

CRC 37/98-9

2 0 0 1

research report

*An Evaluation of Anger Management Programs
with Violent Offenders in Two Australian States*

Volume 1

The Forensic and Applied Psychology Research Group



UniSA



365.66
099423

f
EVA
v.1

Educating Professionals - Creating and Applying Knowledge - Serving the Community

**AN EVALUATION OF ANGER MANAGEMENT PROGRAMS
WITH VIOLENT OFFENDERS IN TWO AUSTRALIAN
STATES**

Kevin Howells, Andrew Day, Susan Bubner, Sue Jauncey, Ann Parker, Paul Williamson, Karen
Heseltine

Forensic and Applied Psychology Research Group
University of South Australia



CONTENTS

List of figures	5
List of Tables.....	6
Acknowledgements.....	7
Introductory Comments.....	8
General introduction.....	10
Anger management programs for violent offenders.....	10
Violent crime is a source of great social concern in Australia	10
Treatment programs developed in clinical settings have been shown to be applicable to criminal justice settings.....	11
Offenders with anger problems tend to be more difficult to manage than other offenders	12
The need for program evaluation.....	12
Summary and objectives of proposed research program.....	13
Norming Study	15
Introduction	15
Method.....	17
Participants.....	17
Procedure	17
Results	18
Concurrent validity.....	18
Levels of anger as measured by the STAXI and NAS.....	19
Comparison between SA and WA samples	19
Comparison between Aboriginal and other offenders	20
Levels of anger across sentence	21
Levels of motivation for treatment	21
Level of risk of the sample.....	23
The relationship between risk, needs and motivation.....	23
Discussion.....	25
Pilot study.....	27
Introduction	27
Method.....	27
Participants.....	27
Procedure	28
Results	29

Preliminary Analysis	29
Other measures	30
Discussion	31
Modifications to the design as a result of the pilot	31
Conclusion	32
Main Study	34
Introduction	34
Hypotheses	34
Method	35
Participants	35
Western Australian sample	36
South Australian sample	36
Procedure	37
Pre-intervention	37
During intervention	39
Post-intervention	39
Two month follow-up	40
Six month follow-up	40
Research hypotheses, project design and methodology	41
Summary of research methodology	41
The anger management programs	41
Measures: further details	42
Outcome measures	42
The State-Trait Anger Expression Inventory	43
The Novaco Anger Scale	44
The Modified Watt Anger Knowledge Scale	44
The Readiness to Change Questionnaire	45
Serin Treatment Readiness Scale	45
Serin Treatment Performance Scale – self-report	45
Serin Treatment Performance Scale – facilitator-report	46
The Modified Overt Aggression Scale	46
The Violent Offender Treatment Programme Risk Assessment Scale	47
Incidents and charges	47
Program integrity	48
Results	49
Correlational analyses	49
Descriptive statistics	52
Evaluating treatment effects on all key outcome measures	57

Evaluating the effect of state and prisoner location on treatment effects	61
Evaluation of treatment performance	62
Changes in anger during two and six month follow-ups.....	63
Analyses of differential dropout of participants	68
Analysis of predictors of treatment improvement	69
Relationship between risk, needs and motivation for treatment effectiveness.....	74
Relationship between program integrity and treatment effectiveness.....	78
Discussion and Overview	82
Main Outcome Study.....	84
Does anger management work?.....	86
State and Community/prison differences.....	87
Do program effects last?	88
For whom does anger management work best?.....	89
How do the outcome study results compare with previous studies?.....	91
Explaining low impact of offender programs	93
Summary of the Main Findings from the three studies conducted.....	94
Recommendations for Improving Anger Management Outcomes.....	95
References	97
Appendixes.....	Separate Book

LIST OF FIGURES

Figure 1	Mean change and pre-test post-test means for anger in by group and need/ risk/ motivation category.....	75
Figure 2	Mean change and pre-test post-test means for total NAS scores by group and need/ risk/motivation category.....	76
Figure 3	Mean change and pre-test post-test means for anger knowledge by group and need/ risk/motivation category.....	77

LIST OF TABLES

Table 1	Anger Scale Correlations.....	18
Table 2	Descriptive statistics for the STAXI and NAS	19
Table 3	Mean Scores on STAXI and NAS for SA and WA offenders	20
Table 4	Mean Scores on STAXI and NAS for Aboriginal and non-Aboriginal offenders	20
Table 5	Mean Scores Relating Stage of Sentence to Levels of Anger.....	21
Table 6	Item Loadings for First Three Factors of RCQ.....	22
Table 7	Numbers of Respondents Allocation to each Readiness to Change category	23
Table 8	Descriptive Statistics for the Outcome Measures	29
Table 9	Means and Standard Deviations on STAXI for pre and post-test.....	30
Table 10	Pilot Program Integrity Ratings	31
Table 11	Program Integrity Ratings.....	48
Table 12	Correlation of STAXI subscales and NAS subscales	50
Table 13	Correlation of NAS subscales with other Outcome Measures.....	51
Table 14	Correlation of other Outcome Measures.....	52
Table 15	Mean scores on Outcome Measures of SA and WA.....	54
Table 16	Mean scores on Outcome Measures of Aboriginal and non-Aboriginal offenders.....	55
Table 17	Mean Scores Relating Stage of Sentence to Levels of Anger.....	56
Table 18	Means for Control and Treatment Groups at Pre- and Post-treatment and Repeated Measure ANOVA results on Key Outcomes.....	57-58
Table 19	Frequencies of Prisoners in each Readiness to Change category at Pre- and Post-treatment for control and treatment groups.....	60
Table 20	Evidence of Linear and Quadratic Trends in Outcome Changes over time for the Treatment Group.....	66
Table 21	Correlations between Pre-treatment Scores and Corresponding Pre-treatment Post-treatment Change Scores for Treatment and Control Participants.....	71
Table 22	Correlations between Possible Predictors of Treatment Effectiveness and pre- post Changes in Outcome Variables for Treatment and Control Groups.....	72-73
Table 23	Mean Change Scores on Outcomes in Low and High Integrity Programs (client ratings).....	80
Table 24	Mean Change Scores on Outcomes in Low and High Integrity Programs (facilitator ratings).....	81

ACKNOWLEDGEMENTS

The successful completion of this project required the assistance and collaboration of a number of agencies and individuals. We would like to thank the Criminology Research Council for a grant made to Kevin Howells and Andrew Day to support the project. We are also grateful to the Department for Correctional Services in South Australia and the Ministry of Justice in Western Australia, who contributed financially and administratively.

We would also like to thank John Paget (CEO for the Department of Correctional Services) and Terry Simpson (CEO for the Ministry of Justice) for their support for the project. Both their Departments have demonstrated their commitment to evaluation and research. Michael Brownjohn, Jenni MacKenzie and Joanne Favretto provided us with considerable assistance in South Australia. Peter-Upton Davies, Audrey Baker, Graham Chapman and Peter Marshall were invaluable to the project in Western Australia.

We are grateful to Professor Raymond Novaco from the University of California at Irvine, whose work continues to inspire our efforts and who kindly allowed us to use the new form of his anger scales and provided computer scoring of them. Thanks also to Louise Warren for assistance with the computer scoring in the United States. Dr Ralph Serin from Correctional services in Canada was helpful in providing information and ideas relating to his work on readiness and on violence programs.

We would like to thank the program facilitators and many other correctional staff for their hard work and cooperation. Finally, and most importantly, our thanks are due the many program participants who completed lengthy questionnaires. Without them the project would never have been completed.

INTRODUCTORY COMMENTS

The project to be described in this report was conducted in South Australia and Western Australia from February 1999 to January 2001, with the active support of the Department for Correctional Services in SA and the Ministry of Justice in WA. Three studies are described in the report: the Normative Study, the Pilot Study and the Main Outcome Study. The first two studies were on a small scale and form only a backdrop to the large outcome study which was the main purpose of the project. The data were collected in SA by Susan Bubner and in WA by Susan Jauncey over the two years. The remainder of the FAPRG team was involved in project design, data analysis and assembling the final report.

In writing the report we have been mindful of the fact that it may be read by a range of people, some of whom may require only a "snapshot" of the major results and the ensuing recommendations. Others may wish to read the methods and findings in detail. Some specialists in the area may be interested to see materials in detail, particularly questionnaires designed for the project.

For those interested to begin with the "snapshot" we recommend reading the summary and recommendations in pages 94 to 96.

The Discussion and Overview Section (pages 82 to 96) provides a longer account of the main findings and themes emerging from the project. For those requiring detail about measures and other aspects we have prepared Appendices in a separate volume.

Our general conclusion is that anger management programs have only a very modest impact in general, but that some particular offenders benefit more than others. For this reason we have recommended a constructive, developmental approach whereby the improvements brought about by anger management interventions can be enhanced. The future targeting of treatment on suitable participants appears to be the way forward.

The views contained in this report are the authors' and do not necessarily represent the views of the Department for Correctional Services (SA) or the Ministry of Justice (WA).

Professor Kevin Howells, on behalf of the project team:

Dr Andrew Day

Ms Susan Bubner

Ms Sue Jauncey

Ms Ann Parker

Mr Paul Williamson

Ms Karen Heseltine

GENERAL INTRODUCTION

Anger Management Programs for Violent Offenders

The 1990s have witnessed a renaissance of interest in the rehabilitation of offenders in many criminal justice systems throughout the world. The causes of this re-awakening of interest are many, but include the increasing evidence that rehabilitation programs have a significant impact on rates of recidivism (Losel, 1996; McGuire, 1995, 1998). As this evidence has accumulated, the notion that "nothing works" in offender rehabilitation has slowly given way to an emphasis on identifying the characteristics of programs that are likely to be effective and, conversely, the characteristics of those that are likely to have no effect or even an adverse effect on rehabilitation rates (Andrews & Bonta, 1994; Howells & Day, 1999).

In recent years, correctional administrations have increasingly identified violent offenders as a key target group for rehabilitation programs. The reasons for this are clear:

1. Violent crime is a source of great social concern in Australia.

Whilst there is limited evidence to suggest that violent crime is increasing (Indermaur, 1999) the numbers of offenders imprisoned for violent offences has risen over the past few years. However, it has been well established that violent offences are not necessarily angry offences (e.g. Mills, Kroner & Forth, 1998) and it has been argued that poor anger control often plays a role in violent offending and can be considered a criminogenic need for many violent offenders (Howells, Watt, Hall, & Baldwin, 1997). Similarly, Novaco (1995) suggested that anger can be used as a risk factor for the prediction of violence. However, there is an emerging alternative view (Loza & Loza-Fanous, 1999a; 1999b) that raises the question of the usefulness of anger in the prediction of violent and non-violent recidivism.

As a group, offenders commonly experience difficulties with anger. On average, prison inmates score substantially higher on measures of anger arousal and expression than other populations (Spielberger, 1991), with violent offenders scoring higher than non-violent offenders (Mills, et al., 1998).

2. Treatment programs developed in clinical settings have been shown to be applicable to criminal justice settings.

The experience and expression of anger have been studied in a wide range of clinical and non-clinical populations, including students (Deffenbacher, Story, Brandon, Hogg, & Hazaleus, 1988; 1990), community, health-care clients (O'Leary, 2000), psychiatric/residential patients (Stermac, 1986) and adolescents in residential settings (Frey, Hirschstein, & Guzzo, 2000; Swaffer & Hollin, 1997). Cognitive-behavioural anger management programs have been developed for use with many of these populations (eg. Medd & Tate, 2000; Reilly & Shopshire, 2000) and initial research suggests that they are effective in reducing problems with anger expression (Beck & Fernandez, 1998). There is reason to believe that similar cognitive-behavioural programs may also be appropriate for offender groups. A number of studies have highlighted the role of cognitive factors in anger arousal and expression in offenders. For example, Gembora (1986) found that inmates with chronic anger control problems perceive and interpret anger provoking situations differently from those without anger problems. Similarly, Ford (1991) found that anger in inmates was related to cognitive variables in anger arousal and expression, a finding supported by Copella and Tata (1990) and McDougall, Venables and Roger (1991).

Findings such as these have led to the widespread implementation of anger management programs in prison settings around the world. These programs tend to be brief (up to 10 sessions) cognitive behavioural programs designed to reduce anger arousal and improve anger control (eg. Novaco, 1994; 1997). Anger Management programs take a skills approach and attempt to help program participants develop alternative strategies in the control and expression of angry impulses (for a discussion of the rationale for anger management with violent offenders see Howells, 1998; Levey & Howells, 1990; Novaco, 1997). Whilst these programs have been shown to be reasonably effective with different clinical populations (Beck & Fernandez, 1998; Edmondson & Conger, 1996), there has been only limited research investigating program outcomes with offender groups (e.g. Day, Maddicks & McMahon, 1993; Watt & Howells, 1999).

3. Offenders with anger problems tend to be more difficult to manage than other offenders

Anger appears to be a particularly important emotion in residential settings with offenders. Anger problems have been linked with prison adjustment, disciplinary problems, assaults, and violence. Anger is a strong predictor of aggression amongst incarcerated adolescents (Cornell, Peterson & Richards, 1999) and has been shown to be associated with physical assault on care staff (Gentry & Ostapuik, 1989). Institutional staff rate anger as the primary problem in secure psychiatric facilities (Rice, Harris, Quinsey & Cyr, 1990). Kroner and Reddon (1995) found that interpersonal problems in prisoners were strongly related to anger expression and arousal, and the inward expression of anger was significantly related to dimensions of psychopathology.

While research on prison adjustment suggests that negative emotions (such as anxiety and depression) decrease over time, this does not appear to be the case for anger. In one study, prisoners reported two episodes of anger per week during the initial stages of their incarceration. The frequency of anger experiences increased the longer they were in prison (Zamble & Porporino, 1990). The finding that anger is a stable and present feature of long-term imprisonment appears to be robust (Bonta & Gendreau, 1990).

The Need for Program Evaluation

The lack of evaluation of Anger Management programs with violent offenders is a cause for concern. Many jurisdictions in Australasia (for example, Western Australia, South Australia and New Zealand) run substantial, well organised programs that would meet recognised standards of good practice and which, in many respects, lead the field internationally. Little information is available, however, about what impact such programs have, and for which offenders they are most effective. Given the large numbers of offenders completing such programs (we estimate upwards of 700 annually in these three jurisdictions alone) a major opportunity for useful research exists which could have significant implications for the planning and implementation of future cost-effective rehabilitation initiatives across Australasia and internationally.

An important issue that needs to be considered in offender rehabilitation is the heterogeneity of the offender population. Offenders differ not only in the patterning and nature of their offences but also in terms of the variables that have contributed to the onset and maintenance of their offending behaviour. For this reason it is necessary to tailor rehabilitation programs to the characteristics and criminogenic needs (Andrews & Bonta, 1994) of particular offender groups.

Andrews & Bonta have argued that program outcomes are likely to be improved if, amongst other things, programs target offenders who have a high risk of re-offending and a high need for intervention. This seems a particularly important area of selection for intervention programs that often select participants on the basis of offence.

A second area of interest arises from attempts to describe the stages through which a person moves in attempting to resolve a problem, and in matching different intervention strategies to clients at different stages of change. Andrews & Bonta (1994) argue that services should be targeted not only at those who have a high level of need, but also at those who are likely to respond to the particular intervention offered. Thus, individual characteristics such as the extent to which an individual recognizes a problem or is motivated to change form an important element of pre-service assessment. Whilst there have been several attempts to measure these individual factors (e.g. Heather, Rollnick & Bell, 1993; McMurrin, Tyler, Hogue, Cooper, Dunseath, & McDaid, 1998), these have largely focused on measuring aspects of treatment readiness in clinical populations with drug and/or alcohol problems.

Summary and objectives of research program

Despite the renewal of interest in offender rehabilitation and the increased optimism about programs attempting to prevent re-offending, violent offenders are rarely the focus for intervention in the published literature. The current study involved an evaluation of the effectiveness of anger-management programs for violent offenders. The research participants were offenders referred for anger management in South Australia's Department for Correctional Services (DCS) and Western Australia's Ministry of Justice (MOJ). These offenders have a high incidence of previous violent offending.

In the main outcome study to be reported, pre-treatment, post-treatment and follow-up assessments were conducted with offenders undergoing the treatment programs and comparisons made with a 'waiting-list' control group. The objectives of the study were two-fold. First, to determine whether anger management is more effective than no-treatment in producing change. Second, to investigate whether improvement in treatment can be predicted from pre-treatment offender characteristics.

There are three studies to follow:

1. The Normative Study
2. The Pilot Study
3. The Main Outcome Study

Experiences of Anger within an Australian male adult prisoner populations

Introduction

Previous research has focused primarily upon levels of anger experienced by prisoners shortly after reception in prison. In addition much of the reported research on anger assessment has been carried out on North American populations with prisoners recently received into prison on assessment or intake units (e.g. Kroner & Reddon, 1992; 1995; Mills et al., 1998). These studies show that on average prison inmates score substantially higher on measures of anger arousal and expression than other populations (Spielberger, 1991), with violent offenders scoring higher than non-violent offenders (Mills et al., 1998). These comparisons are drawn from norms obtained from North American samples. These may or may not be representative of Australian prison samples because of possible differences in either criminal justice systems (eg. sentencing of violent offenders) or cultural differences in the population (see Anastasi & Urbina, 1997 for a discussion of norming data).

The primary aim of this preliminary study was to gather some descriptive data on levels of anger experienced by a sample of Australian prisoners. This was important for several reasons: first, there may be cultural differences in the expression of anger that prohibit generalisation from North American and European studies. One cultural factor that may be relevant is ethnicity. Aboriginal offenders form a relatively high proportion of the Australian prison population (Walker & McDonald, 1995) and may differ systematically in both their experiences of anger and need for intervention (Mals, Howells, & Day, 1999). Findings from most previous studies on anger are based on predominantly white samples. Sharkin (1996) argued that culture may “*not only influence what are considered appropriate and inappropriate forms of anger expression, but also what makes people angry, what is perceived as anger, and how people feel about their anger*” (p168). Similarly, Averill (1982) argued that culture not only affects when it is appropriate to be angry, but also the extent to which violence is an understandable response to anger. Sharkin noted that for some cultural groups, anger may be based in the context of histories of discrimination and oppression. Indermaur (1996) argued that, on an individual level, focussing on offender perceptions of violence may help to identify specific cultural values. Whilst there has been some work on anger arousal in African-Americans (eg Armstead, Lawler, Gorden, Cross &

Gibbons, 1989) and Mexican-Americans (Deffenbacher & Swain, 1999), to our knowledge there has been little work to date on an Australian Aboriginal population.

Second, current anger management programs are delivered to inmates at various stages of their sentences. To gain a representative picture of anger experiences, norms should be established for a general prison population group, rather than just those inmates who have been recently received into prison. It is likely that experiences of anger change over the course of a sentence as prisoners adjust to imprisonment or encounter additional stresses over time. Zamble and Porporino (1990) reported that anger experiences increased the longer prisoners were in prison. This contrasted to other emotions (anxiety and depression), which decreased in frequency as time went on. The finding that anger is a stable and present feature of long-term imprisonment appears to be robust (Bonta & Gendreau, 1990).

Third, such normative data may assist in service planning and the development of selection protocols that optimise treatment outcomes. The identification of prisoners who not only experience high levels of anger, but also are classified at high risk of re-offending, and are motivated to address anger problems may help to provide an estimate of the need for programs in a prison population at any one time. Whilst the normative measures used in this study help to identify those with a need for intervention, and risk data can be gathered largely from archives, there is need to develop an appropriate screening measure of treatment motivation. One method of assessing motivation is to establish an individual's readiness to accept treatment. Recently, there has been much interest in describing the stages through which a person moves in attempting to resolve a problem, and in matching different intervention strategies to clients at different stages of change. Thus, individual characteristics such as the extent to which an individual recognizes a problem or is motivated to change form an important element of pre-service assessment. Whilst, there have been several attempts to measure these individual factors (eg. Heather, et al., 1993; McMurrin, et al., 1998), these have largely focused on measuring aspects of treatment readiness in clinical populations with drug and/or alcohol problems. This study will determine the extent to which existing measures may be adapted for use with anger problems in an offender population, and to examine the relationship between level of need (as measured by anger inventories) and self-reported treatment readiness.

Method

Participants

Participants were 121 adult male convicted prisoners, in five different prisons in two states of Australia: South Australia and Western Australia. Participants were selected randomly from the prison population. Fifty-two were in maximum-security prisons, 49 in medium security, and 20 in minimum security. Eight of the participants surveyed in one of the maximum-security prisons (Yatala Labour Prison) were serving short sentences (less than six months) for fine defaults. These eight were subsequently incorporated into the minimum-security group.

The mean age of the sample was 30.04 years (SD= 10.4, range = 18-68). Sentences being served ranged from 1 month to 23 years with a mean of 58.7 months (SD=61.9). Eighty-three described their ethnic background as Australian, including 34 Australian Aboriginals, and 29 from European backgrounds. Seventy-four described themselves as single, 29 as married/defacto (with the remainder divorced/widowed or separated). Eighty-two were unskilled/labourer/unemployed, with 21 describing themselves as tradesperson/semi-skilled. Nearly half of the sample (56) of the sample reported current offences that included violence. Comparison with census data from the Department for Correctional Services (SA) suggested that the current sample is reasonably representative.

Procedure

Male adult prisoners in five prisons across two different Australian states were invited to participate in the study. Prisoners were approached randomly in residential units and the purpose of the study explained. Participation was voluntary and written consent obtained. Full ethical approval was obtained before the study commenced. Test results were confidential and had no bearing on subsequent treatment or services received in the prison. Participants were asked to complete the measures described above in addition to a short demographic questionnaire. One further questionnaire on treatment motivation was also completed. A total of 101 prisoners agreed to participate in the research. Subsequently a sample of 10 additional Aboriginal and Torres Strait Islander offenders from each state were asked and agreed to participate. This resulted in a total sample of 121 prisoners. Missing data for items on the measures were replaced on an item by mean basis (Tabachnick & Fidell, 1996).

Measures

All participants completed the following measures:

(a) The Novaco Anger Scales (Novaco, 1994)

For the purposes of the present corrective study only the NAS Part A and Part B scores were completed.

(b) The State-Trait Anger Expression Inventory (STAXI; Spielberger, 1991)

This measure is described more fully in the Main Outcome Study later in this report. For the present study only the State Anger and Trait Anger scores were calculated.

(c) Treatment Readiness Scales

All participants completed the Readiness to Change Questionnaire, which was designed and adapted for the present study and the TCU Treatment Readiness Questionnaire. The development of RCQ is described in more detail below.

(d) A violence risk measure – the VOTRAS – was completed for each participant using file information. This test was developed in Western Australia as a predictor for future violent recidivism (Ward and Dockerill, 1999).

Results

Concurrent Validity

Correlations between the anger scales were conducted to assess concurrent validity. Results are reported in Table 1.

Table 1. Anger Scale Correlations

	1	2	3	4
1. NAS Part A	-			
2. NAS Part B	.78	-		
3. STAXI State Anger	.49	.42	-	
4. STAXI Trait Anger	.79	.79	.42	-

Note: N=121. All correlations are significant at $p < .01$. NAS = Novaco Anger Scale; STAXI = State-Trait Anger Expression Inventory

2. *Levels of Anger as measured by the STAXI and the NAS*

Despite the difficulties in comparing samples selected under different conditions (Kendell & Grove, 1988), it was possible to compare the current sample with previous male prison groups. Generally, scores on all of the subscales were comparable with previously published norms. The Australian sample scored slightly lower on all six subscales of the STAXI than the North American prison sample reported by Spielberger (1991), but scoring slightly higher on the Novaco Anger Scale (Parts A and B) than the North American prison sample reported by Mills, et al., (1998). Scores were lower than those reported by Watt and Howells (1999) from a sample of Australian prisoners that had already been selected for treatment (intervention group). Refer to Table 2 for descriptive statistics.

Table 2. Descriptive statistics for the STAXI and NAS

	Present study N=121		Watt & Howells, 1999 N=18		Spielberger, 1991 N=563	
	mean	SD	mean	SD	mean	SD
STAXI						
State Anger	11.88	4.1			15.06	6.6
Trait Anger	18.73	7.2	20.61	7.6	21.66	6.7
Anger Control	22.21	6.0	21.89	6.3	24.79	5.0
Ax Effects	26.27	11.6				
Ax In	16.56	4.8	18.61	6.6	18.06	4.6
Ax Out	15.93	4.6	16.61	3.9	16.52	5.0
NAS					Mills et al., 1998 n=102	
Part A	87.64	18.5			80.38	15.8
Part B	58.43	16.6			53.01	15.8

Note: NAS = Novaco Anger Scale; STAXI = State-Trait Anger Expression Inventory; SD = Standard Deviation

3. *Comparison between South Australian and Western Australian Samples*

Mean scores were calculated comparing participants from the two states. These scores can be found in Table 3. The results indicate that respondents in South Australia scored higher on all measures of anger than respondents from Western Australia.

Table 3. Mean Scores on Anger Measures between South Australian and Western Australian Offenders

	STAXI Anger	State	STAXI Anger	Trait	NAS Part A	NAS Part B
Western Australian (N=61)	11.77 (SD 4.1)		16.81 (SD 5.7)		82.70 (SD 17.2)	54.76 (SD 16.4)
South Australian (N=60)	11.99 (SD 4.2)		20.69 (SD 8.0)		92.67 (SD 18.6)	62.16 (SD 16.1)

Note: N=121. NAS = Novaco Anger Scale; STAXI = State-Trait Anger Expression Inventory; SD = Standard Deviation

An analysis of covariance, controlling for security level of the participants' prison, comparing the differences between South Australian and Western Australian prisoners these differences were significant on the measure of trait anger ($F=5.01$, $p=.008$), NAS part A ($F=5.21$, $p=.007$) and NAS Part B ($F=6.50$, $p=.035$), even when the security level of the prison was controlled for.

4. *Comparison between Aboriginal Offenders and Other Offenders*

A comparison was then made between levels of anger as measured by the STAXI Trait Anger Scale, and Novaco Part A and Part B scales between respondents describing themselves as Aboriginal or Torres Strait Islander (ATSI) (N=14) and others (N=85, 2 missing). Mean scores on each of the scales are presented in Table 4.

Table 4. Mean Scores on Anger Measures of Aboriginal and Non-Aboriginal Offenders

	STAXI Anger	State	STAXI Anger	Trait	NAS Part A	NAS Part B
Non-ATSI (N=85)	11.72 (SD 3.8)		18.07 (SD 7.3)		85.25 (SD 18.4)	56.90 (SD 16.6)
ATSI (N=34)	12.33 (SD 4.8)		20.49 (SD 6.7)		93.61 (SD 17.2)	62.15 (SD 16.3)

Note: N=119, 2 cases missing data. NAS = Novaco Anger Scale; STAXI = State-Trait Anger Expression Inventory; SD = Standard Deviation

These results indicated that the Aboriginal offenders in this sample tended to score higher on measures of anger than non-Aboriginal offenders. Analyses of variance revealed that these differences were significant on the NAS Part A scale ($F=5.203$, $p=.024$) and approached significance on the Trait Anger Scale ($F=2.77$, $p=.099$). No significant differences were found on the State Anger and NAS Part B scales.

5. *Levels of Anger across Stage of Sentence*

The stage of sentence was obtained by calculating how long each participant had served in prison, then dividing by the overall sentence. From the 99 participants for whom data was available, a group 45 participants was identified who had served less than one quarter of their total sentence. Their scores on the anger measures can be found in Table 5.

Table 5. Mean scores relating stage of sentence to levels of anger

Anger measure	Less than 25% served (n=45)		More than 25% served (n=54)	
	mean	SD	Mean	SD
State Anger	11.52	3.5	12.11	4.9
Trait Anger	16.90	5.9	19.41	8.0
Part A	84.56	18.1	89.50	19.8
Part B	53.87	15.3	60.52	16.2

Analyses of variance revealed that these differences were significant for the Part B measure ($F= 4.38$, $p=.039$), and approached significance for Trait Anger ($F=3.20$, $p=.077$).

Levels of Motivation for Treatment

Two measures of treatment readiness and motivation were used in this study, the Readiness to Change Questionnaire and the TCU Treatment Readiness Questionnaire. The TCU questionnaire consists of four subscales problem recognition, treatment readiness, desire for help and external pressures. The Readiness to Change Questionnaire was devised for the present study.

The utility of both measures was examined by factor analysis. Factor analysis of the TCU questionnaire failed to confirm the proposed four factors and was consequently discarded. Examination of the correlation matrix of the 16-item Readiness to Change Questionnaire (RCQ) indicated the matrix was suitable for factor analysis. A principal components analysis with

varimax rotation yielded a four-factor solution explaining 65.3% of the variance. Deletion of the additional four Determination items yielded a clear three factor solution explaining 60.9% of the variance. This three-factor solution corresponded closely to the three stages of change (Pre-contemplation, Contemplation and Action) obtained by Miller and Rollnick (1991) and Bubner (1999). As such, the additional Determination items were deleted, leaving a scale corresponding closely to the original. Each of the items loaded on their corresponding factor, with the exception of item 1, a Pre-contemplation item, which loaded negatively on Factor 1. Examination of item 1 revealed that a negative score on the item would make it consistent with a Contemplation item, as earlier defined by Bubner. The factor loadings for the scale, applying a critical value of 0.542 (Stevens, 1996) can be found in Table 6.

Table 6. Item loadings for first three factors extracted from Varimax rotation with percentage variance accounted for by each.

Item and hypothesised stage in brackets	Factor 1 Contemplation 31.7% of variance	Factor 2 Action 16.7% of variance	Factor 3 Pre-contemplation 12.5% of variance
11. My anger is a problem sometimes (C)	.862		
4. I'm entitled to get angry, but sometimes I go too far (C)	.783		
5. Sometimes I think I should try to control my anger (C)	.688		
10. I am at the stage where I should think about managing my anger (C)	.678		
1. I don't think I have too many problems with anger (P)	-.654		
9. Anyone can talk about wanting to do something about anger, but I am actually doing something about it (A)		.769	
14. I am actually changing how I deal with anger right now (A)		.719	
8. I have just recently changed how I deal with anger (A)		.708	
2. I am trying to control my anger more than I used to (A)		.577	
15. Controlling anger better would be pointless for me (P)			.795
6. It's a waste of time thinking about anger (P)			.669
13. There is no need for me to think about changing how I deal with anger (P)			.570

Note: P= Precontemplation item; C=Contemplation; A=Action

Given the factor structure of the adapted Readiness to Change Questionnaire (RCQ), participants were classified according to Readiness to Change category using the Quick Method of Stage Allocation outlined by Miller and Rollnick (1991). Refer to Table 8 for the proportion of respondents in each of the readiness to change categories (Pre-contemplation, Contemplation and Action). The Readiness to Change category of respondents was then checked against whether the sample had previously completed the anger management program. These results showed that of the 53 people who had completed at least one anger management program, 38 (72%) were in the Action Readiness to Change category compared with 38% of the untreated group.

Table 7. Numbers of Respondents allocated to each Readiness to Change category

Number of Respondents	Readiness to Change category		
	Precontemplation	Contemplation	Action
Total sample (N=119, 2 missing data)	19 (16%)	37 (31%)	63 (53%)
Completed anger program (N= 53)	3 (6%)	12 (23%)	38 (72%)
Not completed at least one anger program (N=66)	16 (24%)	25 (38%)	25 (38%)

Level of Risk of the Sample as measured by the VOTRAS

Ward and Dockerill (1999) suggested a cut-off point of 11 for the VOTRAS, reporting that this cut-off point correctly classifies the risk of re-offending in 70% of cases. Seventy-four of the current sample would be identified as high risk using this criterion.

8. The Relationship between Risk, Needs and Motivation in Planning Anger Management Programs

One way of estimating the need for a service is to look at problem severity. Problem severity can be estimated using the Trait Anger Scale of the STAXI, although relatively arbitrary cut-off points for classification have to be used. One possibility is to select clients using standard deviations from the mean from the prison samples. High anger offenders can be classified as scoring over one standard deviation above the mean (giving a cut-off score of 25 on the Trait Anger Scale). Another possibility is to use the top quartile of scores (Spielberger, 1991; Watt & Howells, 1999), which for the present sample would mean a cut-off point of 23. Previous work

with incarcerated adolescents has suggested a cut-off point of 86 on the NAS Part A, and 20 on the STAXI Trait Anger Scale as predictive of institutional aggression (Cornell et al, 1999). Given that the reduction of institutional aggression is an important target of prison-based anger management programs, the criteria adopted by Cornell et al (1999) were adopted here to define a high need group. Thirty percent of the current sample scored over 20 on the STAXI Trait Anger with 50% scoring over 86 on the NAS Part A. Forty-five participants (37%) met the cut off point for both measures.

Of the 45 participants classified as high in anger, only three were classified as in the Pre-contemplation Readiness to Change category. In other words, nearly all of these offenders recognised a problem with anger on at least some level. Twenty-three of the 45 (51%) were classified in the Contemplation stage with the other 19 (42%) in the Action Readiness to Change category.

It was also possible to categorise a low anger group using the lowest quartile on the Trait Anger Scale of the STAXI. This gives a cut-off point of 16 for prison norms quoted by Spielberger (1991). Of the 47 participants classified as low in need, 14 (30%) were classified as in the Pre-contemplation Readiness to Change category, 5 (11%) were classified in the Contemplation stage with the other 28 (60%) in the Action Readiness to Change category. These results imply that two broad groups exist in low anger prisoners - those who do not experience any problems and see no need to think about anger (Pre-contemplators) and those who feel they are successfully managing their anger (action).

The VOTRAS gives an estimate of the risk of re-offending for each offender. High-risk offenders are defined by scoring over 11 on the VOTRAS. Across this sample, this would mean that 74 could be classified as high risk. Of the 45 respondents identified as high in need, 27 were classified as also being high risk and responsive to change.

In summary, of a sample of 121 adult male prisoners, 27 (22%) could be argued to have high risk, high need and some motivation for treatment. These clients could be considered suitable candidates for anger management using the principles articulated by Andrews & Bonta (1994).

Discussion

The present data would support previous work suggesting that adult male prisoners experience generally higher levels of anger than other populations (Kroner & Reddon, 1998; Spielberger, 1991). As a population, prisoners can be considered as having a higher need for anger management programs than other groups. Using the classification procedure described earlier, approximately 30% of the general prison population could be considered to have a need for intervention in the way they manage anger. Data from the risk and motivation scales collected indicates that this same group of prisoners may also be at high risk of re-offending and recognise that they have some problems with anger. As such, they would be ideal candidates for anger management programs using the principles outlined by Andrews & Bonta (1994).

The current sample was drawn randomly from prisons in two Australian states in different security prisons and represents prisoners at different stages of their sentences. The data suggested that prisoners in the later stages of their sentences may experience greater problems with anger. Another group that might be considered for targeting resources is Aboriginal and Torres Strait Islander offenders. These participants scored slightly higher on the anger measures, indicating an increased need for intervention with this group. The apparent under-representation of this group in treatment programs would, therefore, be a cause of concern (Mals, et al., 1999).

Finally, the modification of the Miller and Rollnick (1991) measure of motivation to reduce problem drinking to make it applicable to anger problems has met with success. The new Readiness to Change Questionnaire has demonstrated that offenders are able to be classified reliably according to Readiness to Change category with anger problems. Future work into this area should examine the predictive validity of such a measure (ie. whether those in Contemplation and Action stages show better program outcomes than those in the Pre-contemplation stage).

The study gives further weight to the utility of the Spielberger and Novaco Anger Scale in assessing anger with this population and identifying problems with anger that can be targeted for intervention. The selection of suitable candidates for intervention using the measures described above offers the possibility of improving program outcomes and reducing re-offending.

Assuming the sample is representative of the broader prison population, it is possible to make some broad estimates of the need for anger management program. The total prison population in South Australia at the time of testing was 1,279, with 2,660 in Western Australia. If our estimates of 22% as being high risk, high need and recognising a problem with anger are accurate, this would mean that approximately 870 offenders within the system across both states would be ideally targeted for anger management.

PILOT STUDY

Evaluation of Anger Management Programs

Introduction

Before commencing the main study, a decision was made to pilot the battery of measures and testing protocols to confirm their utility. While previous work by Watt and Howells (1999) indicated that many of the measures were appropriate for this type of evaluation, this project introduced several new measures, planned to implement them on a much wider scale (across two states), and included programs run in community correctional settings. A total of four separate anger management groups were selected for the pilot study: one community based group and one prison based group, in each state. In total, 30 men were allocated to the intervention group, with 18 in the no-treatment control group. Measures were administered pre- and post- the program.

Method

Participants

Prison participants were drawn from referral to programs run in two medium security prisons (Mobilong Prison, SA and Woorooloo Prison, WA). Community participants were from Whyalla Community Correctional Services in SA and Bentley Community Correctional Services in WA. In total, 23 offenders participated in the prison pilot project, with 25 in the community project.

The mean age of the sample was 29.63 years ($SD= 8.18$, range = 18-62). Sentences being served ranged from 5 months to 8 years with a mean of 30.44 months ($SD=24.05$). Forty-two described their ethnic background as Australian, including 9 Australian Aboriginals, and 4 from European backgrounds. Thirty-two described themselves as single, 10 as married/defacto (with the remainder divorced/widowed or separated). Thirty-two were unskilled/labourer/unemployed, with 12 describing themselves as tradesperson/semi-skilled. Twenty of the 30 people who completed the current offence section reported a current offence that included violence.

Scores on the risk assessment measure (the VOTRAS) indicated that both prison and community groups could be considered as high-risk groups. (Prison group, mean=16.29, $SD=6.54$; Community group mean=15.60, $SD=5.88$).

Procedure

Correctional clients referred for anger management programs were invited to participate in the pilot study. Participation was voluntary and written consent obtained. Test results were confidential and had no bearing on subsequent treatment or services received in the prison. Participants were asked to complete a short demographic questionnaire and the series of outcome measures on two occasions, in the week before attending the program and in the week following completion of the program. A total of 48 clients attending the anger management programs agreed to participate in the research (23 prison, 25 community). Two did not complete the program and no post data was available. Missing data for items on the measures was replaced on an item by mean basis (Tabachnick & Fidell, 1996). In addition to the self-report measures, officers were asked to rate behaviour for the week before the programs and the week following completion. Correctional records were reviewed to obtain data regarding the risk of participants and any officially recorded incidents.

Measures

All treatment and control participants completed the Novaco Anger Scales and the State Trait Anger Experience Inventory, described in the Normative Study above.

In addition, all participants completed the Watt Anger Knowledge Scale (WAKS), a measure of basic knowledge about anger described in Watt and Howells (1999) and described in more detail in the Main Outcome Study later in this report.

The Program Integrity Checklist was a measure designed for the present study and is described in Appendix E. The checklist was completed for the treatment programs involved in the pilot study.

The Modified Overt Aggression Scale (described in the Main Outcome Study below and in Appendix B) was completed by all participants.

Results

Preliminary Analysis

Comparisons between the control and intervention groups revealed no significant differences on demographic, offence, or pre-test data. Table 8 describes the scores on pre and post-measures.

Table 8. Pre and Post Measure Scores for Total Sample

Group		State		Trait		Part A		Part B		WAKS		MOAS client	
		pre	post	pre	post	Pre	post	pre	post	pre	post	Pre	post
Prison Control	Mean	12.9	11.0	19.9	20.8	87.4	86.3	55.4	59.0	23.3	24.3	2.7	1.0
	N	7	6	7	6	7	6	7	6	7	6	7	6
	SD	5.0	2.0	6.9	8.2	15.9	18.5	16.6	19.7	6.0	5.8	4.0	0.0
Exper.	Mean	12.8	12.8	20.3	19.5	90.0	87.7	59.7	63.9	18.6	21.8	4.7	1.0
	N	16	12	16	12	16	12	16	12	16	12	16	12
	SD	4.1	6.7	8.4	6.5	20.3	14.6	18.4	14.7	6.1	6.5	9.2	0.0
Total	Mean	12.8	12.2	20.2	19.9	89.2	87.2	58.4	62.3	20.0	22.6	4.1	1.0
	N	23	18	23	18	23	18	23	18	23	18	23	18
	SD	4.3	5.6	7.8	6.9	18.8	15.4	17.6	16.1	6.3	6.2	8.0	0.0
Commun Control	Mean	12.8	17.5	19.0	19.5	92.9	85.6	59.1	67.9	18.2	22.1	14.7	1.0
	N	10	8	10	8	10	8	10	8	10	8	3	8
	SD	4.0	8.0	7.3	4.1	18.9	17.0	18.9	17.3	6.4	5.8	14.6	0.0
Exper.	Mean	13.9	13.1	19.9	18.9	86.6	98.4	58.9	57.8	19.7	21.5	8.1	0.0
	N	14	10	14	10	14	10	14	10	14	10	8	10
	SD	6.3	6.0	6.0	7.0	15.4	23.0	13.6	15.1	6.8	8.1	9.1	0.0
Total	Mean	13.3	15.1	19.8	19.2	90.1	92.7	59.4	62.3	19.3	21.8	11.6	1.0
	N	25	18	14	10	25	28	25	18	25	18	12	18
	SD	5.3	7.1	6.5	5.7	17.1	21.0	12.5	16.5	6.5	7.0	11.5	0.0
Total Control	Mean	12.8	14.7	19.3	20.1	90.6	85.9	57.6	64.1	20.3	23.1	6.3	1.0
	N	17	14	17	14	17	14	17	14	17	14	10	14
	SD	4.3	6.9	6.9	5.9	17.4	16.9	13.5	18.2	6.6	5.7	9.6	0.0
Exper.	Mean	13.3	12.9	20.1	19.2	88.4	92.5	59.3	61.1	19.1	21.6	5.8	1.0
	N	30	22	30	22	30	22	30	22	30	22	24	22
	SD	5.2	6.2	7.3	6.6	18.0	19.2	16.1	14.8	6.3	7.0	9.1	0.0
Total	Mean	13.1	13.6	20.0	19.6	89.7	89.9	58.9	62.3	19.7	22.2	6.7	1.0
	N	48	36	48	36	48	36	48	36	48	36	35	36
	SD	4.8	6.5	7.1	6.3	17.7	18.4	15.0	16.0	6.4	6.5	9.9	0.0

Note: NAS = Novaco Anger Scale; STAXI = State-Trait Anger Expression Inventory; WAKS = Watt Anger Knowledge Scale; MOAS = behavioural rating scale; SD = Standard Deviation

A series of Analyses of Covariance, comparing change scores on the outcome measures between control and intervention groups, using social desirability scores as a covariate, revealed that differences between the intervention and control groups were not statistically significant (State Anger: $F= 1.168$, $p=.32$; Trait Anger: $F=1.762$, $p=.19$; NAS Part A: $F=2.937$, $p=.07$; NAS Part B: $F=1.208$, $p=.31$; WAKS: $F=.868$, $p=.43$; MOAS: $F=.51$, $p=.48$).

A closer examination of the STAXI scales showed no significant differences between pre- and post-test on the scales. Table 9 shows the means and standard deviations for treatment and control groups.

Table 9. Means and Standard Deviations on STAXI for both pre- and post-test and Control and Treatment Groups

		Pre test		Post test	
		mean	sd	mean	sd
State anger	C	12.40	3.6	14.94	7.5
	T	13.88	5.7	13.59	6.8
Trait anger	C	18.76	6.6	18.00	4.5
	T	20.40	6.9	19.94	7.2
Anger in	C	16.79	4.0	16.93	5.3
	T	17.65	3.9	16.41	5.0
Anger out	C	17.85	4.9	17.86	4.3
	T	17.50	4.0	17.06	6.4
Anger control	C	20.56	3.4	20.64	6.3
	T	18.84	4.7	20.65	5.9
Anger expression	C	30.08	9.8	30.16	11.8
	T	32.31	9.0	28.82	12.5

Note: C= control group; T= treatment group

Other Measures

Program Integrity Checklists

Ratings on the integrity measures showed that the four pilot programs adhered closely to the program manuals, as rated by both a client and a facilitator. Ratings can be found in Table 10 below.

Table 10. Pilot Program Integrity Ratings

Coding category	Client Ratings								Facilitator Ratings							
	Prison				Community				Prison				Community			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Timings	66	7	5	0	83	6	13	0	80	18	6	1	75	10	20	0
Content	6	7	102	2	11	35	101	7	10	24	125	1	5	47	106	3
Process	0	2	24	0	0	12	17	1	3	11	16	0	2	12	15	1
Exercise	9	4	55	1	17	14	49	2	14	12	56	2	14	20	49	3
Discussion	3	1	35	0	12	13	23	2	1	9	28	3	4	11	26	1

Note: Timing ratings 1=yes, followed timing in manual; 2=no; 3=unsure; 4=unable to rate. For other categories, 1: not covered at all; 2=partially covered; 3=thorough; 4=unable to rate.

Discussion

The pilot project has demonstrated the utility of the data collection protocols for use in the main study. No significant problems were experienced with data collection in either prison or community settings. The measure of program integrity revealed that each of the programs were delivered consistently with the program manual as rated by both the facilitator and a program participant.

Modifications to the design as a result of the pilot

1. The behavioural rating scale (MOAS) produced inconsistent findings. Self-report ratings of behaviour over the week before and after completion of the programs by the offenders appeared to give the most realistic rating of behaviour. Correctional officer ratings for the prison sample were on the whole not able to pick up any changes in behaviour due to a floor effect on the measure (lack of positive observations). This may in part be due to the prison environment, where aggressive behaviour (particularly verbal aggression) is more common and hence not seen as noteworthy. Community correctional officers have limited opportunity to observe offenders' behaviour directly and thus are basing responses largely on interview contacts with the individual. As a way of increasing the sensitivity of the rating scale, a modification was introduced for the main study whereby officers rated behaviour for two time periods - over the previous week and

over the previous month. It was anticipated that this would increase the numbers of positive ratings on the scale so that comparisons could be made between pre- and post- program.

2. As a total package of measures, some concerns were raised regarding the length of time taken to complete the assessments. In addition to the measures reported here, the final measures for use in the main study included a brief questionnaire on treatment readiness and treatment performance, as well as the Readiness to Change Questionnaire developed in the norming study. In an attempt to reduce the length of the questionnaire, the Watt Anger Knowledge Scale was reduced from a 35-item to a 19-item measure. Eleven Items (4, 6, 9, 13, 14, 15, 17, 20, 24, 27, 29) were deleted, as they were not sensitive enough for use as a change measure. Over 80% of the pre-treatment pilot sample answered these questions correctly. Item 16 was removed, as it repeated item 14. Item 5 was removed, as it was formatted differently to other items (open ended), and a four-part question (items 31-34) was combined into one new item on problem solving asking respondents to list the order of problem solving tasks. The wording of one other item was re-worded changing from wife to partner. It was anticipated that these changes, whilst significantly reducing the length of the scale, would not significantly affect its use as an outcome measure of knowledge acquired through the program.

3. The addition of a treatment readiness and treatment performance scale to the design of the pilot would enhance the ability to assess program responsiveness. The semi-structured interviews of treatment readiness and performance developed by Serin (1998) were adapted for use as brief questionnaires. In addition to the Readiness to Change Questionnaire described earlier, each offender was asked to rate his assessment of how he approaches the program and how he felt he performed (see Appendix B). Program facilitators were asked to rate treatment performance for each person completing the program.

Conclusion

The main aim of the first 6 months of the project was to develop and pilot measures, and to perform a pilot study to assess feasibility and to modify protocol for the main study. The pilot work has been successful in meeting these aims.

The norming study provided normative data on the two main measures of anger experience and expression to be used as outcome measures in the program evaluation. These Australian prison norms will enable comparisons to be made between a general prison population and those selected for anger management programs. In addition, this study allowed for the development of a measure of motivation (the Readiness to Change Questionnaire), which has established psychometric properties.

The pilot study demonstrated the utility of the evaluation protocols and measures and provided preliminary data on the outcome measures. As a result of this study, several modifications were made to the design of measures for the main study.

MAIN OUTCOME STUDY

Introduction

The aim of the Main Outcome Study was to determine the effectiveness of the Anger Management Programs delivered in South Australia and Western Australia using a similar methodology to that described in the Pilot Study. In brief, individual offenders referred to the anger management programs in prisons and community corrections in South Australia and Western Australia would be assessed prior to commencing the program and immediately following completion of the program. A control (non-treated) group would be formed of offenders referred, but not yet completing the program ('waiting list controls'). The controls would be assessed on the same measures and at the same points in time (pre/post) as the treatment ('experimental') group. In addition, a proportion of the treatment group would be followed up and re-tested on selected measures at 2 months and at 6 months to determine the extent to which and changes would be maintained over time.

Hypotheses

This study aimed to investigate the following hypotheses:

- 1) That violent offenders undertaking Anger Management treatments will show greater reductions on measures of anger and aggression than will waiting-list controls.
- 2) That these changes will maintain over a six-month period following cessation of the program.
- 3) That the degree of therapeutic change will be predictable from baseline psychological, criminological and developmental measures. Examples of predictor variables are: a) pre-treatment levels of anger b) pre-treatment commitment to change. It was hypothesised that therapeutic gains will be greater when variable a) and b) are high.
- 4) That Anger Management outcomes, and predictors of these outcomes, would be the same across two different jurisdictions: South Australia and Western Australia. This last hypothesis is an important one for the project. The comparison of the two states allows for an assessment of the generalisability and likely replicability of any findings. The demonstration of consistency across the states would make it more likely that the results would be of relevance nationally, or even internationally.

Method

Participants

All participants were male. Participants were drawn from referrals to prison- and community-based Anger Management programs in South Australia and Western Australia. In total, 418 offenders participated in the main study, of whom 86% were from prison-based Anger Management Programs.

The mean age of the total sample was 28.8 years ($SD = 8.4$; range: 18-62). Sentences being served ranged from one month to 26 years and 4 months. Sixty-five percent of the sample described their ethnic background as Australian/New Zealand, 19% as Aboriginal or Torres Strait Islander (ATSI), and 16% as European/Asian/Other. Fifty-nine percent of the sample described themselves as single, 30% defacto/married, and 11% divorced/separated/widowed. The mean level of education was 9.4 years ($SD = 2.2$). Occupational descriptions were 56% unemployed/unskilled, 35% semi-skilled/tradesperson, and 9% professional. Seventy-three percent of the participants had not previously completed an anger management course, 20% had completed one anger management course, 4% had completed more than one anger management course, and 3% had attempted but did not complete, an anger management course.

Scores on the risk assessment measure (VOTRAS) indicated that both prison (mean = 13.3; $SD = 6.7$) and community samples (mean = 12.6; $SD = 5.7$) could be considered high-risk groups. Fourteen percent of offenders had committed non-violent offence, 42% violence without bodily harm, 30% violence with bodily harm, 8% grievous bodily harm, and 6% injuries causing death. Thus, the vast majority of referrals were offenders with convictions for violence.

Of the 418 initial participants, 285 completed the post-intervention assessment, and 86% of this latter group were from prison-based Anger Management programs. The mean age of the post-test sample was 29 years ($SD = 8.6$; range: 18-62). Sixty-five percent described themselves as Australian/New Zealand, 17% ASTI, and 18% European/Asian/Other. Sixty percent described themselves as single, 29% defacto/married, and 11% divorced/widowed/separated. The mean level of education was 9.5 years ($SD = 2.2$). Fifty-seven percent described themselves as unemployed/unskilled, 35% semi-skilled/tradesperson, and 8% professional. Seventy-three percent of the sample had not completed an anger management course, 20% had completed one, 5% had completed more than one, and 2% had attempted, but did not complete, an anger management course.

The 2-month follow-up assessment was completed by 78 of the 418 original participants, and 93% of this sample was from prison-based Anger Management programs.

Twenty-one participants completed the 6-month follow-up assessment and this entire sample was from prison-based Anger Management programs.

Western Australian Sample

In total, 203 offenders participated in the main study from WA, with 96% of participants recruited from prison Anger Management programs. The mean age of the WA sample was 29 years (SD= 9.3; range: 19-59). Sixty-three percent described their ethnicity as Australian/New Zealand, 23% ATSI, and 14% European/Asian/Other. Fifty-eight percent described themselves as single, 33% defacto/married, and 9% divorced/separated/widowed. The mean level of education was 9.4 years (SD = 1.9). Forty-eight percent described themselves as unemployed/unskilled, 44% semi-skilled/tradesperson, and 8% professional. Seventy-two percent had not completed a previous anger management course, 20% had completed one, 4% had completed more than one, and 4% had attempted, but not completed, an anger management course.

South Australian Sample

In total, 215 offenders participated in the main study from SA, with 75% of participants from prison-based anger management programs. The mean age was 29 years (SD = 8.6; range: 18-62). Sixty-six percent of participants described their ethnicity as Australian/New Zealand, 15% ATSI, and 19% European/Asian/Other. Fifty-nine percent described themselves as single, 29% defacto/married, and 12% divorced/separated/widowed. The mean level of education was 9.4 years (SD = 2.5). Sixty-five percent described themselves as unemployed/unskilled, 27% semi-skilled/tradesperson, and 8% professional. Seventy-three percent had not completed an anger management course, 20% had completed one, 5% more than one, and 2% had attempted, but did not complete, an anger management course. The substantial reduction in numbers completing the post treatment assessments was due to a number of factors, including transfers to other settings, unavailability for assessment and incomplete assessments due to missed items.

Forty-five participants completed the two-month follow-up, with 96% of this sample from prison-based anger management programs. The mean age was 29.7 (SD = 8.2; range: 18-55). Sixty-five percent described their ethnicity as Australian/New Zealand, 9% ATSI, and 26%

During the administration of the Pre-Intervention Assessment Package, the researcher introduced the Participant Integrity Checklist to one intervention participant (randomly selected) and asked them to be responsible for completing it for each session of the program. It was explained that the Participant Integrity Checklist for each session would be left with the facilitator with an accompanying envelope. On completion of the session they were to ask the facilitator for the appropriate checklist for the session and an envelope. The participant was then required to complete the checklist and return it to the facilitator in the sealed envelope. The facilitator was informed that at the end of the program the researcher would collect the completed Participant Integrity Checklists.

Immediately following the completion of the Pre-Intervention Assessment Package with the participants, the researcher recruited two Correctional Officers (for prison participants) or one staff member (for Community Correctional participants). The recruited staff were asked to complete the Staff Rating Scale for Aggression (Appendix B) for each participant. This involved the recruited staff completing the scale, by thinking about the last week and last month, and recording any aggressive behaviour. It was emphasized that all aggressive behaviours should be recorded, even if they seemed minor or happened only once. In addition, participants were asked to report on their own behaviour over the same reporting period. Previous work has indicated that self and observer ratings of aggression are highly correlated when spouses report on partner behaviour (Mayne & Ambrose, 1999), but it is unclear whether this finding would be replicated in offender-correctional officer observations.

Before the program, the researcher also introduced and explained the completion requirements for the Attendance Record (Appendix E), Participant and Facilitator Integrity Checklist (Appendix E), and the Serin Treatment Performance Scale-Facilitator Report (Appendix E) to the facilitator/s.

The researcher, using the Justice Information System or equivalent, completed the Risk Assessment (Appendix B) for each participant. In addition, the number and severity of incidents and charges for the six months before the program were documented.

During Intervention

The intervention participants completed the Anger Management/STAC program (sessions 1-10). At the completion of each session, the facilitators completed the Integrity Sheet (Appendix E) and recorded attendance on the Attendance sheet (Appendix E).

Post-intervention

During the week following the completion of the Anger Management/STAC program, the intervention and control participants completed the Post-Intervention Assessment Package. To ensure consistency between measures, the Post-Intervention Assessment Package was administered in the same format, which was individual-group, as the Pre-Intervention Assessment Package. The Post-Intervention Assessment Package included:

- Readiness to Change Questionnaire (RCQ) (Appendix B)
- Serin Treatment Performance Scale-Self-rating (STPS-SR) (Intervention participants only) (Appendix B)
- Modified Watt Anger Knowledge Scale (WAKS) (Appendix B)
- Client Self-rating Form (MOAS) (Appendix B)
- State-Trait Anger Expression Inventory (STAXI)
- Novaco Anger Scale (NAS-PI)

Following the completion of the Post-Intervention Assessment Package, the researcher recruited two Officers (for prison participants) or one staff member (for Community Correctional participants) to complete the Staff Rating Scale for Aggression for each participant. As previously described, this involved the staff member(s) thinking about the last week and last month and recording any aggressive behaviour, even if it seemed minor or happened only once.

The researcher would collect the Attendance Record, Participant and Facilitator Integrity Checklists and the Serin Treatment Performance Scale-Facilitator-report (For Intervention participants only) from the facilitator/s at the completion of the Anger Management Program.

Two month Follow-up

Two months following the completion of the AM/STAC Program the Two-month Follow-up Assessment was completed for the randomly selected and available Intervention participants. This involved the completion of the Two-Month Follow-up Assessment Package in the previously described conditions:

- Readiness to Change Questionnaire (RCQ) (Appendix B)
- Modified Watt Anger Knowledge Scale (WAKS) (Appendix B)
- Client Self-rating Form (MOAS) (Appendix B)
- State-Trait Anger Expression Inventory (STAXI)
- Novaco Anger Scale (NAS-PI)

Following the completion of the Two-month Follow-up Assessment Package, the researcher recruited two Officers (for prison participants) or one staff member (for Community Correctional Participants) to complete the Staff Rating Scale for Aggression for each participant. As previously described, this involved the staff member(s) thinking about the last week and last month and recording any aggressive behaviour, even if it seemed minor or happened only once.

Six-month Follow-up

Six months following the completion of the AM/STAC Program, the Six-month Follow-up Assessment was completed for the Intervention participants. This involved the completion of the Six-Month Follow-up Assessment Package, in the format previously described, and included:

- Readiness to Change Questionnaire (RCQ) (Appendix B)
- Modified Watt Anger Knowledge Scale (WAKS) (Appendix B)
- Client Self-rating Form (MOAS) (Appendix B)
- State-Trait Anger Expression Inventory (STAXI)
- Novaco Anger Scale (NAS-PI)

Following the completion of the Six-month Follow-up Assessment Package, the researcher recruited two Officers (for prison participants) or one staff member (for Community Correctional

Participants) to complete the Staff Rating Scale for Aggression for each participant. As previously described, this involved the staff member(s) thinking about the last week and last month and recording any aggressive behaviour, even if it seemed minor or happened only once.

In addition, the researcher, using the Justice Information System or equivalent, collected data regarding the number and severity of incidents and charges for the six months following the program.

Summary of research methodology

The research project has two (related) components. First, an evaluation of the changes brought about by the intervention and their robustness over time. Second, the prediction of such changes from pre-treatment measures.

Intervention (treatment) participants completed measures immediately before and immediately after the Anger Management Program.

Control participants were selected from the same pool as those forming the intervention group, but were on the waiting list to begin the program rather than engaged in the program. Thus, participants were not randomly allocated to intervention/control conditions. Pilot work has shown that total random allocation is rarely feasible, for logistical reasons, in correctional systems but that the waiting list controls are comparable and equivalent in their characteristics to the intervention group (Watt & Howells, 1999). Control participants were tested at the same points in time as the intervention group (pre/post) and were offered the program, some weeks later.

The Anger Management Programs

A description of the objectives of the South Australian and Western Australian Anger Management programs is outlined in Appendix A. The Anger Management programs have high correlative face validity. That is, in both states the program:

- 1) is based on a cognitive approach to behaviour change;
- 2) runs for 10 sessions, each session for 2 hours duration;
- 3) uses trained facilitators to conduct the programs with groups of offenders;
- 4) uses a Treatment Manual based on material developed in New Zealand and derived from Novaco's framework (Novaco, 1975, 1997);

- 5) content includes structured exercises focussing on skills, such as identifying provocations, relaxation, cognitive restructuring, assertion, and relapse prevention.

Measures: further details

A battery of measures had been developed in previous work in Western Australia (Watt, 1997; Watt & Howells, 1999). The battery included pre- and post- program measures of trait/state anger and anger expression (Spielberger scales); measures of the provocational, cognitive, physiological, and behavioural components of anger (Novaco Anger Scale); a measure of knowledge of anger management skills (Watt Anger Knowledge Scale); staff ratings of aggressive and violent behaviour; and officially recorded incidents and rule infractions. The reliability, internal consistency, and intercorrelations of these measures within a violent offender population has been established (Watt, 1997; Watt & Howells, 1999).

The first phase involved collaborating with the Programs Standards Committee in DCS in South Australia and with the MOJ in Western Australia, to further refine such measures, and pilot them across both jurisdictions.

Outcome Measures

The use of self-report measures has characterised research in the area of anger and violence. Novaco (1994) argued that self-report measures potentially offer the most useful method of measuring anger, which was defined as an essentially subjective reaction to a provocative event or stressful circumstance. Whilst some have criticised the use of such measures on the grounds that respondents may attempt to present themselves in a positive light (Saunders, 1991), such effects can be minimised by controlling for response biases. Two self-report measures of anger were chosen for this study, the Spielberger State-Trait Anger Expression Inventory and the Novaco Anger Scale. The Spielberger scale was designed to distinguish between state and trait anger and measure anger expression and control. Whereas, the Novaco scale assesses the intensity of anger, and reactions to different type of provocations. Both measures have been used extensively in previous research (eg. Forgays, Spielberger, Ottaway & Forgays, 1998; Greenglass & Burke, 2000; McMurrin, Richardson, Street, Ahmadi, & Cooper, 2000; O'Neill, 1995; Smith, Smith & Beckner, 1994; Swaffer & Epps, 1999).

The State-Trait Anger Expression Inventory (STAXI; Spielberger, 1991)

It is a 44 item self-report measure of the experience and expression of anger. The STAXI comprises six scales

- State Anger, measuring the individual's present feeling of anger;
- Trait Anger, measuring differences in the dispositional experience of anger;
- Anger-In, measuring the frequency of suppression of angry feelings;
- Anger-Out, measuring the frequency of external expression of anger;
- Anger Control, measuring the frequency of attempts to control the expression of anger;
- Anger Expression, a general index of the frequency of the expression of anger with two subscales of Trait Anger;
- Angry Temperament, measuring the general tendency to experience and express anger, without provocation;
- Angry Reaction, measuring the general dispositional experience of anger when provoked.

The STAXI has been widely used by researchers to measure anger in offender populations (Dalton, Blain, & Bezier, 1998; Myers & Monaco, 2000; Slaton, Kern & Curlette, 2000). Spielberger's (1988) conceptualisation of anger distinguishes between anger as an emotional state (state anger) and anger as an enduring propensity to become angry (trait anger). He later distinguished between the experience of anger (control and effects) and the expression of anger (inward or outward). The scale has acceptable reliability and validity (Spielberger, 1991) and the factor structure of the anger expression scales has been confirmed in a prison population (Kroner & Reddon, 1992). Cornell et al. (1999) suggested that trait anger and anger-out should be most predictive of aggression. McDougall, et al., (1991) related scale scores to staff ratings of anger to demonstrate the validity of the trait scale.

The Novaco Anger Scale (NAS-PI; Novaco, 1994)

The Novaco Anger Scale (NAS-PI) was developed because of perceived lack of theoretical basis for instruments measuring anger (Novaco, 1994). The NAS-PI is a 121 item self-report scale, comprising three parts. Part A contains 60 items that measure cognitive, arousal, and behavioural elements of anger. The three clinically oriented scales each have four subscales. The cognitive domain can be broken down into attentional focus, suspicion, rumination and hostile attitude. The arousal domain measures intensity, duration, tension and irritability. Finally, the behavioural domain measures impulsive reaction, verbal aggression, physical confrontation and indirect experience. Responses are made on a 3-point scale (never true-always true) reflecting how representative the item is. Part B contains 27 items measuring anger intensity and generality across a range of anger-provoking situations and has five subscales: disrespectful treatment, unfairness/injustice, frustration/interruption, annoying traits and irritations. Responses range from not at all angry to very angry on a 4-point scale to measure the intensity of anger whereby a higher score indicates greater anger.

The most obvious limitation of using self-report measures in this type of evaluation is their vulnerability to bias (Mayne & Ambrose, 1999). One source of bias, which may be particularly relevant for offender populations, is the possibility of socially desirable responses. Socially desirable response bias occurs when respondents give answers that they think will be socially accepted. Arias and Beach (1987) reported that social desirable responding was associated with a lower rate of admitting aggression towards partners. Saunders (1991) describes techniques for adjusting self-reports of violence to reduce this bias, including the use of a social desirability measure as a covariate in ANOVA designs. The Novaco has, as its part C, a measure to detect socially desirable responses. It contains 32 true/false items that measure social desirability. The scales have been shown to have acceptable test-retest reliability (Cronbach alpha $>.86$) (Mills, et al., 1998; Novaco, 1994) and consistent concurrent validity within an adult prison population (Mills, et al., 1998).

The Modified Watt Anger Knowledge Scale (WAKS; Watt & Howells, 1999)

The WAKS is a 35-item questionnaire (Watt & Howells, 1999) designed to measure knowledge of techniques for effectively dealing with anger. The questionnaire was developed in collaboration with facilitators running anger management programs in prisons, and items relate

closely to material covered in the programs. Participants are asked to choose one of four possible responses to each question. Summing the total correct responses provides a composite score ranging from 0 to 35 with higher scores indicating greater anger knowledge. Watt and Howells (1999) have used the scale with offenders completing anger management programs.

For the main study, the WAKS was shortened to 19 items (Appendix B). The process of modification involved the original WAKS being administered to the pilot group and the items most sensitive to treatment effects being retained.

The Readiness to Change Questionnaire (RCQ; Heather & Rollnick, 1993).

The RCQ is a 12-item questionnaire based on Prochaska, DiClemente and Norcross' (1992) stages of change model designed to identify the stage of change reached by an excessive drinker of alcohol. The three stages identified by the measure, Pre-contemplation, Contemplation, Action, were subsequently augmented with a fourth stage, Determination, in a follow up study by Bubner (1999) giving a total of 16 questionnaire items. To adapt the measure for use with anger problems, the wording of each item was changed from alcohol to anger (eg 'My drinking is a problem sometimes' changed to 'My anger is a problem sometimes').

Serin Treatment Readiness Scale (STRS; adapted from Serin & Kennedy, 1997)

The STRS is an 11-item scale. Participants are asked to choose the most applicable response regarding their anger and management of their anger on a 5-point scale, ranging from strongly disagree to strongly agree. The STRS was adapted from the Treatment Readiness Rating Scale (TRRS; Serin & Kennedy, 1997), which is administered in an interview format. Each question of the STRS addressed the main two components of each construct of the TRRS. Questions were formulated using face validity of content.

Serin Treatment Performance Scale – Self-report (STPS-SR; adapted from Serin & Kennedy, 1997)

The STPS-SR is a 15-item scale. Participants are asked to choose the most applicable response regarding management of anger on a 5-point scale, ranging from strongly disagree to strongly agree. The STPS-SR was adapted from the Treatment Responsivity Rating Scale (TrRS) (Serin & Kennedy, 1997), which is administered in an interview format. Each question of the

STPS-SR addressed the main two components of each construct of the TrRS. Questions were formulated using face validity of content.

Serin Treatment Performance Scale – Facilitator report (STPS-FR; adapted from Serin & Kennedy, 1997)

The STPS-FR is a 15-item scale, with content identical to the STPS-SR. The only difference between the two measures is the STPS-FR was written in the third person. Similarly, facilitators were asked to choose the most applicable response regarding management of the participant's anger on a 5-point scale, ranging from strongly disagree to strongly agree.

The Modified Overt Aggression Scale (MOAS; Kay, Wolkenfeld, & Murrill, 1988)

The Modified Overt Aggression Scale (MOAS) (Kay et al., 1988) is an attempt to accurately and reliably measure aggression. The MOAS is a modification of a scale developed by Yudofsky, Silver, Jackson, Endicott & Williams (1986) designed to assess four categories of aggression in psychiatric patients: verbal aggression, aggression against property, auto-aggression, and physical aggression. A total of 16-items, four items from each category, are rated on a five-point scale representing increasing levels of severity. These scores are then weighted to give a total score reflecting the overall seriousness of the aggression. The scale is reported to have acceptable internal reliability, inter-rater reliability, and validity (Kay et al., 1988) and has been used in a prison setting with prisoner behaviour being rated by correctional officers (Watt & Howells, 1999).

The MOAS was adapted to form the Client Self-Rating Form and Staff Rating Scale of Aggression. Both scales retained the four categories of aggression as previously mentioned. A total of 20-items, five items for each category. The raters (participant or staff) were asked to rate the highest level of aggressive behaviour engaged for each category over the last week and last month. The highest possible score in each category are summed to give a total score reflecting aggression.

The Violent Offender Treatment Programme Risk Assessment Scale (VOTRAS; Ward & Dockerill, 1999)

The VOTRAS is a screening measure designed to estimate risk in violent offenders. The scale comprises seven risk factors:

- Current violent offence
- Most serious offence
- Previous violent offences
- Previous non-violent offences
- Age at first arrest
- Alcohol related offences
- Other drug misuse

In two subscales:

- Level of harm
- Probability of re-offence

Each item can be rated from archival material and is weighted. The scale is an acceptable predictor of violent re-offence (Ward & Dockerill, 1999).

Incidents and Charges

Previous attempts to develop behavioural measures of aggression have centred around staff incident reports and disciplinary infractions. In Western Australia and South Australia staff are required to record all incidents of abusive or threatening behaviour, assault, and/or self-injury (eg. DCS System Operating procedure No. 4, 1996). Official recording systems have been criticised for only measuring the most consequential acts of violence (Yudofsky, et al., 1986) and underestimating aggressive behaviour by ignoring incidents of verbal abuse, aggressions against property and unsuccessful attempts at violence (Arboleda-Florez, Crisanti, Rose, & Holley, 1994). Even so, McDougall, et al., (1987) demonstrated that such reports could be used to measure the effectiveness of anger management programs. Therefore, data was collected for all research participants in the period six months before and six months after the program from the Justice Information System or equivalent.

Program Integrity

Program integrity refers to the extent to which a program is delivered in practice as intended in theory and design (Hollin, 1995). Waltz, Addis, Koerner, & Jacobsen (1993) suggested that assessing integrity includes the therapist's adherence to the treatment protocol and competence in delivery of treatment. As a measure of treatment adherence, tasks from the program manuals were developed into a checklist for completion by a program facilitator from each group. Checklists were then adapted for completion one participant in each administered program.

Whilst facilitators are likely to have some biases in their perceptions of sessions, and clients may not have the level of knowledge required to assess integrity accurately, these sources of data are commonly utilised in checking for integrity (Moncher & Prinz, 1991).

For both checklists, items were coded according to five separate categories (items relating to timing, content, group processes, exercises, and discussion). Items relating to timing (whether or not time allocated was adhered to) were rated on a 3-point scale (yes, no, unsure). All other ratings were on a 4-point scale, which assessed degree to which the task was complied with (not at all, partial, thorough, unable to rate). Two independent raters coded each item according to these criteria. Correlations revealed acceptable levels of inter-rater reliability for the four scales (MOJ Facilitator Scale .89, MOJ Client Scale .81, SA Facilitator .83, SA Client .87). All correlations were significant at the .01 level. Disagreements were then recoded to reach a final coding for each item. The frequency of items occurring in each of the five categories as measured by the Integrity Checklists can be found in Table 11.

Table 11. Program Integrity Ratings

Coding	SA Manual		WA Manual	
Timing	62	26%	44	23%
Content	104	44%	57	30%
Group process	14	6%	16	9%
Exercise	43	18%	43	23%
Discussion	14	6%	28	15%
Total	237		188	

Results

The results section is divided into two parts. In the first part, descriptive statistics and correlations are given. In the second part, the research aims are addressed through statistical analyses.

1. Correlational Analyses

The correlation matrices for the pre-test measures were examined to assess the extent to which variables were related. A preliminary check for multicollinearity and singularity was made by examining the bivariate correlations of all the variables. Tabachnick and Fidell (1996) advise caution if two variables are correlated above .70 in the same analysis. All but one of the bivariate correlations were below .70. The one that correlated higher than .70 could be expected to do so by chance. The correlation matrices are presented in Tables 12 - 14.

Where not all items in a scale or test were completed or where the responses were ambiguous or difficult to score, that subject was omitted from that particular correlation. Thus there are variations in the numbers for each correlation.

Table 12. Correlation^a of pre-test STAXI Subscales with NAS Subscales on pre-test sample

Variable	State anger	Trait anger	Anger in	Anger out	Anger control	Anger expression
Angry Cognitions	.29**	.55**	.40**	.45**	-.32**	.57**
Anger Arousal	.31**	.59**	.40**	.43**	-.38**	.59**
Angry Behaviour	.25**	.64**	.35**	.59**	-.43**	.67**
Anger Regulation	-.09	-.22**	-.12*	-.27**	.44**	-.43**
Total NAS	.31**	.64**	.41**	.53**	-.41**	.66**
Provocations (disrespectful treatment)	.24**	.64**	.35**	.41**	-.29**	.50**
Provocations (unfairness)	.22**	.57**	.35**	.40**	-.18**	.45**
Provocations (frustration)	.22**	.59**	.32**	.40**	-.27**	.48**
Provocations (annoying traits of others)	.27**	.58**	.32**	.39**	-.24**	.46**
Provocations (irritability)	.26**	.58**	.36**	.45**	-.33**	.55**
Total PI	.67**	.60**	.33**	.39**	-.29**	.49**

Note. ^atwo-tailed; * $p < .05$; ** $p < .01$; NAS = Novaco Anger Scale; PI = Provocation Inventory
Ns ranged from 280 to 369

Table 13. Correlation^a of pre-test NAS Subscales with other measures for pre-test sample

Variable	WAKS	RCQ	Serin	VOTP RISK	MOAS (self) week	MOAS (self) month
Angry cognitions	-.17*	-.03	.09	.20**	.27**	.23**
Anger arousal	-.11**	-.03	.18**	.18**	.26**	.21**
Angry behaviour	-.18**	-.03	.15**	.23**	.33**	.30**
Anger regulation	.23**	.02	.01	.03	-.20**	-.17**
Total NAS	-.17**	-.03	.16**	.22**	.31**	.27**
Provocations (disrespectful treatment)	-.09	-.00	.19**	.13*	.20**	.15**
Provocations (unfairness)	.00	.00	.18**	.13*	.22**	.20**
Provocations (frustration)	-.07	-.02	.20**	.16*	.25**	.19**
Provocations (annoying traits of others)	.04	-.01	.25**	.17**	.25**	.21**
Provocations (irritability)	-.13*	-.01	.19**	.14*	.27**	.22**
Total PI	-.08	-.04	.19**	.16**	.25**	.21**

Note. ^atwo-tailed; * $p < .05$; ** $p < .01$; WAKS = Watt Anger Knowledge Scale; RCQ = Readiness to Change Questionnaire; Serin = Serin Treatment Readiness Scale; VOTP=Risk Assessment Scale; MOAS = behavioural rating scale

Ns ranged from 253 to 418

Table 14. Correlation^a Matrix of other Variables for pre-test sample

Variable	1	2	3	4
1. VOTP	-			
2. WAKS	-.05	-		
3. MOAS week	.19**	-.19**	-	
4. MOAS month	.16**	-.11*	.83**	-
5. SCQ	-.07	.03	-.02	.00

Note. ^atwo-tailed; * $p < .05$; ** $p < .01$; VOTP=Risk Assessment Scale; WAKS = Watt Anger Knowledge Scale; MOAS = behavioural rating scale; RCQ=Readiness to change Questionnaire
Ns ranged from 309 to 418

A number of interesting and statistically significant correlations are apparent in Tables 12 to 15. As might be expected, many of the NAS Scores correlate positively with the STAXI scores. Both tests were constructed to assess similar constructs, (various aspects of anger experience). Significant positive correlations suggest convergent validity of the two tests.

The NAS Scores allow for the most detailed analysis of anger and correlations between subscales and other measures are summarised in Table 13. It is clear that high anger tends to be associated with low anger knowledge (WAKS), with high self-reported anger incidents in real life (MOAS) and with high statistical risk of recidivism (VOTP).

Descriptive Statistics

Tables 15 - 17 show pre-test scale ranges, means and, standard deviations for outcome measures, broken down to show differences between states, differences between ethnicity and differences between stage of sentence served. A series of Analyses of Covariance, comparing scores on the above measures, using social desirability scores as a covariate, revealed that some differences between the above measures were statistically significant (see Appendix F)

Many of the differences between SA and WA achieve statistical significance. This confirms the findings of the Normative Study that anger levels are higher in SA samples.

Many differences are statistically significant in the comparisons for ethnicity. Again, as in the Normative Study, there is a tendency for ATSI to have higher anger scores.

Broardly, the same often was found for stage of sentence as occurred in the Normative Study but many results failed to achieve statistical significance.

Table 15. Mean Scores on Pre-test Measures for South Australian and Western Australian Subjects (total pre-test sample)

Outcome variable	South Australia		Western Australia	
	mean	sd	mean	sd
<u>STAXI</u>				
State anger	12.89	5.5	11.48	3.6
Trait anger	20.22	7.5	16.98	6.1
Anger in	16.86	4.7	16.52	4.9
Anger out	16.62	5.2	15.83	4.3
Anger control	20.31	5.9	21.26	5.9
Anger expression	29.16	11.4	27.13	9.9
<u>Novaco scales</u>				
Angry cognitions	31.75	5.8	29.76	5.2
Anger arousal	30.05	7.0	28.51	7.5
Angry behaviour	30.09	6.9	27.79	6.2
Anger regulation	24.62	4.1	24.79	4.7
Novaco anger scale total	92.04	18.2	85.99	17.3
Provocations (disrespectful treatment)	12.00	3.5	11.20	3.5
Provocations (unfairness)	15.64	4.2	14.31	4.2
Provocations (frustration)	14.94	4.2	13.99	4.3
Provocations (annoying traits of others)	12.14	4.1	10.82	3.7
Provocations (irritability)	13.92	4.1	12.83	4.1
Total provocation inventory	68.64	18.6	63.15	18.0
<u>Other measures</u>				
Watt Anger Knowledge Scale	8.35	3.9	8.45	3.7
Serin Readiness	39.39	6.6	40.76	6.5
Readiness to change category	2.38	0.7	2.46	0.7
Risk total	16.77	5.2	9.34	5.7
MOAS self-rated weekly	1.81	2.9	0.62	1.2
MOAS self-rated monthly	2.37	3.4	1.03	2.0

Note: *N*s range from 197 to 217 for South Australian sample and from 165 to 201 for Western Australian sample.

Table 16. Mean Scores on pre-test Measures of Aboriginal and Non-Aboriginal Offenders (total pre-test sample)

Outcome variables	Australian/ New Zealand		Aboriginal/ Torrens Strait Islander		Other (inc Asian European)	
	mean	sd	mean	sd	mean	sd
STAXI						
State anger	12.29	4.8	12.35	5.1	11.85	3.8
Trait anger	18.53	7.1	18.97	7.4	18.97	6.7
Anger in	16.52	4.8	17.16	5.0	16.89	4.8
Anger out	16.32	4.8	16.27	4.9	16.01	4.6
Anger control	21.30	6.1	19.17	5.4	20.39	5.8
Anger expression	27.63	11.0	30.16	9.6	28.45	11.0
Novaco scales						
Angry cognitions	30.81	5.4	30.80	6.1	31.03	6.2
Anger arousal	29.15	7.6	30.72	6.8	28.67	6.6
Angry behaviour	29.31	6.6	29.08	6.0	27.95	7.7
Anger regulation	24.63	4.1	24.95	4.4	24.80	5.2
Novaco anger scale total	89.33	18.1	90.59	17.5	87.63	19.2
Provocations (disrespectful treatment)	11.58	3.7	12.54	2.9	10.89	3.3
Provocations (unfairness)	14.97	4.3	16.19	4.3	14.14	4.0
Provocations (frustration)	14.41	4.4	15.62	3.9	13.70	4.6
Provocations (annoying traits of others)	11.50	4.1	12.10	3.5	11.08	4.1
Provocations (irritability)	13.63	4.3	14.56	3.3	12.36	4.0
Total provocation inventory	65.76	19.3	71.23	15.7	62.16	18.1
Other measures						
Watt Anger Knowledge Scale	8.86	3.9	7.25	2.9	7.97	3.9
Serin Readiness	40.30	6.9	40.50	5.9	38.57	6.0
Readiness to change category	2.37	0.7	2.62	0.6	2.37	0.8
Risk total	13.30	6.6	12.61	6.9	13.70	6.2
MOAS self-rated weekly	1.16	2.3	1.39	2.3	1.30	2.4
MOAS self-rated monthly	1.70	2.9	1.70	2.7	1.80	2.9

Note: Ns range from 234 – 266 for ANZ, from 64 to 77 for ATSI and 60 to 75 for Others.

Table 17. Comparison of Measures for Stage of Sentence (total pre-test sample)

	Less than 25% served		More than 25% served	
	mean	sd	mean	sd
<u>STAXI</u>				
State anger	11.77	4.3	12.41	4.8
Trait anger	17.34	6.3	19.43	7.0
Anger in	16.62	4.9	17.19	4.9
Anger out	16.41	5.0	16.16	4.5
Anger control	21.64	5.8	19.94	5.9
Anger expression	27.48	10.5	29.59	10.7
<u>Novaco scales</u>				
Angry cognitions	30.20	5.6	30.50	5.6
Anger arousal	28.50	6.8	29.21	7.5
Angry behaviour	28.20	6.7	29.10	6.4
Anger regulation	24.95	4.6	24.33	4.7
Novaco anger scale total	86.85	17.7	88.96	17.8
Provocations (disrespectful treatment)	11.61	3.6	11.70	3.4
Provocations (unfairness)	15.07	4.3	14.68	4.1
Provocations (frustration)	14.73	4.6	14.34	4.3
Provocations (annoying traits of others)	11.56	4.1	11.53	4.1
Provocations (irritability)	13.36	4.4	13.53	4.1
Total provocation inventory	66.32	19.0	65.77	18.5
<u>Other measures</u>				
Watt Anger Knowledge Scale	8.51	3.7	8.69	4.3
Serin Readiness	40.15	7.5	41.14	6.1
Readiness to change category	2.34	0.8	2.54	0.7
Risk total	11.09	6.5	13.51	6.3
MOAS self-rated weekly	0.91	1.9	0.92	1.7
MOAS self-rated monthly	1.39	2.6	1.34	2.3

Note: Ns range from 117 to 149 for -25% group and from 113 to 121 for +25 group

2. Analyses

2.1 Evaluating treatment effects on all key outcome measures

To evaluate the effectiveness of the anger management program, the change from “pre-treatment” to “post-treatment” was compared for those prisoners receiving the anger management program (the treatment group) and the waiting list prisoners (the control group) for all key outcome variables. The mean scores for each outcome measure taken prior to and after the program as well as the mean change, for both the treatment and control groups, are reported in Table 18 and graphical displays of the means are shown in Appendix G.

Table 18. (Adjusted) Means for control and treatment groups at pre-treatment and post-treatment

Dependent variable	Group	Means (SDs)				Pre-post mean change
		Pre-treatment		Post-treatment		
		m	sd	m	sd	
STAXI						
State anger	C	11.2	3.3	11.9	4.4	+ 0.7
	T	12.4	4.9	12.1	5.6	- 0.3
Trait anger	C	17.2	6.4	17.1	6.7	- 0.1
	T	19.3	7.4	17.6	6.5	- 1.7
Anger In	C	16.1	4.6	15.7	4.5	- 0.4
	T	17.1	4.8	15.7	4.5	- 1.4
Anger Out	C	16.4	4.9	14.7	4.1	- 1.7
	T	16.3	4.7	15.3	3.9	- 1.0
Anger Control	C	22.5	5.5	23.0	6.1	+ 0.5
	T	20.7	5.7	22.2	6.1	+ 1.5
Anger Expression	C	26.3	11.6	23.4	9.6	- 2.9
	T	28.6	10.7	24.2	10.1	- 4.4
Novaco Scales						
Angry cognitions	C	29.3	5.8	28.4	5.4	- 0.9
	T	30.9	5.7	29.0	5.7	- 1.9
Anger arousal	C	27.9	6.3	26.7	6.6	- 1.1
	T	29.2	7.5	27.9	6.5	- 1.4
Angry behaviour	C	27.7	7.3	26.7	6.9	- 1.1
	T	29.0	6.7	27.0	6.3	- 2.0
Anger regulation	C	25.7	4.4	25.7	4.8	+ 0.0
	T	24.6	4.5	25.5	4.6	+ 0.9

¹ C = control, T = treatment

² For the Novaco scales, the means are adjusted, controlling for social desirability. For other outcome variables, raw mean scores are used.

Note: Ns for controls range from 56 to 66 and for treated from 191 to 217

Table 18. (cont.) (Adjusted) Means for control and treatment groups at pre-treatment and post-treatment

Dependent variable	Group	Means (SDs)				Pre-post mean change
		Pre-treatment		Post-treatment		
		m	sd	m	sd	
<u>Novaco Scales</u>						
<u>(cont.)</u>						
Total Novaco	C	84.9	18.2	81.8	18.0	- 3.1
Anger Scale score	T	89.2	18.4	83.8	17.4	- 5.4
Provocations	- C	10.8	3.6	10.5	3.6	- 0.3
disrespectful treatment	T	11.5	3.4	11.2	3.6	- 0.4
Provocations	- C	14.0	4.5	13.4	4.0	- 0.6
unfairness	T	14.9	4.2	14.4	4.1	- 0.5
Provocations	- C	13.4	5.0	13.0	4.6	- 0.4
frustration	T	14.3	4.2	14.0	4.2	- 0.4
Provocations	- C	10.4	4.0	10.0	4.0	- 0.3
annoying traits of others	T	11.4	4.1	10.9	3.7	- 0.5
Provocations	- C	12.3	4.6	12.3	4.4	- 0.01
irritability	T	13.2	4.0	13.1	4.2	- 0.1
Total	C	60.9	19.6	59.4	19.4	- 1.5
Provocations	T	65.4	18.3	63.5	18.5	- 1.9
Inventory score						
<u>Other measures</u>						
Watt Anger	C	9.08	3.8	10.03	4.0	+ 0.95
Knowledge Scale	T	8.55	3.8	10.35	3.6	+ 1.80
SCQ –	C	-0.98	3.3	-1.21	2.8	- 0.23
precontemplation	T	-1.05	3.3	-0.58	3.1	+ 0.47
SCQ –	C	0.73	3.4	1.53	3.2	+ 0.80
contemplation	T	1.96	3.7	1.85	3.5	- 0.11
SCQ – action	C	2.31	3.3	3.12	3.3	+ 0.81
	T	2.80	3.3	4.32	2.7	+ 1.52
MOAS self-rated	C	1.12	2.4	1.09	2.7	- 0.03
weekly	T	0.95	1.8	0.52	1.4	- 0.41
MOAS self-rated	C	1.48	3.1	1.03	2.2	- 0.45
monthly	T	1.63	2.6	1.26	2.9	- 0.37

¹ C = control, T = treatment

² For the Novaco scales, the means are adjusted, controlling for social desirability. For other outcome variables, raw mean scores are used

Note: Ns ranged for controls range from 50 to 66 and for treated from 187 to 217

The data in Table 18 show a general trend of small improvements on all key outcome measures over time, with the treatment group showing slightly better improvement on most outcomes (except for anger out, anger arousal and the monthly self-rated MOAS). To analyse the reliability of these observed differences between the treatment and control groups, repeated measures Analyses of Variance (ANOVA) were carried out with time of measurement (pre-versus post-) as a within-subjects factor and group (treatment versus control) as a between subjects factor. For the Novaco scales, social desirability was statistically controlled for using covariate analysis, as recommended by Novaco (1994). The hypothesis that the treatment group shows evidence of greater improvement over time compared to the control group is tested by the two-way interaction (time of measurement \times group) in the ANOVA. The time of measurement main effect tests the hypothesis that, on average, both groups show change over time. The interaction effect and the time main effect results of all ANOVAs are shown Appendix H.

The ANOVA results, for all but one outcome variable, showed that the trend for a relative improvement in the treatment group compared to the control group was not statistically significant, and hence could not be taken as evidence for a reliable improvement due to the anger management program. However, the results did show that the prisoners who undertook the anger management program ($M_{change} = 1.80$, $SD = 3.1$) showed significantly greater improvement in anger knowledge than did the prisoners in the control group ($M_{change} = 0.95$, $SD = 2.5$), $F(1, 273) = 4.12$, $p < .05$. The ANOVAs also showed a reliable improvement over time on many of the key outcome measures for both the control and treatment groups, as indicated Appendix H.

Although the data showed a trend for the control group participants to increase their contemplation scores relative to the treatment group, and, conversely, for the treatment group participants to increase their action scores relative to the control group, neither effect was statistically significant. To investigate this result further, participants were classified according to their Readiness to Change category (precontemplation, contemplation, or action) using the Quick Method of Stage Allocation (Miller & Rollnick, 1993) at both “pre-treatment” and “post-treatment”. The frequency and percentage of participants in each Readiness to Change category at “pre-treatment” and “post-treatment” are shown in Table 19, separately for the control and treatment groups.

The table can be interpreted easily by noting that the prisoners represented along the diagonal (top left to bottom right) in each sub-table, are those that showed no change in their Readiness to

Change category. Those above the diagonal represent those who have “improved” their Readiness to Change category by moving to a higher stage of change (e.g., moving from precontemplation to contemplation), whereas those below the diagonal represent those who have “worsened” their stage of change category by moving to a lower stage of change (e.g., moving from action to contemplation).

Table 19. Frequencies (percentages) of prisoners in each Readiness to Change category at pre- and post-treatment for the control and treatment groups

		Post-treatment			
CONTROL GROUP		precontemplation	contemplation	action	total
Pre-treatment	Precontemplation	5 (33%)	0 (0%)	10 (67%)	15
	Contemplation	0 (0%)	7 (64%)	4 (36%)	11
	Action	1 (3%)	10 (28%)	25 (69%)	36
	Total	6 (9%)	17 (28%)	39 (63%)	62

		Post-treatment			
TREATMENT GROUP		precontemplation	contemplation	action	total
Pre-treatment	Precontemplation	10 (35%)	2 (7%)	17 (58%)	29
	Contemplation	3 (5%)	17 (26%)	45 (69%)	65
	Action	4 (3%)	12 (11%)	97 (86%)	113
	Total	17 (8%)	31 (15%)	159 (77%)	207

% in brackets represent the percentage within readiness category at pretest

The above data show that in the control group about half of the participants were in the same Readiness to Change category at “pre-treatment” compared to “post-treatment” (59.5%), and among those that did change stages, the changes were relatively “random” with a similar percentage “improving” (22.6%) as “worsening” (17.7%). On the other hand, in the treatment group, 59.9% remained static, and of those who changed stages, 30.9% “improved” compared to only 9.1% who “worsened”. Hence, undergoing an anger management program appears to result in greater moves towards readiness for change.

2.2 Evaluating the effect of state and prisoner location on treatment effects

The previous findings were expanded upon by evaluating treatment effectiveness, taking the state (WA versus SA) and the location (community located versus prison located) of each prisoner into account. Again, repeated measures Analyses of Variance (ANOVA) were carried out with time of measurement (pre- versus post-) as a within-subjects factor and group (treatment versus control) as a between subjects factor but also including state and location as additional between-subjects factors. The key results are the (higher level) interactions that involve time of treatment and group as well as either or both of state and location. These interaction effects test whether treatment effectiveness varies depending on state and/or location.

For several of the Novaco outcome measures, statistically significant three-way interactions between time of measurement, group and state were obtained providing evidence that treatment effectiveness varied somewhat between WA and SA. Significant three-way (time of measurement by group by state) interactions were found for several of the NOVACO scale anger measures; angry cognitions, $F(1,231) = 9.37, p = .002$, anger arousal, $F(1,231) = 3.28, p = .071$, angry behaviour, $F(1,232) = 5.99, p = .015$, the total Novaco Anger Scale score, $F(1,232) = 6.87, p = .009$, and the provocations measures of disrespectful treatment, $F(1,232) = 4.43, p = .036$.

For each of these significant interactions, “pre-treatment” and “post-treatment” means, for both treatment and control groups in WA and SA, are shown in the top graph of each page in Appendix I. Furthermore, the mean change over time for both treatment and control groups in WA and SA are shown in the lower graph of each page. Note that a negative mean change corresponds to a decrease in mean scores over time and a positive mean change corresponds to an increase in the mean scores over time. These graphs show that, in WA, the treatment group showed greater improvement compared to the control group by reporting lower levels of angry

cognitions, anger arousal, angry behaviour, total Novaco Anger Scale scores, and provocations (disrespectful treatment). By contrast, among the SA prisoners, the treatment group did not differ from the control group, and indeed showed a trend towards greater improvement in the control group, although this difference was not reliably significant.

However, one important consideration here, is that the effectiveness of the anger management program in WA compared to SA, in relation to the Novaco scales, is primarily a function of differences between the control groups in the two states rather than between the treatment groups. Unfortunately, in SA, the control groups showed evidence of improvement of a similar magnitude as that obtained by the treatment groups in both states. Hence, some caution is required in interpreting these state differences.

No other significant interactions were found in any of the ANOVAs. In particular, there was no evidence that location (community located versus prison located) had any reliable effect on treatment effectiveness in relation to any of the key outcome measures.

In summary, evaluation of treatment effectiveness at the completion of the anger management program showed a trend of weak improvements in favour of those prisoners who undertook an anger management program, however, a reliable improvement was detected on only one outcome measure in both SA and WA; namely, knowledge about anger and anger management skills (as measured by the Watt Anger Knowledge Scale). Data also revealed a greater tendency for improvements in stages of change (primarily towards the action stage) among those receiving anger management treatment compared to those who do not. In addition, reliable improvements by those receiving anger management training (in contrast to those not receiving anger management training) in WA only, were found on the Novaco scales measuring cognitive, arousal and behavioural elements of anger as well as on most of the Novaco scales measuring anger intensity across a range of anger provoking situations, although some caution is required in interpreting these latter findings.

2.3 Evaluation of treatment performance

For the treatment group, treatment performance was measured using two scales of treatment performance; STPS-SR (self report by prisoner) and STPS-FR (report by facilitator). Initial analyses of the univariate descriptive statistics showed that the prisoners made significantly higher ratings of their treatment performance ($M = 62.30$, $SD = 7.03$) than did their respective

facilitators ($M = 55.42$, $SD = 7.50$), $t(84) = 6.38$, $p < .001$). Moreover, the distribution of STPS-SR scores showed significant negative skewness ($skew = -0.55$, $SE_{skew} = 0.16$) as shown in Fig. 1. However, a normal distribution was evident in the distribution of STPS-FR scores.

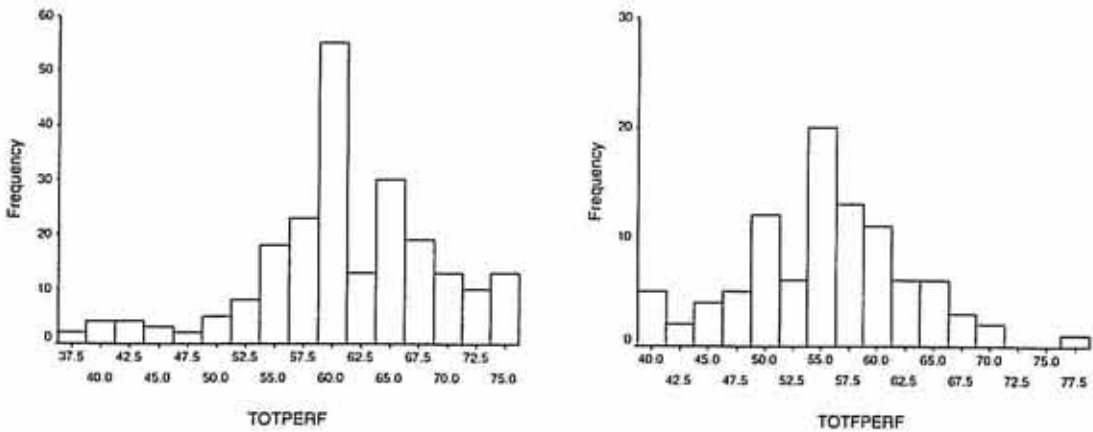


Fig. 1. Histograms showing distribution of STRP-SR (TOTPERF – prisoner self report) and STRP-FR (TOTPERF – facilitator report) scores.

Perhaps more importantly, the correlation between facilitator reports and prisoner self reports was virtually nil ($r(85) = .07$, $p = .56$) showing evidence of almost no correspondence between the facilitator reports and prisoner self reports of treatment performance. Further analyses were carried out on the treatment performance measure but as a result of the low correlation between the two measures (prisoner and facilitator measures), these were carried out separately using the client and facilitator measures of treatment performance. Treatment performance was compared for those who were high on risk, need and motivation with those who were low. Also, treatment performance was compared for those in high and low integrity programs. The results of these analyses are reported later in the report where the effects on treatment effectiveness of risk, need and motivation (section 3.7) and program integrity (section 3.8) are considered in more detail.

2.4 Changes in anger during two and six month follow ups

Some care is needed in interpreting the results of this section. This is because, unlike the comparisons made in the previous section where the treatment group can be explicitly compared to the control group, no follow-up data were available from the control group. This would not have been a problem except that the previous analyses revealed that the control group also

showed improvement on many measures, and for many of the measures, to a similar extent as shown by the treatment group. Therefore, although we can investigate the patterns over time on each outcome, we can not definitively rule out the possibility that the control group would not have shown similar long term improvement or change as shown by the treatment group. On the other hand, we can be more confident about the results for those measures where the treatment group showed a reliable improvement at the end of treatment relative to the control group; that is, on anger knowledge for both SA and WA samples, and on most of the Novaco measures in WA (bearing in mind the previous cautionary comments on this finding).

Due to differential drop out of participants at different stages of the analyses, data was analysed in two stages: (1) for participants from whom data was obtained at two month follow-up (*Ns* varied between 43 and 81, depending on outcome measure), and (2) for participants from whom data was obtained at six month follow-up (*Ns* varied between 10 and 23). Graphs showing mean scores on all key outcome variables at pre-treatment, post-treatment, two-month follow-up and six-month follow-up, are shown in Appendix J.

The longer term effects of anger management treatment were investigated by analysing trends over time using a one-way repeated measures (time of measurement) ANOVA on two samples; (a) using those with complete data at pre-treatment, post-treatment, two month follow-up, and (b) using those with complete data at pre-treatment, post-treatment, two month follow-up, and six month follow-up. Although several qualitatively different types of trends can be evaluated, only two are focussed on here. The first is a linear trend that is indicative of a constant decrease (or increase) in scores over time, and the second is a quadratic trend which is indicative that the trend over time changes. The latter quadratic trend is most likely to reflect an initial improvement from pre-treatment to post-treatment that either weakens over the longer time (i.e., continual improvement that "flattens off" or reduces in intensity) or reverses direction over the longer term (i.e., initial improvement are lost over time). Linear and quadratic trends were evaluated using polynomial contrasts within a repeated measures ANOVA including a between subjects factor, state.

Table 20 shows whether there is any evidence for a linear or quadratic trend in the change in outcome scores over time taken at two and six month follow-up. For anger control, angry cognitions, anger arousal, angry behaviour, total Novaco anger score, unfairness, frustration, annoying traits of others, and action, a significant linear trend (only) was found at two month

follow-up. This indicates evidence that the improvement detected on these measures at post-treatment continues, and hence initial gains are improved upon, at a constant rate, over time. At six month follow-up, evidence was only found on the Novaco provocations subscales of unfairness and frustration for a continuation of the linear trend to six months. However, the small sample size at six month follow up may have contributed to the lack of statistical significance at this point in time.

In addition, quadratic trends were evident for trait anger, anger in, anger knowledge, and the self-rated MOAS (weekly). For trait anger, this corresponded to a constant decrease in trait anger scores from pre-treatment through to two month follow-up ($M_{pre} = 19.6$, $M_{post} = 17.4$, $M_{fu2} = 16.3$) and then a “flattening” (or slight increase) of the effect ($M_{fu6} = 16.6$). Hence, the data provide evidence that the effect gained during treatment continues to about two months after treatment at which point no further improvement occurs but the gains appear to be maintained. In addition, a state by linear trend was evident at two months for trait anger. This finding revealed that there was a constant decrease in trait anger over time ($M_{pre} = 20.0$, $M_{post} = 18.3$, $M_{fu2} = 16.7$) in South Australia but not in Western Australia ($M_{pre} = 16.3$, $M_{post} = 15.8$, $M_{fu2} = 16.0$) where little change was evident.

The quadratic trend for anger in at two month follow-up showed that the improvement from pre-treatment to post-treatment was reversed between post-treatment and two month follow-up ($M_{pre} = 16.8$, $M_{post} = 15.3$, $M_{fu2} = 16.5$). Hence, the improvements in anger in at post-treatment appear to be lost at two month follow-up. However, analysis of the data at six month follow-up showed that this may simply reflect error in measurement and sampling variability because the longer term effect suggested a linear trend. That is, when six month follow-up is taken into account, although there is a general pattern of decrease, increase, and decrease again, the ANOVA results show a significant linear (decreasing) trend ($M_{pre} = 17.2$, $M_{post} = 15.2$, $M_{fu2} = 16.2$, $M_{fu6} = 14.6$).

Strong linear trends and moderate quadratic trends are evident for anger knowledge at both two month and six month follow-up. For both sets of data, the pattern of means show that there is consistent long term improvement in anger knowledge but the level of improvement diminishes with time ($M_{pre} = 8.7$, $M_{post} = 10.6$, $M_{fu2} = 11.1$; $M_{pre} = 8.0$, $M_{post} = 9.7$, $M_{fu2} = 10.3$, $M_{fu6} = 10.4$).

Table 20. Evidence of linear and quadratic trends in outcome changes over time for the treatment group.

Outcome measure	Two-month follow-up sample				Six month follow-up sample		
	<i>N</i>	Linear trend	Quadratic trend	State × linear trend	<i>N</i>	Linear trend	Quadratic trend
State anger	77	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	21	No, <i>p</i> > .10	No, <i>p</i> > .10
Trait anger	73	<i>p</i> = .047	No, <i>p</i> > .10	<i>p</i> = .022	21	No, <i>p</i> > .10	<i>p</i> = .047
Anger in	63	No, <i>p</i> > .10	<i>p</i> = .036	No, <i>p</i> > .10	16	<i>p</i> = .028	No, <i>p</i> > .10
Anger out	69	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	18	No, <i>p</i> > .10	No, <i>p</i> > .10
Anger control	71	<i>p</i> = .006	No, <i>p</i> > .10	No, <i>p</i> > .10	20	No, <i>p</i> > .10	No, <i>p</i> > .10
Anger expression	57	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	15	<i>p</i> = .029	No, <i>p</i> > .10
Angry cognitions	43	<i>p</i> = .010	No, <i>p</i> > .10	No, <i>p</i> > .10	10	No, <i>p</i> > .10	No, <i>p</i> > .10
Anger arousal	43	<i>p</i> = .027	No, <i>p</i> > .10	No, <i>p</i> > .10	10	No, <i>p</i> > .10	No, <i>p</i> > .10
Angry behaviour	43	<i>p</i> = .016	No, <i>p</i> > .10	No, <i>p</i> > .10	10	No, <i>p</i> > .10	No, <i>p</i> > .10
Anger regulation	43	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	10	No, <i>p</i> > .10	No, <i>p</i> > .10
Total NAS score	58	<i>p</i> = .032	No, <i>p</i> > .10	No, <i>p</i> > .10	14	No, <i>p</i> > .10	No, <i>p</i> > .10
Provocations (disrespect)	43	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	10	No, <i>p</i> > .10	No, <i>p</i> > .10
Provocations (unfairness)	43	<i>p</i> = .031	No, <i>p</i> > .10	No, <i>p</i> > .10	10	<i>p</i> = .001	No, <i>p</i> > .10
Provocations (frustration)	43	<i>p</i> = .015	No, <i>p</i> > .10	No, <i>p</i> > .10	10	<i>p</i> = .024	No, <i>p</i> > .10
Provocations (annoying)	43	<i>p</i> = .020	No, <i>p</i> > .10	No, <i>p</i> > .10	10	No, <i>p</i> > .10	No, <i>p</i> > .10
Provocations (irritability)	43	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	10	<i>p</i> = .042	No, <i>p</i> > .10
Total provocations score	59	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	14	No, <i>p</i> > .10	No, <i>p</i> > .10
WAKS	81	<i>p</i> < .001	<i>p</i> = .008	No, <i>p</i> > .10	23	<i>p</i> = .004	<i>p</i> = .041
Precontemplation	77	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	22	No, <i>p</i> > .10	No, <i>p</i> > .10
Contemplation	78	No, <i>p</i> > .10	No, <i>p</i> > .10	No, <i>p</i> > .10	22	No, <i>p</i> > .10	No, <i>p</i> > .10
Action	78	<i>p</i> = .022	<i>p</i> = .078	No, <i>p</i> > .10	23	No, <i>p</i> > .10	No, <i>p</i> > .10
MOAS (weekly)	70	No, <i>p</i> > .10	<i>p</i> = .028	No, <i>p</i> > .10	16	<i>p</i> = .081	No, <i>p</i> > .10
MOAS (monthly)	76	<i>p</i> = .056	No, <i>p</i> > .10	No, <i>p</i> > .10	22	<i>p</i> = .086	<i>p</i> = .069

The data for the self-rated MOAS (weekly) at two month follow-up showed that the improvement from pre-treatment to post-treatment was reversed between post-treatment and two month follow-up ($M_{pre} = 1.17$, $M_{post} = 0.54$, $M_{fu2} = 0.70$), although the scores had not returned to pre-treatment levels. However, similar to the anger in results, this may simply reflect error in measurement and sampling variability because the longer term effect suggested a linear trend (although not statistically significant at the .05 level; $p = .081$). The longer term trend suggested a constant decrease in MOAS (weekly) scores over time ($M_{pre} = 1.69$, $M_{post} = 0.75$, $M_{fu2} = 0.69$, $M_{fu6} = 0.25$).

Finally, of the 23 participants who completed the 6-month follow-up, 10 had incidents recorded against them as at the last testing stage. Of the Western Australian participants ($n=7$) there was no information available for 3 and no pre-test data available. Of the South Australian participants ($n=16$) 5 had pre-test incidents recorded against them – of these 2 did not have post-test incidents (at 6 months). However, 6 participants without pre-test incidents recorded had post-test incidents recorded. It is unclear as to how to interpret this data. Not enough detailed information was recorded about the type of incidents but it seemed that all, except one, could be classified as Type 3 incidents, under the DCS System Operating Procedure No 4 Incident Reporting and Recording criteria. Type 3 incidents are recorded at the discretion of the Prison Officer on duty and could be described as minor incidents. Type 1 incidents are serious incidents. The pre-test incidents (SA) were aggression, aggression against property, prohibited substances, prohibited items, in bed at inspection, and not ready for lockup. The post-test incidents were possession of weapon, prohibited items, no respect regimes, pass letter to visitor, prohibited substances, positive urine, positive test, hoarding medication, assault, escape (Type 1), out of bounds, and not ready for lockup.

In summary, long term effects were found on several of the outcome measures and these either reflected a constant improvement over time (for anger in, anger control, anger expression, angry cognitions, anger arousal, angry behaviour, total Novaco anger score, provocations – unfairness, provocations – frustration, provocations – annoying traits of others, and MOAS weekly) or continual improvement over time but with diminishing improvement over time (trait anger and anger knowledge).

2.5 Analysis of differential dropout of participants

One critical issue in evaluating treatment effects is a consideration of the impact of differential dropout in participants on interpretation of results. This type of analysis is primarily concerned with the question of whether those who drop out of the study (e.g., those who refuse to participate further, those who are released from prison and do not wish to participate further or cannot be located, etc.) are different from those who still remain in the study. If those who dropout do differ from those who remain in the study, and hence those that we draw our conclusions from, then these differences need to be considered in interpreting our final results. In particular, the differences often lead to limiting the generalisation of the results to only those similar to the remaining participants.

The analyses consisted of comparisons between those still in the study and those no longer in the study, at (1) “post-treatment”, (2) two month follow-up, and (3) six month follow-up. The comparisons at “post-treatment” used all subjects and effectively compared those used in subsequent evaluations of treatment effectiveness (both treatment and control groups) and those who only preliminary data was collected from. Comparisons were made on “pre-treatment” measures of all key outcomes as well as other key variables (i.e., Serin Readiness score, risk score). Comparisons at two and six month follow-up compared only those participants initially assigned to the treatment group, and were made on the previously described pre-treatment measures as well as on the change scores of all key outcome measures. All analyses were carried out using independent samples *t*-tests.

At “post-treatment”, the only statistically significant difference found between those measured at “post-treatment” and those who had already dropped out was on the Watt Anger Knowledge scale. The dropouts ($M = 7.66$, $SD = 3.76$) reported lower knowledge about anger management techniques than those used in the study ($M = 8.75$, $SD = 3.74$), $t(416) = 2.76$, $p = .006$. No significant differences were found on any of the “pre-treatment” anger measures.

At two month follow-up, significant differences between dropouts and completers were found on change in irritability (from pre-treatment to post-treatment), and on pre-treatment anger expression, anger regulation, and irritability. Dropouts showed a greater relative improvement in irritability scores (dropouts - $M_{\text{change}} = -0.43$, $SD = 3.3$; completers - $M_{\text{change}} = +0.71$, $SD = 3.6$; $t(235) = 2.34$, $p = .020$), were higher in pre-treatment anger expression (dropouts - $M = 29.7$, $SD = 10.6$; completers - $M = 26.4$, $SD = 10.5$; $t(183) = 2.12$, $p = .036$), and lower in anger control (dropouts - $M = 23.67$, $SD = 4.24$; completers - $M = 25.13$, $SD = 4.71$; $t(160) = 2.04$, $p = .043$).

No significant differences between those who dropped out and completers at six month follow-up. However, the low sample size among completers at six-month follow-up may not have allowed a reliable test of dropout versus completer differences.

2.6 Analysis of predictors of treatment improvement

Although the data have not shown strong improvements as a result of the anger management program, it is quite likely that treatment effectiveness is masked by the fact that anger management programs may not work equally effectively for all those that receive the treatment, and may even be detrimental for some participants. One way to investigate this aspect of treatment effectiveness is to identify predictors of treatment improvement (and, by inference, of treatment failure – no improvement or even a worsening of outcome). Predictors of treatment effectiveness can be easily assessed by correlating likely predictors of treatment effectiveness with change over time scores on each of the key outcomes. Change over time scores were calculated for all participants by subtracting “pre-treatment” scores from “post-treatment” scores. As a result, positive scores reflect an increase in scores from “pre-treatment” to “post-treatment”, and negative scores reflect a decrease. The likely predictors of treatment effectiveness that were used were all pre-treatment scores on the key outcome variables as well as the Serin readiness total, risk total, number of previous times in prison, length of sentence, stage of sentence (% through sentence), and number of prison incidents. The correlations between pre-treatment scores and the equivalent change over time scores are presented in Table 21, for both the treatment and control groups.

The negative correlations in Table 21 show that, for all outcomes, those with the highest “pre-treatment” scores had the most negative change scores, and those with the lowest “pre-treatment” scores had the most positive change scores. The interpretation of these results is dependent on the general pattern of change for each outcome. For most variables, the general pattern among participants in the study was a decrease in scores from “pre-treatment” to “post-treatment”, so for these variables, the negative correlations in Table 21 show that the greatest decreases were reported by those who were initially higher at “pre-treatment”, whereas those who reported lower scores at “pre-treatment” exhibited smaller decreases or slight increases by “post-treatment”. On the other hand, for some variables (e.g., anger control, anger knowledge), participants showed a general increase in scores from “pre-treatment” to “post-treatment”. For these outcomes, those

who reported lower scores at “pre-treatment” showed the greatest increases whereas those who reported higher scores at “pre-treatment” exhibited smaller increases or slight decreases in scores. Taken as a whole, these results simply show that those who had the most room for improvement were the ones who showed the greatest improvement.

However, this conclusion cannot be taken to reflect a useful criterion for selection into treatment since the correlations were similar for both the treatment and the control group on most measures. In other words, the evidence is consistent with the idea that greatest improvement will occur for those who exhibit the most negative symptoms initially, irrespective of whether or not they undertake an anger management program. The exceptions occur for the following outcomes, anger arousal, anger knowledge, action (Readiness to Change score), and the self-rated MOAS (weekly), where predictive validity is better in the treatment group than in the control group, suggesting that better gains in these outcomes can be made by taking those who have poorer scores on these variables.

Further analyses were carried out to investigate several other possible predictors of treatment effectiveness. The variables selected for analysis were the Serin readiness total, risk total, number of previous times in prison, length of sentence, stage of sentence (% through sentence), and number of prison incidents. The correlations between these variables and the change scores are shown in Table 22.

The most consistent pattern evident in Table 22 is that correlations between treatment readiness, as measured by the Serin scale, and change on (most of) the anger outcome measures (STAXI and Novaco scales) are significant and negative, for the treatment group. The two exceptions are for anger control and anger regulation which are both positive and not significant. The direction of the correlation for all these outcomes suggests that greater readiness for treatment is associated with greater improvement on the anger outcomes for those who receive treatment. On the other hand, the correlations are either not significant or in the other direction for the control group. In other words, greater readiness for treatment is associated with less improvement on the anger outcomes for those who don't receive treatment.

Table 21. Correlations between pre-treatment scores and corresponding “pre-treatment” “post-treatment” change scores for treatment and control participants

Correlation between pre- measure and pre-post change scores	Treatment	Control	Fisher Z test
STAXI			
State anger	-.52 ^{***}	-.42 ^{**}	Z = 0.87, p > .05
Trait anger	-.58 ^{***}	-.45 ^{***}	Z = 1.17, p > .05
Anger In	-.60 ^{***}	-.46 ^{***}	Z = 1.24, p > .05
Anger out	-.67 ^{***}	-.60 ^{***}	Z = 0.75, p > .05
Anger control	-.41 ^{***}	-.43 ^{**}	Z = -0.15, p > .05
Anger expression	-.52 ^{***}	-.56 ^{***}	Z = -0.34, p > .05
Novaco scales			
Angry cognitions	-.49 ^{***}	-.51 ^{***}	Z = -0.15, p > .05
Anger arousal	-.60 ^{***}	-.31 [*]	Z = 2.09, p < .05
Angry behaviour	-.52 ^{***}	-.42 ^{**}	Z = 0.72, p > .05
Anger regulation	-.45 ^{***}	-.36 [*]	Z = 0.60, p > .05
Total Novaco Anger Scale score	-.54 ^{***}	-.38 ^{**}	Z = 1.22, p > .05
Provocations – disrespectful treatment	-.40 ^{***}	-.41 ^{**}	Z = 0.07, p > .05
Provocations – unfairness	-.46 ^{***}	-.57 ^{***}	Z = -0.92, p > .05
Provocations – frustration	-.48 ^{***}	-.48 ^{**}	Z = 0.00, p > .05
Provocations – annoying traits of others	-.50 ^{***}	-.44 ^{**}	Z = 0.47, p > .05
Provocations – irritability	-.37 ^{***}	-.40 ^{**}	Z = -0.22, p > .05
Total Provocations Inventory score	-.40 ^{***}	-.40 ^{**}	Z = 0.00, p > .05
Other variables			
Watt Anger Knowledge	-.53 ^{***}	-.26 [*]	Z = 2.24, p < .05
SCQ – precontemplation	-.57 ^{***}	-.57 ^{***}	Z = 0.00, p > .05
SCQ – contemplation	-.52 ^{***}	-.54 ^{***}	Z = -0.18, p > .05
SCQ – action	-.67 ^{***}	-.45 ^{***}	Z = 2.16, p < .05
MOAS self-rated weekly	-.72 ^{***}	-.46 ^{***}	Z = 2.67, p < .05
MOAS self-rated monthly	-.50 ^{***}	-.76 ^{***}	Z = 3.17, p < .01

Note: * p < .05 ** p < .01 *** p < .001 Significant differences are shown in **bold**

Table 22. Correlations between possible predictors of treatment effectiveness and pre-post changes in outcome variables for treatment and control groups

Pre-post change in outcome variable	Group	Serin readiness	Risk total	Previous prison	Sentence length	Stage of sentence
Change in state anger	Treatment	-.16*	-.15*	-.11	-.01	+12
	Control	-.01	-.17	+12	-.22	+25
	Sig (diff)	p < .05	p < .05	p > .05	p > .05	p > .05
Change in trait anger	Treatment	-.26***	-.04	-.12	+10	-.01
	Control	+05	-.01	+14	-.06	+19
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in anger in	Treatment	-.17*	+08	+05	+03	-.09
	Control	+36	+26	-.05	+02	+01
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in anger out	Treatment	-.15	+05	-.10	+11	+03
	Control	+31*	+09	-.15	+26	+08
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in anger control	Treatment	+14	+02	+05	-.13	+08
	Control	+19	+20	-.09	+08	+14
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	p > .05
Change in anger expression	Treatment	-.25**	+08	-.03	+19*	-.09
	Control	+26	+19	-.04	+03	-.04
	Sig (diff)	p < .05	p > .05	p > .05	p < .05	p > .05
Change in angry cognitions	Treatment	-.28**	+05	-.03	+13	-.07
	Control	+17	-.12	+11	-.10	+24
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in anger arousal	Treatment	-.20*	+03	-.04	+15	+05
	Control	+17	-.05	+28	+19	-.13
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in angry behaviour	Treatment	-.29**	+02	-.14	+17	-.07
	Control	+21	-.08	+27	+06	+21
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in anger regulation	Treatment	+03	-.18*	-.02	-.04	-.15
	Control	+17	-.02	+13	-.19	-.35
	Sig (diff)	p > .05	p < .05	p > .05	p > .05	p > .05
Change in total Novaco anger scale score	Treatment	-.26**	+04	-.04	+14	-.04
	Control	+21	-.12	+25	+05	+11
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in provocations (disrespectful treatment)	Treatment	-.22**	+09	-.08	+14	-.10
	Control	+06	-.04	+10	+17	+13
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	p > .05

* $p < .05$ ** $p < .01$ *** $p < .001$ Significant differences are shown in bold

Sig (diff) refers to the p value for Fisher's Z test to test whether the correlations are significantly different

Table 22. (cont.) Correlations between possible predictors of treatment effectiveness and pre-post changes in outcome variables for treatment and control groups

Pre-post change in outcome variable	Predictors of treatment effectiveness					
	Group	Serin readiness	Risk total	Previous prison	Sentence length	Stage of sentence
Change in provocations (unfairness)	Treatment	-.20**	+.05	+.02	+.13	+.01
	Control	-.01	-.02	+.18	+.18	+.23
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	p > .05
Change in provocations (frustration)	Treatment	-.20**	+.02	-.17*	+.11	-.05
	Control	+.30	-.07	+.09	+.23	+.09
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	P > .05
Change in provocations (annoying traits others)	Treatment	-.28**	-.02	-.04	+.15*	-.10
	Control	+.07	-.15	+.09	+.16	+.07
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in provocations (irritability)	Treatment	-.23**	+.02	+.00	+.12	-.06
	Control	+.10	-.13	+.03	+.07	+.03
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	P > .05
Change in total provocations score	Treatment	-.27**	+.04	-.07	+.15	-.07
	Control	+.12	-.06	+.11	+.19	+.13
	Sig (diff)	p < .05	p > .05	p > .05	p > .05	p > .05
Change in anger knowledge (WAKS)	Treatment	-.01	-.10	-.03	-.03	-.02
	Control	-.04	-.03	-.22	-.02	-.18
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	P > .05
Change in precontemplation score	Treatment	+.13	-.04	+.05	-.11	+.08
	Control	-.03	-.03	.00	+.03	-.16
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	P > .05
Change in contemplation score	Treatment	-.13	+.01	+.01	-.03	-.09
	Control	-.06	-.14	-.10	-.06	-.06
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	P > .05
Change in action score	Treatment	-.18	+.01	-.10	-.08	-.08
	Control	+.08	+.08	-.17	+.01	-.13
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	p > .05
Change in MOAS score (weekly)	Treatment	-.02	-.02	-.07	+.04	+.09
	Control	+.08	-.05	-.03	+.00	+.17
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	p > .05
Change in MOAS score (monthly)	Treatment	-.05	+.03	+.00	+.05	+.11
	Control	-.10	+.14	-.03	+.04	+.05
	Sig (diff)	p > .05	p > .05	p > .05	p > .05	p > .05

* p < .05 ** p < .01 *** p < .001 Significant differences are shown in **bold**
 Sig (diff) refers to the p value for Fisher's Z test to test whether the correlations are significantly different

In particular, the results show that the correlations of the Serin readiness scores and the change in outcome scores are significantly different between the treatment and control groups. Hence, readiness for treatment, as measured by the Serin scale, is a good predictor of treatment effectiveness because it has predictive validity only for those who undertake an anger management program.

No other predictors (risk total, previous times in prison, sentence length, or stage of sentence) show any consistent predictive validity of treatment effectiveness.

2.7 Relationship between risk, needs and motivation and treatment effectiveness

A further question of interest is whether treatment is most effective when clients are of high need, high risk and have high motivation to change their behaviour. Subjects were classified as high and low on risk, need and motivation according to the criteria detailed on pages 23 to 25. Two hundred and nineteen research participants reported valid scores on trait anger, Novaco Anger Scale Part A, VOTP risk assessment, and the Readiness to change Questionnaire as well as provided "post-treatment" data on the key outcomes. Of these, 100 (34.4%) were classified as high in need, 185 (63.6%) as high in risk, and 243 (83.5%) as highly motivated. Of these, 71 (24.4%) were classified as high on all three facets. This resulted in 16 (21.3%) high need/risk/motivation prisoners in the control group and 55 (25.5%) high need/risk/motivation prisoners in the treatment group.

The data were analysed using a repeated measures ANOVA with one within-subjects factor, time of measurement (pre-test versus post-test), and two between-subjects factors, group (treatment versus control) and need/risk/motivation category (high versus low). For each of the Novaco scales, the analyses also used social desirability as a covariate. The primary focus in these analyses were on the three way interaction between time of measurement, group, and need/risk/motivation category. The three way interaction is the key test of a differential effect of need/risk/motivation category on treatment effectiveness. In other words, it would be hypothesised that greater improvement will be observed in the treatment group compared to the control group, but only for those high in need, risk and motivation.

Evidence for a three way interaction was found for three outcome measures (although for two of these outcomes, the statistical significance level obtained was over the conventional alpha level of .05). For each of these three outcomes, two graphs were plotted; (1) mean change in the

outcome for the treatment and control groups according to need/risk/motivation category, and (2) mean pre-test and post-test scores for the treatment and control groups according to need/risk/motivation category. These graphs are shown in Figures 1 (anger in), 2 (total NAS), and 3 (Watt anger knowledge).

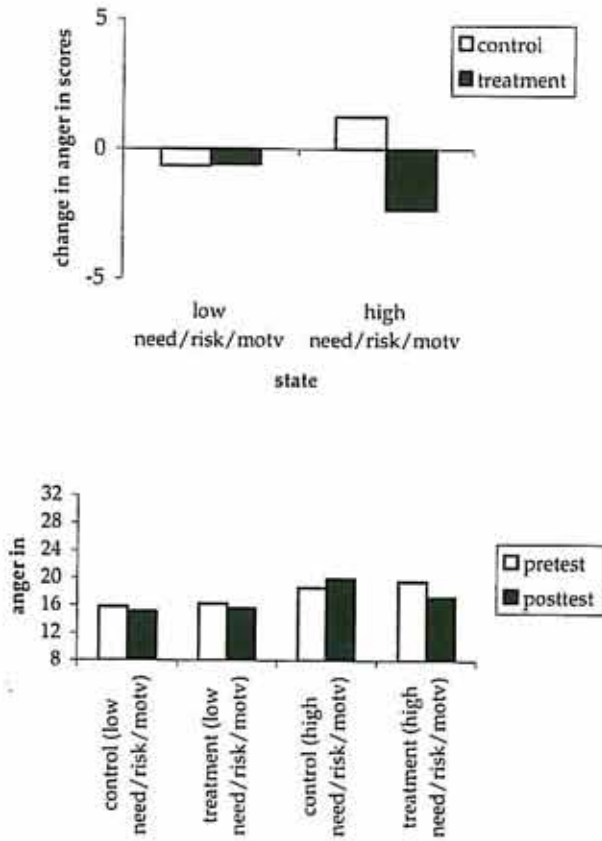


Fig. 1. Mean change and pre-test post-test means for anger in by group and need/risk/motivation category

The ANOVA on anger in revealed evidence of a three-way interaction between time of measurement, group, and need/risk/motivation category, $F(1,192) = 2.8, p = .093$. Figure 1 shows that this interaction is consistent with the hypothesised effects. Among those classified as not being high in risk, need, and motivation, small (essentially negligible) improvements are found on anger in for both control and treatment participants. However, for those classified as high on need, risk and motivation, a modest improvement (effect size equals .47 of a standard deviation)

is observed in the treatment group, whereas a slight decrement (effect size equals .25 of a standard deviation) is observed in the control group.

A similar pattern was observed for total Novaco anger scale scores. Evidence was found for a three way interaction, $F(1,170) = 3.0, p = .084$. Figure 1 shows that this interaction is also consistent with the hypothesised effects. Although, all conditions show a decrease in anger, the only condition where a moderate effect is found is among those high in need, risk and motivation who undertook the anger management program (effect size equals .61 of a standard deviation; less than .20 for all other conditions).

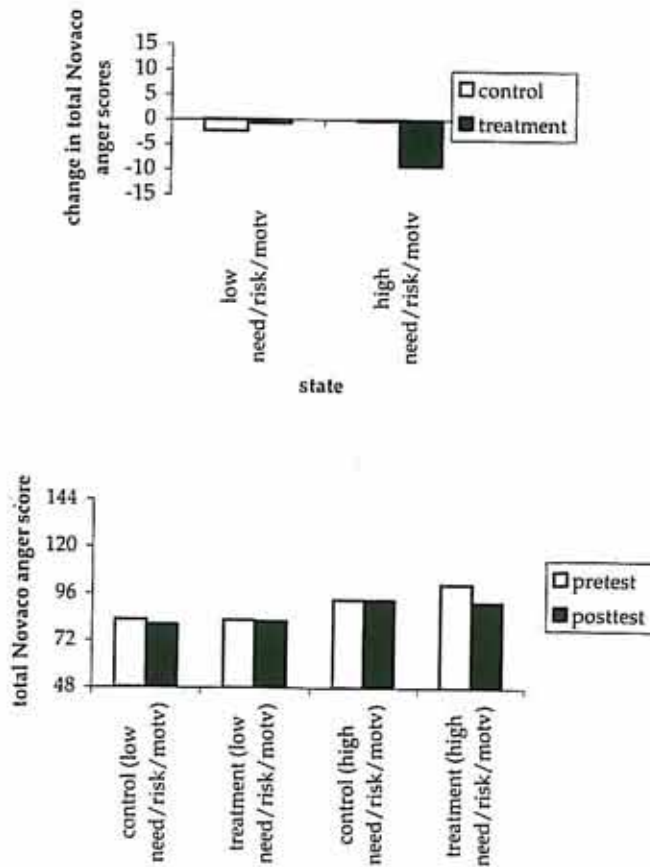


Fig. 2. Mean change and pre-test post-test means for total NAS scores by group and need/risk/motivation category

Both of these sets of results provide some modest evidence that treatment is most effective when targeted at those high in risk, high in need and high in motivation to change their behaviour.

The final outcome variable that showed evidence of a three way interaction between time of measurement, group, and need/risk/motivation category was anger knowledge (WAKS), $F(1,218) = 4.1, p = .045$. The results are shown graphically in Figure 3. Contrary to the hypothesis, these data show treatment is more effective for improving anger knowledge when the participants are not in the high need, risk and motivation category (effect size of .63 of a standard deviation compared to .14 for control group). By contrast, treatment is less effective (than a control condition) for improving anger knowledge when the participants are in the high need, high risk and high motivation category (effect size of .19 of a standard deviation compared to .52 for control group).

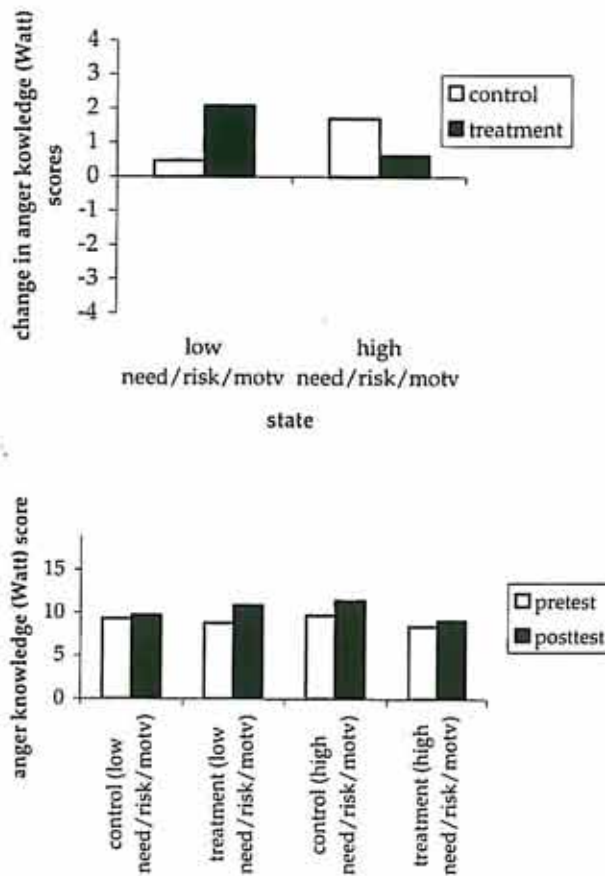


Fig. 3. Mean change and pre-test post-test means for anger knowledge by group and need/risk/motivation category

Similar analyses were planned using the Serin readiness scale instead of the Readiness to change Questionnaire as a measure of motivation for treatment, particularly given the prior results

showing the predictive validity of the Serin readiness scale as a predictor of treatment effectiveness. However, using the median as a cut-off for high and low motivation resulted in only 5 (7.1%) high risk, high need, high motivation participants in the control group and 29 (13.8%) in the treatment group. The small sample size, particularly in the control group, would make it difficult to draw any meaningful conclusions from any analyses.

Finally, treatment performance was compared for those high versus low in risk need and motivation. No significant differences between the two groups were found for both the self report treatment performance measure (low, $M = 61.42$, $SD = 8.31$; high, $M = 60.61$, $SD = 6.70$; $t(99.0) = 0.67$, $p = .51$) and the facilitator rated treatment performance measure (low, $M = 55.63$, $SD = 7.44$; high, $M = 55.89$, $SD = 8.64$; $t(75) = -0.13$, $p = .90$).

2.8 Relationship between program integrity and treatment effectiveness

One important aspect that has rarely been addressed in the previous literature assessing treatment effectiveness is the integrity of the treatment program. It is reasonable to expect that treatment will be more effective when the delivery of the program closely follows the procedures of the original program design (i.e., has high integrity). A total integrity score was calculated to reflect overall program integrity. An integrity classification using a median split was derived from program facilitator ratings and client ratings to produce two integrity factors.

Of the 32 anger management groups, 9 (28%) were rated as low in integrity by both clients and facilitators, 9 (28%) were rated as high in integrity by both clients and facilitators, 9 (28%) were rated as low in integrity by clients and high in integrity by facilitators, and 5 (16%) as high in integrity by clients and low in integrity by facilitators. As a result of the relatively poor correspondence between client and facilitator ratings (only 18 of 32 (56%) agreed), program integrity was evaluated separately on the basis of client and facilitator ratings.

The relationship between program integrity and treatment effectiveness was analysed using repeated measures ANOVA with one within-subjects factor, time of measurement (pre-test versus post-test), and two between-subjects factors, level of integrity (high versus low). Social desirability was controlled for in each of the analyses involving the Novaco anger measures. Each analysis was carried out separately for each outcome and for each integrity measure.

Mean change scores (post-treatment - pre-treatment) on all key outcome measures for participants in high integrity and low integrity programs are reported in Table 23 (based on client

ratings of program integrity) and Table 24 (based on facilitator ratings of program integrity). The results are quite consistent between client rated integrity and facilitator rated integrity. The results show that high integrity programs are associated with significantly greater improvements in anger control, and three of the Novaco provocation measures (disrespectful others, annoying traits of others, and irritability), when compared with low integrity programs.

The effect of program integrity on treatment performance showed no evidence for reliably superior treatment performance in high integrity programs. For self rated treatment performance, there were no significant difference among programs of low and high client rated integrity measure (low, $M = 60.78$, $SD = 7.56$; high, $M = 61.73$, $SD = 7.82$; $t(149) = -0.76$, $p = .45$) or among programs of low and high facilitator rated integrity measure (low, $M = 60.47$, $SD = 7.56$; high, $M = 61.47$, $SD = 7.90$; $t(140) = -0.75$, $p = .46$). Similarly, for facilitator rated treatment performance, there were no significant difference among programs of low and high client rated integrity measure (low, $M = 56.43$, $SD = 7.90$; high, $M = 54.81$, $SD = 6.80$; $t(85) = 1.03$, $p = .31$) or among programs of low and high facilitator rated integrity measure (low, $M = 54.19$, $SD = 6.82$; high, $M = 56.35$, $SD = 7.56$; $t(85) = -1.33$, $p = .19$).

In summary, these data show preliminary evidence that evaluation of treatments should also consider program integrity as a moderator of treatment effectiveness. On the other hand, some caution is required in drawing conclusions from the results because of the relatively poor agreement about program integrity by the client and facilitators.

Table 23. Mean change scores on outcomes in low and high integrity programs (client ratings)

Outcome	Program Integrity		Significance of mean difference
	Low	High	
State anger	- 1.25	+ 0.56	Not significant
Trait anger	- 1.52	- 1.24	Not significant
Anger in	- 0.56	- 2.01	Not significant
Anger out	- 1.05	- 1.16	Not significant
Anger control	+ 0.72	+ 1.92	$F(1,151) = 1.71, p = .19$
Anger expression	- 3.22	- 5.19	Not significant
Angry cognitions	- 2.04	- 2.58	Not significant
Anger arousal	- 1.35	- 2.43	Not significant
Angry behaviour	- 1.90	- 2.99	Not significant
Anger regulation	+ 0.27	+ 0.72	Not significant
Total NAS score	- 5.29	- 8.00	Not significant
Provocations (disrespect)	+ 0.27	- 1.12	$F(1,107) = 5.05, p = .027$
Provocations (unfairness)	- 0.29	- 1.47	Not significant
Provocations (frustration)	- 0.22	- 0.92	Not significant
Provocations (annoying)	+ 0.05	- 1.45	$F(1,107) = 4.12, p = .045$
Provocations (irritability)	+ 0.33	- 1.18	$F(1,107) = 4.43, p = .038$
Total provocations score	+ 0.09	- 5.32	$F(1,135) = 3.37, p = .054$
WAKS	+ 1.55	+ 2.13	Not significant
Precontemplation	+ 0.13	+ 0.24	Not significant
Contemplation	+ 0.45	+ 0.20	Not significant
Action	+ 1.29	+ 1.56	Not significant
MOAS (weekly)	- 0.25	- 0.16	Not significant
MOAS (monthly)	- 0.42	- 0.52	Not significant

Means are adjusted (for social desirability) for the Novaco anger measures

Table 24. Mean change scores on outcomes in low and high integrity programs (facilitator ratings)

Outcome	Program Integrity		Significance of mean difference
	Low	High	
State anger	- 1.25	- 0.11	Not significant
Trait anger	- 1.13	- 1.60	Not significant
Anger in	- 1.02	- 1.22	Not significant
Anger out	- 1.59	- 0.92	Not significant
Anger control	- 0.18	+ 1.83	$F(1,142) = 4.22, p = .042$
Anger expression	- 2.79	- 4.49	Not significant
Angry cognitions	- 1.78	- 2.33	Not significant
Anger arousal	- 1.09	- 2.05	Not significant
Angry behaviour	- 1.77	- 2.56	Not significant
Anger regulation	- 0.36	+ 0.93	$F(1,124) = 2.79, p = .097$
Total NAS score	- 4.64	- 6.94	Not significant
Provocations (disrespect)	+ 0.34	- 0.86	$F(1,105) = 3.72, p = .057$
Provocations (unfairness)	- 0.23	- 1.20	Not significant
Provocations (frustration)	+ 0.15	- 1.06	Not significant
Provocations (annoying)	+ 0.26	- 1.37	$F(1,105) = 5.70, p = .019$
Provocations (irritability)	+ 0.49	- 0.96	$F(1,105) = 4.10, p = .045$
Total provocations score	+ 0.85	- 4.23	$F(1,123) = 3.39, p = .068$
WAKS	+ 1.55	+ 1.92	Not significant
Precontemplation	+ 0.38	+ 0.03	Not significant
Contemplation	- 0.16	+ 0.59	Not significant
Action	+ 0.81	+ 1.61	Not significant
MOAS (weekly)	- 0.60	- 0.04	Not significant
MOAS (monthly)	- 0.88	- 0.32	Not significant

Means are adjusted (for social desirability) for the Novaco anger measures

DISCUSSION AND OVERVIEW

The research described in this report comprises three studies: the Normative Study, the Pilot Study and the Main Study. In this section we will summarise some of the main findings that have emerged and outline their implications for conducting anger management programs within correctional systems.

The Normative Study looked at the experience of anger in a sample of 121 male convicted adult prisoners in South Australia and Western Australia. A range of anger-related psychometric measures were used. The results confirmed that this sample of adult male prisoners had higher levels of anger than other non-criminal populations, though they obtained lower scores than American offender populations. The high anger scores do suggest that anger management is a relevant intervention for the prison population as a whole. However, a wide range of anger scores exists in the sample studied, with some individuals obtaining "normal" or low scores while others have very high scores. This clearly suggests that not all prisoners need anger management treatment.

In the Normative study we attempted to estimate what proportion of the prisoner population might require intervention. The rehabilitation literature suggests that offenders high in Risk, Needs and Responsivity are the most appropriate targets for intervention (Andrews and Bonta, 1994; Howells and Day, 1999). Using the measures available in the Normative Study, we have made a tentative estimate that 22% of the population might reasonably be regarded as good candidates for anger management intervention. This figure needs to be regarded with some caution as it is dependent on a relatively modest sample size. It is also based on resident prisoners rather than on new admissions. It would be useful to build measures of this sort into admission assessments in prisons so that the level of need for treatment could be established. This would be a relatively straightforward task.

Several other serendipitous findings emerged in the Normative Study, though they need to be interpreted with caution and replicated on a larger sample. Indigenous prisoners obtained higher anger scores than non-Indigenous prisoners on some measures. This finding is open to many interpretations and it cannot be ruled out that it is artefactual. Nevertheless, given the links between high anger and violent offending (Howells, 1997; Novaco, 1997, further research is warranted into the nature of the experience and expression of anger in Indigenous offenders. Such research may be relevant to the issue of how standard rehabilitation programs such as anger management need to be adapted to meet the needs of Indigenous offenders (Mals, Howells, Hall and Day, 2000). Given the higher levels of anger in Indigenous prisoners, it would be expected that, other things being equal, Indigenous prisoners would be well-represented amongst those offered anger management treatment.

The data also suggested that those in the later stages of their sentence experienced greater problems with anger than those in the earlier stages. It may be appropriate to target such individuals for treatment.

The Readiness for Change measure, adapted and devised for the present study, proved to be useful and to have a sound factorial structure. This measure will be useful as a possible screening assessment to select prisoners who are ready to address their angry behaviour. It is a brief measure which is straightforward and easily understood. It can be used to define offenders in the Precontemplation, Contemplation and Action stages respectively.

The Pilot Study, though substantial in itself, allowed for a “test run” of all procedures and new measures devised for the main study. Forty eight anger management referrals, recruited from both prisons and community corrections, in South Australia and Western Australia, were studied. Participants were tested pre- and post- the intervention period. Treated participants were compared with waiting-list controls. No significant changes were found when pre- and post-measures were compared for the treatment group and there were no significant differences between treated and untreated participants.

The Pilot Study did demonstrate the utility of the data collection protocols. Few problems were encountered with the measures used. One of the major functions of the study was to pilot the method for assessing program integrity of the anger management groups – that is, the extent to which what occurred in the group sessions complied with what the program Manual suggested should occur. As rated by group participants and the facilitator, program integrity ratings were

generally high. This demonstrates that the Manuals are being effectively implemented in both states. This is an important finding, given the evidence elsewhere that low integrity programs are less effective than those with high integrity (Hollin, 1999). The Program Integrity Checklist devised for the present study provides a useful start in attempting to measure program integrity in a correctional setting, a task internationally recognised as being a difficult one. The exact content of the Checklist would need to be modified for use in other Australian states to reflect the differences in program content that may exist in different states.

The Pilot Study demonstrated the difficulty in obtaining correctional officer ratings of real-life angry and aggressive behaviours by the offenders. Aggressive behaviours were rarely endorsed, even prior to the intervention, thus making it difficult to establish a baseline with which treatment gains could be compared. Lack of familiarity with the offenders' real-life behaviour was a common reason for not reporting angry incidents. The implication for future routine use of these behavioural measures in correctional settings is that they will be more useful when completed by an officer who knows the offender well and who has ways of finding out how the offender has actually behaved over the time period required.

As a result of the Pilot Study some modifications were made to the measures planned for the Main Outcome Study.

MAIN OUTCOME STUDY

The main outcome study describes differences pre- and post-treatment in those offenders undertaking anger management programs and waiting-list controls. A large number (418) completed the initial assessment. Two hundred and eighty five people completed the post assessment and a further 78 were re-assessed at a two-month follow-up and a further 21 were re-assessed after 6 months. The sample were exclusively male and were predominantly Australian /New Zealanders by self identification, with approximately one fifth describing themselves as from an ATSI ethnic background. The vast majority had formal convictions for a violent offence.

The total sample of participants is substantial and the study is one of the largest ever conducted of anger management outcomes. Although the numbers completing each test were often reduced because of non-completion of items or non-scorable responses, the overall data set is nevertheless large enough to allow reasonably reliable conclusions to be drawn from the results.

There was a general tendency for the South Australian participants to obtain higher scores than their Western Australian counterparts on a range of anger measures. The reasons for this difference are not clear as yet. It may reflect sentencing differences between the two states, differential criteria for determining that an individual has an anger problem or other factors. It is unlikely that the difference can be explained in terms of differences in the conditions of prison life in the two states (the notion, for example, that prison life in South Australia might be in some way more frustrating). This explanation can be rejected because the inter-state differences are evident for various aspects of trait anger (the individual's general disposition to become frequently or intensely angry) in addition to state anger (how angry the individual feels at the time precise they are assessed). It is, arguably, more likely that high anger is an enduring feature, possibly present prior to imprisonment, which the South Australian participants are more likely to bring with them on admission to prison. The higher violence risk score in the South Australian sample would also suggest that the SA referred population is a more aggressive one in terms of previous violent offending. Whatever the explanation of the differences between the two states, the need for anger management interventions is clearly slightly greater in South Australian prisons.

That the experience of imprisonment may have some effect on levels of anger is, however, suggested by the finding that prisoners who are towards the end of their sentence tend to have higher anger scores on some measures than those who are early in their sentence. This finding replicates a similar finding for a non-selected prison population sample in the Normative Study. What the present study cannot reveal is the nature of the frustrations and problems that elevate anger later in a sentence. Further qualitative research may throw some light on this issue and on how this sentence-related anger may be remediated.

The Main Outcome Study also confirmed the pattern detected in the Normative Study – that there is a trend towards higher anger scores in Indigenous (ATSI) offenders. Our results suggest that future research needs to be conducted to illuminate the nature of angry experience in Indigenous offenders. Whether and how anger is qualitatively different in Indigenous offenders needs to be investigated. It will also be important to determine whether anger programs for Indigenous people might need a different focus and context.

Does anger management work?

The core question for the outcome study is clearly - What is the impact of anger program participation on offenders? With some consistency the results demonstrate that the overall impact of the anger management interventions is small. Although the treated group consistently makes changes in the expected direction, the changes are not large enough to be of real clinical significance. It is also the case that similar small changes in the direction of improvement can be observed for the control group on many measures. This would suggest that the act of completing anger assessments may have a small beneficial effect in itself, even when treatment is not provided. The tendency for problem behaviours to be reactive to the assessment process itself (that is for problems to decrease following testing) is a well known phenomenon for psychological and psychiatric treatment interventions. This finding does highlight the need for a control group in any future evaluations of anger management (or any other) programs in correctional environments. Without a control group, it is possible to make an incorrect inference that a pre-test/post-test improvement in a program evaluation is attributable to the particular program implemented.

The critical issue is whether the improvements that occur in the treated group are significantly greater than those that occur in the controls. The present results clearly show that there are very few statistically significant differences between the treatment and control groups. There are only two exceptions to this pessimistic conclusion; the findings for Anger Knowledge (WAKS) and, to a less clear extent, for "Readiness to change". The treated group improve their anger knowledge more than do the controls, though, again, the difference is very small in absolute terms. It might be argued that the changes in anger knowledge are unsurprising, given that the content of the WAKS measure (see Appendix B). The items of the WAKS relate closely to the content of what is taught in the anger management program and were originally devised by Watt to reflect what is covered in anger programs. It would be surprising indeed if treatment participants did not show some increased elementary knowledge of this sort.

The differential changes in the "Readiness to change" classification for the two groups are not great but they do suggest a slight tendency for treated subjects to move in the direction of greater readiness to undertake change. The question arises as to how the (small) positive results of treatment on these two measures can be squared with the absence of a treatment effect for trait and state anger and the myriad other features of anger assessed by the measures. One

interpretation is that these positive effects are no more than chance occurrences reaching statistical significance only because of the large number of measures being assessed. An argument against this hypothesis is that a similar effect was found for Anger Knowledge in an earlier, smaller and less comprehensive study in Western Australia by Watt and Howells (1999). Again, in this latter study, there were no significant differences on anger between the treatment and control groups, with the exception of Anger Knowledge, where a positive effect for treatment was found.

A second interpretation, and one we would endorse, is that the anger management programs are having an effect only at an “educational” level (knowledge about what anger is and its effects and, possibly, commitment to change). They have only a small effect however on the direct experience of anger itself (frequency and intensity), or on the important physiological, cognitive or behavioural components of anger. A “therapeutic” effect for anger management programs would require change at these latter’ levels. We will argue later in this report that significant therapeutic change may not be possible for the current anger management programs in South Australia and Western Australia because they are too brief and insufficiently intensive to produce therapeutic change.

State and community/prison differences

Some differences between the two states were observed. Nevertheless, the broad pattern of results discussed above held equally for both states. Some inter-state differences were observed in the treatment versus control comparisons but these are difficult to interpret (see Results section above). Overall, it is reasonable to conclude that the broad pattern of treatment having a very small or minimal impact holds for both states. The fact that findings are similar across two jurisdictions, notwithstanding some detailed differences, may suggest that the results are generalisable to other states that offer similar programs. The results clearly suggest a degree of caution is required in an Australian correctional context, before accepting that anger management programs of this sort, delivered to offenders without substantive pre-treatment assessment, will have a significant impact on those who participate. The question of how program delivery might be adapted to make it more effective will be discussed further below.

No significant differences were observed for community versus prison participants in either the Pilot Study or the Main Outcome Study. However, the number of community participants was very small. Future research should re-examine community versus prison differences in a more

substantial way, given the general evidence that rehabilitation programs in the community are more effective than in institutions.

Do program effects last?

The question of whether treatment gains brought about by intervention (for example, anger reduction) endure over the subsequent months after the program is ended is a very important one for any psychological intervention. It is equally important for anger management in correctional settings. In the present study selected participants in the treatment group were followed up for 2 months and 6 months respectively. Given that the changes brought about by the program were modest (see above) it is not easy to determine whether improvements are maintained in the follow-up period and some caution is required in interpreting trend analyses. The analyses conducted suggested that different follow-up trends occurred for different measures. It is encouraging that the improvements in Anger Knowledge (WAKS) are maintained and even increased over the 6-month follow-up. Linear trends were also found for general measures of anger, angry cognitions and other aspects of angry behaviour which suggested that the positive changes brought about by the program were further improved upon at 2-month follow-up. Some other anger measures also showed a tendency to improve further at the 6 months follow-up.

Taken as a whole, the follow-up data provide some encouragement that the rate of improvement brought about by the anger intervention programs is maintained on at least some measures. The mechanisms that may underlie maintenance or improvement are not known. It may be that participants rehearse and remember their experiences in the groups or that group participation encourages them to work further on their problems or to discuss them constructively with correctional staff. In any future development or adaptation of the anger management programs in the two states, it will be important to consider the issue of how treatment gains can be maintained and built upon. There is increasing evidence that follow-up and “booster” sessions add to the impact of correctional programs (Day and Howells, 2001). The present study suggests that maintenance of gains over time is possible.

For whom does anger management work best?

The second major question addressed in the Main Outcome Study, and in the project as a whole, is - What participant characteristics are associated with making treatment gains in the anger management programs? This is an important question for a number of reasons. It is widely acknowledged that violent offenders are a heterogeneous group with a multiplicity of criminogenic needs. Given that having a history of violence is likely to be an important criterion in referring an offender for anger management, it is likely that anger management referrals are equally heterogeneous. Howells, Hall, Baldwin and Watt (1997) have previously argued that some offenders referred to anger management programs in correctional systems world-wide may not actually have high anger problems —the “instrumental”, the “psychopathic” and the “overcontrolled” violent offender are all possible examples of this phenomenon. If this is true, then it would be expected that the measured impact of anger management programs (for example, in pre-post comparisons) would be significantly diminished. Thus understanding the effect of individual differences amongst offenders is crucial.

Knowledge of which individuals benefit most is also crucial for correctional managers who need to know for which offenders referral to anger management is cost effective.

The results of the present study support the notion of individual differences being important. The extent of change (improvement) of an individual undertaking the programs was shown to be predictable from a number of pre-treatment measures. Across a range of anger measures, those high in anger and low in anger control at the pre-treatment assessment showed the greatest change at the post program re-test. In everyday terms, the worse you are, the better you do! The results suggest that measures of anger arousal (that is, the extent of physiological arousal when angry), Readiness to Change and the extent of self rated angry experiences and other measures are good discriminating psychometric variables. The readiness scale, based on Serin's work, proved to be a very consistent predictor of improvement in treatment. Offenders who are motivated and ready to work on their anger problems show much greater improvements on a wide range of anger measures, Conversely, those who are poorly motivated do show less or no change. Risk level, sentence length, previous imprisonment and stage of sentence showed no clear pattern of association with improvement in treatment and are unlikely to be useful as screening instruments.

International practice in offender rehabilitation over the past decade has increasingly emphasised that certain rehabilitation variables are likely to be associated with good rehabilitation

outcomes, as measured by reductions in recidivism. In particular, the Risk, Needs and Responsivity principles have been recommended (Andrews and Bonta, 1994; Day and Howells, 2001; Hollin, 1999, Howells and Day, 1999). These principles suggest that rehabilitation be focussed on those of high risk, high criminogenic needs and high responsivity. In our Normative Study we made a crude estimate of the prevalence of such offenders in a prison sample (see above). In the Main Outcome Study it was demonstrated that participants classified as high in risk, needs and responsivity (motivation) did indeed show greater improvement on some anger measures, compared to a group low on these factors, though this pattern was reversed for the anger knowledge scale. Why anger knowledge should be different is not clear and requires further investigation.

A fourth influential principle in correctional rehabilitation has been Program Integrity, which suggests that programs high in program integrity typically have greater impact than those low in integrity (Day and Howells, 2001; Hollin, 1999). Integrity refers to the extent to which the program is delivered in practice in the way in which it was designed and planned in principle. Our Pilot Study suggested that integrity was generally high for various aspects of integrity in both South Australian and Western Australian programs. Nevertheless, it is inevitable that some variation will occur in practice, with some facilitators introducing variations from the Manual more than others. In our study some differences in outcome were shown to relate to program integrity. Although differences were not apparent on all measures, low integrity programs were associated with less positive outcomes, particularly in the areas of anger relating to perceived provocations and anger control. More detailed analysis will be possible at a later stage into the various components of the integrity measure we developed, to see if any particular aspect of integrity is particularly important. The present results confirm that the monitoring of program integrity is an important feature of the delivery of effective programs in correctional settings. A major impediment to monitoring integrity in correctional settings is that the most obvious methodology (the video-recording and subsequent content analysis of session content) is likely to be unfeasible because of problems of labour intensiveness of the task, confidentiality, compliance by participants and equipment availability. The simple integrity assessment devised for the present study (see Appendix E) using both participant and facilitator ratings of each session has some merit, in that it has been shown to have some relationship to program outcome and has the potential to be further developed for use in the routine correctional environment.

How do the outcome study results compare with previous studies?

How do the results of the present study compare with previous evaluations of the effectiveness of anger management programs? Programs for violent offenders, such as anger management, have been provided in a variety of forensic and clinical settings and a number of workers have described the rationale for, and content of, anger/violence reduction programs but the evaluation of the effects of interventions has been limited.

One of the problems in assessing the anger management literature is that many studies are conducted with non-offender populations. It cannot therefore be assumed that the findings from these studies can be generalized to offender populations.

Of the studies that have been conducted with offenders, many have methodological problems, including lack of control groups, absence of comprehensive measures or poorly specified comparison groups.

Stermac (1986) evaluated the effects of a program including cognitive skills, relaxation and assertiveness training, with 40 forensic psychiatric patients. Participants with a history of anger control problems or aggressive behaviour were randomly assigned to treatment or to a control group. In comparison to the control group, at post-testing the treated group reported less angry feelings, more cognitive change and less self-denigration in response to provocation.

McDougall, Barnett, Ashurst and Wills (1987) have reported the effects of a 6-week anger management program at a Youth Custody Centre in England. Pre/post intervention changes were shown in disciplinary reports, suggesting improved anger control. Treatment was for 6 sessions of 1.5 hours.

McDougall and Boddis (1991) evaluated a brief anger management program for offenders with anger-control problems, as identified by prison staff. Participants were randomly assigned to either a treatment or control group. Greater improvements were found for the treated group on self-reported aggression, anger and governor's reports.

Dixon and Polascheck (1992) included recidivism data in their evaluation of a cognitive-behavioural intervention for violent offenders in New Zealand. A reduction in recidivism followed completion of the program, but no control group was included in the study.

One of the best evaluated aggression control interventions is that reported by Goldstein and Glick (1996) with juvenile delinquents. Their Aggression Replacement Training (ART) has been developed and evaluated over a 10 year period, with encouraging results. This is a multimodal

intervention including "skillstreaming", moral education and anger control training components. A series of controlled evaluations, using a range of treatment outcome measures, have provided evidence that ART is more effective than no treatment and other control conditions. Goldstein and Glick's program is clearly wider ranging and more intensive than typical anger interventions in correctional systems for adult offenders.

Dowden, Blanchette and Serin (1999) have reported a substantial study of the effectiveness of an anger-management program with offenders in Canada. The program itself was a reasonably substantial one – 25 two-hour sessions, approximately double the length of the programs in the present study. The program was shown to have an impact in reducing recidivism over a three-year period, though this improvement was found only for high-risk offenders

Beck and Fernandez (1998) conducted a meta-analysis of the effectiveness of cognitive-behavioural programs for the treatment of anger. An average weighted effect size of .85 was found. This suggests that inmates receiving cognitive-behavioural treatment did better than approximately 80% of those not receiving treatment. Clearly, the results reported in the present study suggest an impact far smaller than that found by Beck and Fernandez. It needs to be borne in mind that only 6 of the 50 studies evaluated by Beck and Fernandez involved offenders. Thus it may be the case that that smaller effects of treatment are to be expected for offender populations, particularly for the group of violent offenders who form most of the participant group in the present study.

None of the studies reviewed above was conducted in Australia, though Dixon and Polascheck (1992) did their work in New Zealand. A recent small-scale, controlled study by Watt and Howells (1999) was conducted in Australia (WA) and suggests a need for caution before applying anger management indiscriminately with violent prisoners. The results of the WAKS and Howells study are similar to those in the present study and tend to reinforce our conclusion that the impact of such programs is modest.

Explaining low impact of offender programs

Watt and Howells (1999) put forward several possible explanations for the low effectiveness of the programs they studied, including a) poor motivation of participants b) the content of programs being too complex for the limited program time available c) low program integrity and d) limited opportunity to practice the skills learned in the program . The results from the present study allow us to rule out explanation c, in that program integrity was relatively high. Explanations a, b and d still stand as potential explanations. It could be argued that explanation a is given increased credibility by the present study in that motivational-type factors were shown to predict whether improvement occurred. Explanations b and d amount, arguably, to the suggestion that the programs are too short for the amount of work that needs to be done (low intensiveness).

Both the motivational and low intensiveness explanations are credible and are not mutually exclusive. To these two explanations we would like to add a possible third - multiple problems in offender populations. Again, this explanation does not exclude the other two. Indeed all three factors may interact to diminish program effectiveness.

Motivational problems on the part of program participants are readily identified by most correctional staff as a major factor determining progress in program sessions. Motivational issues have been curiously neglected in the anger management literature. Howells (1998) has argued that anger management needs to be preceded by an analysis of the "goal structure" of the offender and has suggested that considerable variation in goal structures occurs within offender populations. Ralph Serin's work in Canada has also begun to unravel some of the important dimensions of motivation and Renwick, Black, Ramm and Novaco (1997) have described the problems at a clinical level.

Renwick et al (1997) point to the therapeutic pessimism felt by both clients and therapists in correctional and high security settings and to enduring problems of low motivation, treatment resistance and avoidance. These authors note the resentful, distrustful and even combative style of some offender participants in therapeutic groups. Additionally, the clients had realistic concerns about the effects of disclosure of their emotions and past behaviour on release or parole plans. Novaco (1997) similarly, highlights the long histories of failure, institutionalization and social rejection that characterize such clients and which entrench their anger and aggression.

The issue of multiple problems in offenders is also a self-evident one for many correctional staff but, again, it has been neglected, until recently in the anger management field. It is a truism that offenders, particularly high risk offenders, have multiple psychological and social problems. It is a very different task conducting anger management with someone who has no other serious problems apart from anger control itself than it is conducting the same program with an offender who has, for example, an antisocial personality disorder, severe substance abuse problems, limited verbal skills and absence of family support. Establishing a working therapeutic alliance with such a person may itself be a time consuming but necessary task before the specifics of anger control can be addressed (Howells, 1998).

Given these two factors (low motivation and multiple problems), it would not be surprising if anger management with offenders had far less impact than it does with non-offenders. The remedy would be to make offender programs (or at least high risk/needs offender programs) intensive enough to allow for offender problems to be addressed in a significant way. It is noteworthy that the Canadian program described by Dowden, Blanchette and Serin (1999) lasts for 50 hours and that, internationally, rehabilitation programs of 100 hours or more are typically recommended for offenders with high levels of need.

SUMMARY OF THE MAIN FINDINGS FROM THE THREE STUDIES CONDUCTED

This report has described a large number of findings in the three studies. The attached Appendices also contain a range of further, more detailed data. It is important that major conclusions are not lost amongst so much detail. We wish to summarise the most significant findings and outcomes as follows:

- 1) The studies confirm that high levels of anger exist in the prison population. Thus effective anger management programs are required.
- 2) The anger management programs studied are, in general, producing very small effects, though the changes are in the right direction. Few statistically significant changes occur from pre-to post group assessments, with the exception of improved anger knowledge. The treated group do not improve significantly more than the untreated control group.

- 3) South Australia and Western Australia show broadly the same pattern in terms of impact of intervention.
- 4) No marked differences in outcome occur for prison compared to community programs.
- 5) Where improvements occur with treatment, they are often maintained over the follow-up period
- 6) Some program participants have much better outcomes than others.
- 7) Participants who have high pre-treatment levels of various aspects of anger show more benefit from the programs.
- 8) Participants who are motivated and “ready” for treatment have better outcomes.
- 9) Participants classified as high in risk, need and responsivity shows a tendency to have better outcomes than those not so classified.
- 10) Programs that are high in program integrity have better outcomes,
- 11) A series of measures have been developed during the project which may be useful assessment and screening devices for use in future programs.

RECOMMENDATIONS FOR IMPROVING ANGER MANAGEMENT OUTCOMES

In conclusion, we would like to consider some implications of the findings for correctional policy and for the future development of anger management programs. We recommend the following:

- A) **Given the high scores obtained by the offender participants on a range of anger measures, anger management interventions should be maintained as an important component of any portfolio of “core programs”.** The links between anger and violence are increasingly recognised in the research and rehabilitation literatures. Virtually all well-developed correctional systems internationally deliver programs of this sort. Both South Australia and Western Australia have well developed and managed systems for the delivery of anger management programs and for staff training which will be of enormous benefit in future development of the programs.
- B) **There should be a move away from the strategy of “blanket” delivery of the programs to all offenders referred because of violent histories or because they have been**

informally deemed to be suitable. The results of the present study show this approach to be ineffective.

- C) **Anger management should be offered to offenders on the basis of the likelihood they will benefit. Thus all referrals should receive a pre-treatment assessment to determine suitability.**
- D) **The suitability assessment should be based on the findings of the present study and should comprise, as a minimum, formal psychometric measures of anger proneness and of readiness/motivation for treatment.** Some of the measures used in and developed for the present study (see Appendices) could be used routinely, with some adaptation, for this purpose. Cut-off points would need to be developed for the various anger and motivational scales.
- E) **Given the low impact of the current programs, they should be developed further and made more intensive.** Intensiveness can be addressed in two (inter-related) ways – by extending the length of the programs and by revising the content to ensure they have a stronger “therapeutic” and less of an “educational” focus. We recommend that the programs should be at least 50 hours in length. If “high risk/high needs” offenders are targeted in the future, then even longer programs (100 hours plus) are likely to be required for such groups. The costs of increasing intensiveness are likely to be offset by the savings deriving from a more targeted and less general approach to service delivery.
- F) **Evaluation measures of the sort used in the present study should be “built-in” to the anger programs so that effectiveness can be monitored in an ongoing way.**
- G) **Program integrity monitoring needs to be developed as a routine practice.** Few correctional systems internationally have developed integrity assessment methodologies. The integrity assessments developed for the present study offer only one approach to this difficult task.

References

- Andrews, D. A., & Bonta, J. (1994). Psychology of Criminal Conduct. Cincinnati, OH: Anderson.
- Arboleda-Florez, J., Crisanti, A., Rose, S., & Holley, H. (1994). Measuring Aggression on Psychiatric Inpatient Units: Development and Testing of the Calgary General Hospital Aggression Scale. International Journal of Offender Therapy and Comparative Criminology, *38*(3), 183-204.
- Arias, I., & Beach, R.H. (1987). Validity of self reports of marital violence. Journal of Family Violence, *2*, 139-142.
- Armstead, C. A., Lawler, K. A., Gorden, G., Cross, J., & Gibbons, J. (1989). Relationship of racial stressors to blood pressure responses and anger expression in Black college students. Health Psychology, *8*, 541-556.
- Averill, J. R. (1982). Anger and Aggression: An Essay on Emotion. New York: Springer-Verlaag.
- Beck, R., & Fernandez, E. (1998). Cognitive-behavioural therapy in the treatment of anger: a meta-analysis, Cognitive Therapy and Research, *22*, 63-74.
- Bonta, J., & Gendreau, P. (1990). Reexamining the cruel and unusual punishment of prison life. Law & Human Behavior, *14*, 347-372
- Bubner, S. (1999). Utility of motivational screening in the assessment and treatment of substance use with offenders. Masters thesis. University of South Australia, Adelaide, South Australia, Australia.
- Copello, A. G., & Tata, P. R (1990). Violent behaviour and interpretive bias: An intervention study of the resolution of ambiguity in violent offenders. British Journal of Clinical Psychology, *29*, 417-428.
- Cornell, D. G., Peterson, C. S., & Richards, H. (1999). Anger as a Predictor of Aggression Among Incarcerated Adolescents. Journal of Consulting and Clinical Psychology, *67*, 108-115.
- Dalton J. E., Blain, G. H., & Bezier, B. (1998). State-Trait Anger Expression Inventory scores of male sexual offenders. International Journal of Offender Therapy & Comparative Criminology, *42*(2), 141-148.

Day, A., Maddicks, R., & McMahon, D. (1993). Brief psychotherapy in two-plus-one sessions with a young offender population. Behavioural & Cognitive Psychotherapy, 21, 357-369.

Day, A., & Howells, K. (2001). Psychological treatments for rehabilitating offenders: Evidence-based practice comes of age. Australian Psychologist, (in press).

Deffenbacher, J. L., Story, D. A., Brandon, A. D., Hogg, J. A., & Hazaleus, S. L. (1988). Cognitive and Cognitive-Relaxation Treatments of Anger. Cognitive Therapy and Research, 12, 167-184.

Deffenbacher, J.L., & Swain, R.C. (1999). Anger expression in Mexican American and white non-hispanic adolescents. Journal of Counseling Psychology, 46, 61-69.

Department for Correctional Services. (1996). System operating procedure No. 4. Incident reporting and recording. Adelaide: DCS.

Dixon, B., & Polaschek, D. (1992). Development and evaluation of a treatment programme for violent offenders. Paper presented to conference on Progress in Forensic Psychiatry, Auckland, New Zealand.

Dowden, C., Blanchette, K., & Serin, R. (1999). Anger Management Programming for Federal Male Inmates: An Effective Intervention. Research Report R82e. Correctional Service Canada.

Edmondson, C.B., & Conger, J.C. (1996). A review of treatment efficacy for individuals with anger problems: Conceptual, assessment and methodological issues. Clinical Psychology Review, 16, 251-275.

Ford, B. D. (1991). Anger and irrational beliefs in violent inmates. Personality & Individual Differences, 12, 211-215.

Forgays, D. K. Spielberger, C. D., Ottaway, S. A., & Forgays, D. G. (1998). Factor structure of the State-Trait Anger Expression Inventory for middle-aged men and women. Assessment, 5(2), 141-155.

Frey, K. S., Hirschstein, M. K., & Guzzo, B. A. (2000). Second step: preventing aggression by promoting social competence. Journal of Emotional & Behavioral Disorders, 8(2), 102-112.

Gembora, L. B. (1986). Anger and prisoners. Unpublished doctoral dissertation. Queen's University, Kingston, Ontario.

Goldstein, A.P., & Glick, B. (1996). Aggression replacement training: methods and outcomes. In C.R. Hollin & K. Howells (Eds.) Clinical approaches to working with young offenders (pp. 151-164). Chichester: Wiley.

Greenglass, E. R., & Burke, R. J. (2000). The relationship between hospital restructuring, anger, hostility and psychosomatics in nurses. Journal of Community & Applied Social Psychology, 10(2), 155-161.

Heather, N., Rollnick, S., & Bell, A. (1993). Predictive validity of the Readiness to Change Questionnaire. Addiction, 88, 1667-1677.

Hollin, C.R. (1995). The meanings and implications of programme integrity. In J. McGuire What Works: Reducing Reoffending. John Wiley and Sons: Chichester.

Hollin, C. R. (1999). Treatment programs for offenders. International Journal of Law and Psychiatry, 22, 361-372.

Howells, K. (1996). The psychological management of violence in clinical and forensic settings: pitfalls and remedies. Psychiatry, Psychology and Law, 3, 71-76.

Howells, K. (1998). Cognitive behavioural therapy for anger, aggression and violence. In N. Tarrier (Ed.) Cognitive Behavioural Therapy for Complex Cases. Chichester: Wiley.

Howells, K., & Day, A. (1999). The Rehabilitation of Offenders: International Perspectives Applied to Australian Correctional Systems, Trends and Issues in Crime and Criminal Justice, 112, Australian Institute of Criminology: Canberra.

Howells, K., Watt, B., Hall, G., & Baldwin, S. (1997). Developing programmes for violent offenders. Legal and Criminological Psychology, 2, 117-128.

Indermaur, D. (1996). Offender psychology and sentencing. Australian Psychologist, 31, 15-19.

Indermaur, D. (1999). Patterns of violence. Paper presented at the 3rd National Outlook Symposium on Crime in Australia. Canberra: Australian Institute of Criminology.

Kay, S. R., Wolkenfeld, F., & Murrill, L. M. (1988). Profiles of aggression among psychiatric patients: I. Nature and prevalence. The Journal of Nervous and Mental Disease, 176, 539-546.

Kroner, D.G. & Reddon, J. R. (1995). Anger and psychopathology in prison inmates. Personality & Individual Differences, 18, 783-788.

Levey, S., & Howells, K. (1990). Anger and its management. Journal of Forensic Psychiatry, 1, 305-327.

Losel, F. (1996). Working with young offenders: the impact of meta-analyses. In C.R. Hollin and K. Howells (Eds). Clinical Approaches to Working with Young Offenders, Chichester: Wiley.

Loza, W., & Loza-Fanous, A. (1999a). The fallacy of reducing rape and violent recidivism by treating anger. International Journal of Offender Therapy & Comparative Criminology, 43(4), 492-502.

Loza, W., & Loza-Fanous, A. (1999b). Anger and prediction of violent and nonviolent offenders' recidivism. Journal of Interpersonal Violence, 14(10), 1014-1029.

Mals, P., Howells, K., & Day, A. (1999). Adapting violence programs for the Aboriginal Offender. Journal of Offender Rehabilitation, 30, 121-135.

Mals, P., Howells, K., Day, A., & Hall, G. (2000). Adapting violence rehabilitation programs for the Australian Aboriginal offender. Journal of Offender Rehabilitation, 30, 121-135.

Mayne T.J., & Ambrose T.K. (1999). Research review on anger in psychotherapy, Journal of Clinical Psychology, 55, 353-363.

McDougall, C., Barnett, R. M., Ashurst, B., & Willis, B. (1987). Cognitive control of anger. In. Applying psychology to imprisonment: Theory and practice (pp. 303-313). London: Her Majesty's Stationery Office Books.

McDougall, C., & Boddis, S. (1991). Discrimination Between Anger and Aggression: Implications for Treatment. Issues in Criminological and Legal Psychology, 2, 101-106.

McDougall, C., Venables, P., & Roger, D. (1991). Aggression, anger control and emotion control. Personality & Individual Differences, 12, 625-629.

McGuire, J. (1995) (Ed). What Works in Offender Rehabilitation? Chichester: Wiley.

McGuire, J. (1998). Alternative to custodial sentences: Effectiveness and potential for development, Memorandum of evidence to the Home Affairs Committee, House of Commons.

McMurrin, M. E. V., Richardson, C., Street, H., Ahmadi, S., & Cooper, G. (2000). Referrals for anger and aggression in forensic psychology outpatient services. Journal of Forensic Psychiatry, 11, 206-213.

McMurrin, M, Tyler, P., Hogue, T., Cooper, K., Dunseath, W., & McDaid, D. (1998). Measuring motivation to change in offenders. Psychology, Crime and Law, 4, 43-50.

Medd, J., & Tate, R. L. (2000). Evaluation of an anger management therapy programme following acquired brain injury: a preliminary study. Neuropsychological Rehabilitation, 10, 185-201.

Miller, W. R., & Rollnick, S. (1991). Motivational Interviewing: Preparing People for Change. New York: Guildford Press.

Mills, J. F., Kroner, D. G., & Forth, A. E. (1998). Novaco Anger Scale: Reliability and validity within an adult criminal sample. Assessment, 5, 237-248.

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

Myers, W. C., & Monaco, L. (2000). Anger experience, styles of anger expression, sadistic personality disorder, and psychopathy in juvenile sexual homicide offenders. Journal of Forensic Sciences, 45(3), 698-701.

Novaco, R. (1975). Anger Control: The Development and Evaluation of an Intervention Treatment. Lexington, MA: D.C. Health.

Novaco, R. (1994). Anger as a risk factor for violence among the mentally disordered. In J. Monahan & H.J. Steadman (Eds). Mental Disorder and Violence: Developments in Risk Assessment. Chicago: University of Chicago Press.

Novaco, R. (1997) Remediating anger and aggression with violent offenders. Legal and Criminological Psychology, 2, 77-88.

O'Leary, C. A. (2000). Reducing aggression in adults with brain injuries. Behavioral Interventions, 15(3), 205-216.

O'Neill, H. (1995). Anger: The assessment and treatment of problematic anger, part 1. British Journal of Occupational Therapy, 58(10), 427-431.

Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: applications to addictive behaviors. American Psychologist, 47, 1102-1114.

Reilly, P. M., & Shopshire, M. S. (2000). Anger management group treatment for cocaine dependence: preliminary outcomes. American Journal of Drug & Alcohol Abuse, 26(2), 161-177.

Renwick, S. J., Black, L., Ramm, M., & Novaco, R. W. (1997). Anger treatment with forensic hospital patients. Legal and Criminological Psychology, 2, 103-116.

Saunders, D.G. (1991). Procedures of adjusting self-reports of violence for social desirability bias. Journal of Interpersonal Violence, 6, 336-344.

Serin, R. (1998). Treatment responsivity, intervention, and reintegration: A conceptual model. Forum on Corrections Research, 10, 29-32.

Serin, R., & Kennedy, S. (1997). Treatment readiness and responsivity: Contributing to effective correctional programming. Research Report, R-54. Ottawa, ON: Correctional Services of Canada.

Sharkin, B. S. (1996). Understanding anger: Comment on Deffenbacher, Oetting, et al. (1996), Deffenbacher, Lynch, et al. (1996), and Kopper and Epperson (1996). Journal of Counselling Psychology, 43, 166-169.

Slaton, B. J., Kern, R. M., & Curlette, W. L. (2000). Personality profiles of inmates. Journal of Individual Psychology, 56, 88-109.

Smith, L. L., Smith, J. N., & Beckner, B. M. (1994). An anger-management workshop for women inmates. Families in Society, 75(3), 172-175.

Spielberger, C.D. (1988). State-Trait Anger Expression Inventory: Revised research edition. Odessa, FI: PAR.

Speilberger, C. D. (1991). State-trait Anger Expression Inventory Revised Research Edition, Professional Manual. Odessa, FL: Psychological Assessment Resources.

Stermac, L. E. (1986). Anger control treatment for forensic patients. Journal of Interpersonal Violence, 1, 446-457.

Stevens, J. (1996). Applied Multivariate Statistics for the Social Sciences (3rd Ed). New Jersey: Lawrence Erlbaum.

Swaffer, T., & Epps, K. (1999). The psychometric assessment of anger in male and female adolescents resident at a secure youth treatment centre. Journal of Adolescence, 22(3), 419-422.

Swaffer, T., & Hollin. C. R. (1997). Adolescents' experiences of anger in a residential setting. Journal of Adolescence, 20, 567-575.

Tabachnick, B. G., & Fidell, L. S. (1996). Using Multivariate Statistics. New York: Harper Collins.

Walker, J., & McDonald, D. (1995). The over-representation of indigenous people in custody in Australia. Trends and Issues in Crime and Criminal Justice; No. 47. Canberra: Australian Institute of Criminology.

Waltz, J., Addis, M. E., Koerner, K., & Jacobsen, N. S. (1993). Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. Journal of Consulting and Clinicial psychology, 61, 620-630,

Ward, A., & Dockerill, J. (1999). The predictive accuracy of the Violent Offender Treatment Program Risk Assessment Scale. Criminal Justice and Behaviour, 26, 125-140.

Watt, B. (1997). Skills training for aggression control. Masters Thesis. Edith Cowan University, Perth, Western Australia.

Watt, B., & Howells, K. (1999). Skills training for aggression control: evaluation of an anger management programme for violent offenders. Legal and Criminological Psychology, 4(2), 285-300.

Yudofsky, S. C., Silver, J. M., Jackson, W., Endicott, J., & Williams, D. (1986). The Overt Aggression Scale for the objective rating of verbal and physical aggression. American Journal of Psychiatry, 143, 35-39.

Zamble, E., & Porporino, F. (1990). Coping, imprisonment, and rehabilitation: Some data and their implications. Criminal Justice & Behavior, 17, 53-70.

5.66099423

A

AUST INST OF CRIMINOLOGY



MON18939

4005583

