by minimal operational statements or program/treatment label only
- 2 Treatment strategy. Some information on nature of treatment to be delivered, program tactics or intent, but little explanation for why those tactics should lead to desired outcome, e.g. provide supportive counseling; give experience in self-reliance. Focus and some detail on what treatment provides/does, not so much on how or why subject is expected to respond or improve.
- 3 Treatment rationale. General statement about why treatment supposed to work but not as specific as describing causal mechanism, intervening variables, etc. Belief that a warm relationship helps one to learn self-control. More emphasis on how and why subject expected to respond than above, but not fully conceptualized, measured or tested. Provides some underlying psychological or sociological processes within the recipient or client, but not as explicit theory.
- 4 Hypothesis testing? Propositions having to do with the relationship between treatment and outcome are specified and tested on the data; may be somewhat ad hoc, specific to the particular study context and its data (tests whether degree of attitude change on an intervening variable correlates with recidivism measure subsequent to treatment). Stays close to the data and variables involved. Not really a complete a priori theory but does test some theoretical/conceptual propositions.
- 5 Integrated theory. General proposition(s) about the causal mechanism that connects treatment to expected outcome; specifiable a priori and generally coherently related to program design research plan and analysis. May develop own theory or put study within existing broad theory (e.g. social learning). E.g. links role-playing treatment with empathy and delinquency outcome measures. This involves broad, well thought-out concepts from which the treatment was developed.

12. Treatment etiological orientation: Enter value of best judgement. Note: This will rarely be stated overtly: must be inferred from nature of treatment and descriptions given of subjects, treatment issues, rationale, etc. TREATOR PICRTHEO 0,1 (5,6,7-1,2,3,4)

- 1 Personal. The cause of the delinquent behavior is due to some personal characteristic of the individual, e.g. disease, personality problem, psychological disturbance, low IQ, lack of self control, etc. Treatment focuses on remedying the problem characteristic Note Use this category only when no elements other than personal are detected; if personal with other elements, use the "Personal, Mixed" category.
- 2 Personal, mixed. Mostly personal orientation as for item above, but some elements of one or more of the other categories e.g. an individualistic therapy orientation for which some attention to family issues, peers, etc. is indicated.
- 3 Interpersonal. Delinquent behavior is due to interpersonal problems that are endemic to immediate social surroundings, e.g. family, peers, etc.
- 4 Structural/cultural/economic. Problem behavior is due to sociological problems that are endemic to delinquents' general social conditions, e.g. lack of economic opportunity, lower class norms, etc. Treatment may still focus on the individual, but with regard to coping with the sociological condition at issue, e.g. teaching job

finding skills. Alternatively.

- 5 Labeling. Delinquency is a matter of getting caught and being labeled an offender. Otherwise, offenders are assumed to be very similar to non-offenders. Treatment may focus on helping the client handle the negative impact of the label, altering the labeling or the system that labels, etc.
- 6 Social, Mixed. There are some elements of interpersonal, structural and labeling perspectives; no strong leanings to individual treatment
- 7 Other, specify _____

PROGRAM IMPLEMENTATION

Program Director is defined as the individual currently responsible for all aspects of the clinical services delivered to the offenders or the person responsible for the overall management of the program

13. Is program directors the sole originator of the program or was instrumental in helping set up the program INITIA

- 0 No
- 1 Yes

14. Is program director professionally trained (university degree) in one of the helping professions (e.g., education, nursing, psychology, social work) QUALIF

- 0 No
- 1 Yes

15. Does program director has experience (management or delivery) with an offender treatment program of any description.? PEXPER

- 0 No
- 1 Yes

16. Is program director directly involved in hiring and providing clinical training to staff? HIRTRA

- 0 No
- 1 Yes

17. Is program director involved in conducting some aspects of the program that involves a) direct service delivery to client, or b) direct supervision of staff? CONDUCT

- 0 No
- 1 Yes

18. Does article has an extensive literature search conducted? LITER

- 0 Literature search was not conducted or included only few articles/books that are not major works in the field
- 1 Literature search was conducted that included the major work in the field and/or some review of treatment effectiveness literature was conducted. (Note: For the articles that are published prior to 1980 look for an attempt to cover available liter-

ature in the filed at that time)

19. Before formal program began, is pilot program undertaken to sort out program logistics and content? PILOT.

- 0 Pilot program was not conducted at all
- 1 Pilot program was conducted

20. The need for the program was assessed (e.g., surveys, focus groups stakeholders, management, staff) NEED

- 0 No
- 1 Yes

21. Are values and goals of program consistent with existing values and practices of management, staff, and stakeholders? VALUE

- 0 Not consistent
- 1 Consistent

22. Is the program perceived as cost-effective* by stakeholders, management, staff and/or community (*to be less costly than other alternatives).COST

- 0 No
- 1 Yes

23. Funding is considered adequate for the task (operational and personnel). FUND

- 0 No
- 1 Yes

CLIENT PRE-SERVICE ASSESSMENT

24. The type of client presently received by the program is appropriate. If there is a statement that certain number of clients was included even though the program is not considered to be adequate by the programmers, then item will be scored as "0". All other cases will be scored as "1". CLIENT

- 0 No
- 1 Yes

25. There is a rational clinical or legal basis for the exclusion of certain types of clients from program participation. EXCL

- 0 No exclusion criteria reported or exclusion criteria without rational clinical/legal basis
- 1 Exclusion criteria with rational clinical/legal basis

26. Risk factors (from the perspective of the programmers) were taken into account. Please note the factors nominated do not necessary have to be those with the strongest predictive validity. If the program took into account previous criminal history, i.e. program was done on repeat offenders, it should be scored that risk factors were surveyed. RASSESS

- 0 No
- 1 Yes

27. Are risk factors measured by a psychometric scale/test (e.g., LSI-R)? File records and clinical notes are acceptable if the relevant information is quantified. RMETHOD

- 0 Risk factors are not scored at all
- 1 Risk factors are scored with a psychometric scale/test or from clinical notes that have relevant information quantified.

28. Clients personal characteristics are summarized as to level of risk - either qualitatively (high-medium-low) or an actual score such as LSI-R of 24+ = high risk or Wisconsin of 2= low risk. RLEVEL

- 0 No
- 1 Yes

29. Risk factors are taken into account in assigning offenders to particular treatment group. ASSRISK

- 0 No
- 1 Yes

30. Need factors (i.e., dynamic attributes of offenders and their circumstances that are linked with recidivism, anti-social attitudes) were taken into account. Please note the factors nominated do not necessary have to be those with the strongest predictive validity SNEED

- 0 No
- 1 Yes

31. Are needs factors measured by a psychometric scale/test? File records and clinical notes are acceptable if the relevant information is quantified. NMETH

- 0 Need factors are not scored at all
- 1 Need factors are scored with a psychometric scale/test or from clinical notes that have relevant information quantified.

32. Clients personal characteristics are summarized as to level of needs - either qualitatively (high-medium-low) or an actual score such as LSI-R of 24+ = high risk or Wisconsin of 2= low risk. NLEVEL

- 0 No
- 1 Yes

33. Need factors are taken into account in assigning offenders in different treatment/control group ASSNEED

- 0 No
- 1 Yes

34. The responsivity of offenders to different styles and modes of service. How characteristics of offenders may interact with style and mode of service. Examples: low conceptual level offenders respond better to higher levels of program structure, higher anxiety offenders respond poorly to confrontation, highly verbal approaches would be inappropriate with those of low verbal ability. Score if any of responsivity characteristics has been taken into account by programmers. * List of potential specific responsivity variables - I Level, Conceptual Level, Psychopathy, Intelligence, Motivation, Age, Gender, Ethnicity, Other). Also, general responsivity factors were taken into account. Example: if programs are using behavioral intervention then they address general responsivity principles. RESPON

- 0 None of responsivity characteristics was considered
- 1 At least one of responsivity characteristics was considered.

35. Responsivity factors as assessed by some psychometric instrument (if applicable) e.g. aptitude test (General Aptitude Test Battery), PCL-R Psychopathy, Beck (depression) (Note: this is applicable for general responsivity principle). PSYHCO

- 0 No
- 1 Yes

36. Client's personal characteristics are summarized as to level of responsivity characteristics - either qualitatively (e.g. psychopaths./nonpsychopaths) or an actual score such as PCL-R of 30 or more for psychopaths (Note: for special responsivity characteristics). RSLEVE

- 0 No
- 1 Yes, please specify how responsivity characteristics were summarized

37. Were responsivity characteristics a factor in assignment of offenders to particular treatment group? (For special responsivity characteristics) ASSRESPN

- 0 No
- 1 Yes

PROGRAM CHARACTERISTICS

38. Program targets criminogenic needs and the number of targeted criminogenic needs exceeds number of targeted noncriminogenic needs. CRIMNON2

- 0 No
- 1 Yes

39. Type of behavioral intervention used? TYPEBEH (0,3)
BEH (0,1)

- 0 No
- 1 Yes

40. In a prison, program participants are separated from the rest of the population unless the entire institution is involved in the program. For a community-based program the whereabouts of clients and peer associations are closely monitored by staff or other means LOCAT

- 0 No. Program participants are not separated from the rest of prison population (institution) or whereabouts of participants and peer associations are not closely monitored (community)
- 1 Yes. Program participants are separated from the rest of prison population (institution) or whereabouts of participants and peer associations are closely monitored (community).

41. Does the program have a printed manual? ANYPRINT(O,1)

- 0 No
- 1 Yes

42. How were the sessions distributed? DISTRIB (1-5) DISTRIB2 (0,1)

- 1 Minimum dose for effectiveness is unspecified
- 2 Minimum dose specified, but number of sessions completed unavailable.
- 3 Approximately half of subjects received minimum dose
- 4 Most of subjects received the minimum dose (60-80%)
- 5 Almost all subjects received the minimum dose (81-100%)

43. The higher risk clients receive the highest intensity/ duration of service. NEWRISK2

- 0 No everyone receive the same treatment
- 1 The higher risk clients receive the highest intensity and/or duration of service

44. Participants are assigned to programs that mach up best with their interests and style of learning. CLIPRO

- 0 No
- 1 Yes

45. Is staff assigned to a program their skills match up with best (e.g., staff who function best with higher degree of structure work in a contingency management program). STPROG

- 0 No
- 1 Yes

46. Is staff assigned to clients they can work with effectively (e.g. high conceptual level staff assigned to like clients) STACLI

- 0 No
- 1 Yes

47. There is clearly outlined completion criteria as to when program terminates for each client. Adequate completion criteria will be defined by progress in acquiring pro-social behaviors, beliefs while in program or engaging in behavior that seriously jeopardized the safety of staff and other clients. COMPLE

- 0 No
- 1 Yes

48. Client can modify some features of rules and regulations of the program MODIFY

- 0 No
- 1 Yes

49. Theory supporting the type of punishment used is described (e.g., negative law of effect, competing response theory). Effective punishing stimuli are used in the program (e.g. electric shock, drug induced aversion, mild aversive stimuli, covert sensitization, overcorrecting, time out, and response cost) ACCTDISA(0,1)

- 0 No
- 1 Yes

Note:

Following four items pertain to relapse prevention strategies, which are commonly applied to substance abusers and sex offenders. Please note #53-55 is infrequently found in prison based setting.

50. Insuring that the Client is able to Recognize Risky Situations, and has a Concrete and Well Rehearsed Plan for Dealing with those Situations: this involves teaching the offender to identify situations which are conducive to criminal activity or responses and teaching the offenders how to avoid them. It may also involve teaching the offender how to deal with these situations if they are caught in one. These types of models are usually based on relapse prevention. RISKY.A

- 0 No
- 1 Yes

51. Client plans and rehearses alternative prosocial responses and practices new prosocial behaviors in increasingly difficult situations. ACCTSLP.A

- 0 No
- 1 Yes

52. Is there evidence of brokerage by the therapist (job referrals, medical referrals, seeking out services, arranging services BROKER2(0,1)

- 0 No
- 1 Yes

53. Are significant others (e.g., family, friends) trained to provide support. SIGN.

- 0 No
- 1 Yes

54. Does client return to relearn/reinforce skills originally taught in formal treatment phase of program? BOOST

- 0 No
- 1 Yes

STAFF CHARACTERISTICS

55. The following items refer to full/part time staff (excluding the program director) Majority of the staff with an undergraduate degree and/or an advanced (MA) degree. EDUCAT.

- 0 No
- 1 Yes

56. Majority of staff that has training in criminal justice, education, nursing, psychology, social work or specialized fields (e.g., addictions). ARSTUD

- 0 No
- 1 Yes

57. Majority of staff has previous working experience in treatment programs with offenders? STAEX

- 0 No
- 1 Yes

58. When hiring staff, where any of personality characteristics taken into account (e.g. sense of humor, flexible, warm, empathetic). RELFACTR(0,1)

- 0 No
- 1 Yes

59. When hiring staff, where any of skill factors taken into account SKILLFAC(0,1)

- 0 No
- 1 Yes

60. Is staff assessed on clinical skills related to service delivery and/or do they receive regular clinical supervision. SUPBTRS2(0,1)

- 0 No
- 1 Yes

61. Therapist are trained in the model applied. THERTRA2 (0,1)

- 0 No
- 1 Yes

62. Is staff able to modify program structure with approval by program supervisor/review board? INPUT

- 0 No
- 1 Yes

63. Majority (over 50%) of the staff/management have remained on the job for at least two years. OVERT

- 0 No
- 1 Yes

EVALUATION

64. Is a management audit system mechanism in place that consists of a) within program checks, file review, problem oriented records, etc., that monitor treatment progress, b) clinical supervision and c) client feedback .AUDIT

- 0 No
- 1 Yes

65. Are clients surveyed as to satisfaction with services SATIS

- 0 No
- 1 Yes

66. Is there periodic, objective, standardized assessment of clients on program target behaviors. To score this item tangible evidence in files (e.g. change scores on risk /need measure) must be present)MONITOR2(0,1)

- 0 No
- 1 Yes

67. Does researcher/evaluators contribute to the design and management of the services?EVALUIN2(0,1).

- 0 No
- 1 Yes

68. Are client re-arrest, reconviction, or reincarceration data gathered at 6 months or more after leaving program? RECID.

- 0 No
- 1 Yes

69. As above, with at least one evaluation in last 5 years, comparing treatment outcome with a risk control comparison group of some sort. COMPAR

- 0 No evaluation with not comparison group
- 1 Yes evaluation with comparison group

70. Is there a document that containing introduction, method, results, discussion on file that details effectiveness of program available. DOCU

- 0 No
- 1 Yes

71. Is there a document containing introduction, method, results, discussion on file that details effectiveness of program available that is published in an edited journal. PUBLI

- 0 No
- 1 Yes

OTHER

72. Client records are kept in confidential file (in order to score yes records must have social history, record of presenting problem, assessment data, program progress notes, etc) CLREC

- 0 No
- 1 Yes

73. There is documented code of ethics of intervention (e.g., least intrusive intervention) ETHICS

- 0 No
- 1 Yes

74. Is there a noticeable change in program components in last 2 years that jeopardized smooth functioning of the program? CHCOM

- 0 No
- 1 Yes

75. Is there noticeable changes in program funding in last 2 years that jeopardized smooth functioning of the program CHFUN

- 0 No
- 1 Yes

76. Is there a noticeable change in community support for program in last 2 years that jeopardized smooth functioning of the program? CHSUP

- 0 No
- 1 Yes

77. Does the program have a group of individuals officially designated to advise program (e.g. Board of directors) BOARD

- 0 No
- 1 Yes

Appendix K

Values for Eta for Individual CPAI Items

CPAI item	<u>%</u>	<u>eta</u>
Program director- initiated	1.1	.13*
Program director-qualification	1.5	.11
Program director-experience	.8	.02
Program director-hiring/training	.4	.01
Program director-supervision	.8	.03
Literature review	10.9	.13*
Pilot project	10.2	.11
Need for program	6.8	-.01
Value consistency	1.5	-.06
Cost effectiveness	4.5	.04
Guarantied budget	.4	-.04
Appropriate Clients	86.1	.30**
Exclusion Criteria	30.8	.06
Risk Assessment	80.1	.18**
Method of risk assessment	3.4	.07
Risk level summarizes	9.4	.01
Assigned based on risk	75.6	.14**
Needs assessment	57.5	.44**
Method of needs assessment	3.8	.04
Needs level summarizes	2.6	.06
Assigned based on need	23.7	.32**
Responsivity Assessment	30.1	.30**
Method of Responsivity assessment	3.4	-.08
Level of responsivity	3.0	-.08

*p<.05

**p<.01

CPAI item	%	eta
Assigned based on responsivity	4.5	.24**
Program Targets	60.9	.48**
Type of intervention	28.2	.33**
Location	29.3	.01
Printed Manual	25.9	.25**
Intensity	37.6	.02
Duration	.4	.06
Matching Client/Program	4.9	.08
Matching Staff/Program	1.5	.23**
Matching Staff/Client	3.0	.21**
Disapproval	2.6	.16**
Relapse Prevention Recognize/Rehearse	6.0	.12**
Relapse Prevention Plan/Practice	12.4	.32**
Brokerage	9.8	.15*
Significant Others	17.7	.18*
Booster sessions	5.3	.01
Staff Education	13.2	.07
Staff training	12.8	.10
Staff Experience	6.4	.17**
Relationship factors	4.5	.28**
Skills factors	15.0	.30**
Supervised staff	19.9	.20**
Trained Therapists	50.0	.25**
Staff input	3.4	.01
Staff turnover	1.1	-.04
Quality Assurance	2.6	.21

* $p < .05$

** $p < .01$

CPAI item	<u>%</u>	<u>eta</u>
Client satisfaction	7.1	.08
Evaluator Involved	27.1	.36**

* $p < .05$

** $p < .01$

Appendix L

Ranges within the CPAI Sections Scores (Meta-analysis)

CPAI sections	Treatment group		Control group	
	Min	Max	Min	Max
Program Implementation	.00	.36	.00	.27
Pre-service Assessment	.00	.79	.00	.64
Program Characteristics	.00	.53	.00	.29
Staff Characteristics	.00	.78	.00	.56
Program Evaluation	.50	.88	.50	.88
Other	.00	.00	.00	.00
Treatment Score	.00	.58	.00	.45
Total Score	.08	.40	.06	.45

Appendix M

Informed Consent - (CPAIQ)

Dear participant:

The purpose of an informed consent is to insure that you understand the purpose of the study and the nature of your involvement. The informed consent must provide sufficient information in such a way that you have the opportunity to determine whether you wish to participate in the study.

This study is being conducted by Aleksandra Milovic as part of her Doctoral thesis at Carleton University and was approved by the Correctional Service of Canada/Solicitor General of Ontario. The purpose of this research is to evaluate correctional treatment programs using the questionnaire (CPAIQ) that was adapted from the Correctional Program Assessment Inventory (CPAI). Your program has been randomly selected among federal/provincial correctional programs and your name has been obtained from the Correctional Service of Canada/Solicitor General of Ontario.

This study will require your participation in two phases. In the first phase, you will be asked to answer questions on the CPAIQ. These questions are pertaining to the implementation, delivery and evaluation of the correctional program that you are managing. Some questions will require only yes or no answers, others will require more specific program details. Please elaborate as much as you can regarding the program details. In the second phase, you will be contacted again by the principal investigator to set the time and date for the interview that will be audiotaped. In this interview you will be asked to answer detailed questions about the correctional program that you are managing.

The completion of the questionnaire and the participation in the interview is voluntary. The completion of the questionnaire is expected to take 45 to 60 minutes of your time. The duration of the interview is estimated to be 2 hours. Your name and your answers will be seen only by researchers associated with this project and they will be kept strictly confidential. All information will be coded and analyzed in such way that your name will not be used.

You may withdraw from the study at any time. Please use the enclosed, self-addressed stamped envelope to return the completed questionnaire. Your response within one-month period would be greatly appreciated.

If you have any questions or concerns regarding your participation or regarding the research project, you can contact the principal investigator Aleksandra Milovic at (613) 721-9387, or Dr. Don Andrews, Professor, Carleton University, at (613) 520-2600 ext. 2662. If you have any ethical concerns about this research then please contact Dr. M. Gick, Chair, Department of Psychology Ethics Committee, at (613) 520-2600, ext. 2664 or Dr. K. Matheson, Chair, Department of Psychology, 520-2600, ext. 2648.

Informed Consent Form

I have read above form and understand the condition of my participation. In addition, the research "Psychometric evaluation of the Correctional Program Assessment Inventory" was explained to me by the principal investigator Aleksandra Milovic. I agree to participate in both phases of this study (to complete the questionnaire and to have an interview that will be audiotaped) and I grant Aleksandra Milovic permission to use my answers in Her Doctoral dissertation.

Participant's name _____

Participant's signature _____

Date _____

Appendix N

Informed Consent (Student sample) - (ETEQ)

Dear Participant:

The purpose of an informed consent is to ensure that you understand the purpose of the study and the nature of your involvement. The informed consent must provide sufficient information such that you have the opportunity to determine whether you wish to participate in the study.

This study is being conducted by Aleksandra Milovic as part of her Doctoral thesis at Carleton University and was approved by the Ethics Committee of Carleton University. Your name has been obtained from the experiment sign-up sheets for this study. You will obtain one credit for the participation in this study.

The enclosed questionnaire consists of a list of treatment elements that are an integral part of many correctional programs and that are potentially related to its effectiveness. You will be asked to state for each individual treatment element your opinion about its potential importance for program effectiveness.

Completion of the questionnaire is voluntary and is expected to take 20 to 30 minutes of your time. Your name and answers will be seen only by researchers associated with this project and will be kept strictly confidential. All information will be coded and analyzed in such a way that your name will not be used.

You may withdraw from the study at any time. Please use the enclosed, self-addressed stamped envelope to return your completed questionnaire. Your response within one-month period would be greatly appreciated.

If you have any questions or concerns regarding your participation or the research project, you can contact the principal investigator Aleksandra Milovic at (613) 721-9387, or Dr. Don Andrews, Professor, Carleton University, at (613) 520-2600 ext. 2662. If you have any ethical concerns about this research then please contact Dr. M. Gick, Chair, Department of Psychology Ethics Committee, at (613) 520-2600, ext. 2664 or Dr. K. Matheson, Chair, Department of Psychology, 520-2600, ext. 2648.

Appendix O

Informed Consent (Psychologist and Treatment Staff Sample) - (ETEQ)

Dear Participant:

The purpose of an informed consent is to ensure that you understand the purpose of the study and the nature of your involvement. The informed consent must provide sufficient information such that you have the opportunity to determine whether you wish to participate in the study.

This study is being conducted by Aleksandra Milovic as part of her Doctoral thesis at Carleton University and was approved by the Correctional Service of Canada and the Solicitor General of Ontario. Your name has been randomly selected from the list of personnel that was obtained from the Correctional Service of Canada and the Solicitor General of Ontario.

The enclosed questionnaire consists of a list of treatment elements that are an integral part of many correctional programs (both institutional and community) and that are potentially related to their effectiveness (i.e. programs that reduce reoffending). You will be asked to state for each individual treatment element your opinion about its potential importance for program effectiveness. Completion of the questionnaire is voluntary and is expected to take 20 to 30 minutes of your time. Your name and answers will be seen only by researchers associated with this project and will be kept strictly confidential. All information will be coded and analyzed in such a way that your name will not be used.

You may withdraw from the study at any time. Please use the enclosed, self-addressed stamped envelope to return your completed questionnaire. Your response within one-month period would be greatly appreciated.

If you have any questions or concerns regarding your participation or the research project, you can contact the principal investigator Aleksandra Milovic at (613) 721-9387, or Dr. Don Andrews, Professor, Carleton University, at (613) 520-2600 ext. 2662. If you have any ethical concerns about this research then please contact Dr. M. Gick, Chair, Department of Psychology Ethics Committee, at (613) 520-2600, ext. 2664 or Dr. K. Matheson, Chair, Department of Psychology, 520-2600, ext. 2648.

Informed consent form

I have read the above form and understand the condition of my participation. In addition, the research "Psychometric evaluation of the Correctional Program Assessment Inventory" was explained to me by the principal investigator Aleksandra Milovic. I agree to participate in this study and I grant Aleksandra Milovic permission to use my answers in Her Doctoral dissertation.

Participant's Name _____

Participant's Signature _____

Investigator's Name _____

Investigator's Signature _____

Date _____

Appendix P

Correctional Program Assessment Inventory Questionnaire-Debriefing

Thank you for your participation in this study. This research is conducted by Aleksandra Milovic through Carleton University, Ottawa, as a part of her Doctoral dissertation. The main purpose of this study is to determine whether the Correctional Program Service Inventory (CPAI) can be used in the evaluation of correctional treatment programs. The questionnaire you completed and the interview you had with the principal investigator will provide data on the best administration and scoring method for this inventory. In addition, the CPAI will be scored based on an audiotaped interview by two independent raters in order to obtain data on inter-rater agreement. Once the study is completed we will be pleased to share results with you.

If you have any questions or concerns regarding your participation or regarding the research project, you can contact the principal investigator Aleksandra Milovic at (613) 721-9387, or Dr. Don Andrews, Professor, Carleton University, at (613) 520-2600 ext. 2662. If you have any ethical concerns about this research then please contact Dr. M. Gick, Chair, Department of Psychology Ethics Committee, at (613) 520-2600, ext. 2664 or Dr. K. Matheson, Chair, Department of Psychology, 520-2600, ext. 2648.

Appendix Q

Effective Treatment Elements Questionnaire-Debriefing

Thank you for your participation in this study. This research is conducted by Aleksandra Milovic through Carleton University, Ottawa, as a part of her Doctoral dissertation.

The main purpose of this study is to determine whether the Correctional Program Service Inventory (CPAI) can be used in the evaluation of correctional treatment programs. The questionnaire you completed will provide information as to whether or not the treatment elements that are included in the CPAI are considered by correctional personnel and the general public as important for program effectiveness. Your answers will provide data as to whether or not this inventory covers effective treatment elements within correctional programs. Once the study is completed we will be pleased to share results with you.

If you have any questions or concerns regarding your participation or regarding the research project, you can contact the principal investigator Aleksandra Milovic at (613) 721-9387 or Dr. Don Andrews, Professor, Carleton University, at (613) 520-2600 ext. 2662. If you have any ethical concerns about this research then please contact Dr. M. Gick, Chair, Department of Psychology Ethics Committee, at (613) 520-2600, ext. 2664 or Dr. K. Matheson, Chair, Department of Psychology, 520-2600, ext. 2648.