

How do risk factors for sexual and non-sexually violent offenders manifest and change
in custody?

Tamara Sweller

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Centre for Forensic Behavioural Science
Swinburne University of Technology
2018

Supervisors
Professor Michael Daffern
Professor Stuart Thomas

© 2018 Tamara Sweller

All rights reserved

Under the *Copyright Act 1968*, this thesis must be used only under the normal conditions of scholarly fair dealing. In particular, no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

Abstract

The current research investigated the behavioural manifestations of dynamic risk factors and their change over time through three studies. Both of these important issues contribute to decision-making throughout sexual and non-sexually violent offenders' incarceration. In Study 1, participants were members of Australian and North American organisations for registered psychologists. Through an online survey, 34 participants generated examples of predicted behavioural manifestations of dynamic risk factors and their prosocial equivalent behaviours for incarcerated sexual offenders. Risk factors were drawn from two established sexual offender risk assessment instruments: Risk for Sexual Violence Protocol (RSVP) and Violence Risk Scale: Sexual Offender version (VRS:SO). Results revealed that psychologists could identify potential behavioural manifestations, with various degrees of ease. For instance, more manifest dynamic risk factors were identified than prosocial equivalent behaviours, which may be attributed to the custodial context where there is an emphasis on risk and punishment. Study 2 was a preliminary validation of the Satisfactory Behaviour Rating Guide (SBRG), which was subsequently used in Study 3. The SBRG was developed to assist determination of satisfactory behaviour for Corrections Victoria, Australia, and ultimately to guide parole board decision-making. Participants were sexual and non-sexually violent incarcerated offenders. This study evaluated whether caseworkers' (custodial officers') ultimate behaviour rating of participants (Satisfactory, Of Minor Concern, or Of Major Concern) corresponded with official records (e.g., institutional misconducts), within each of four behaviour categories. Results revealed that at the group-level, ratings of satisfactory behaviour were associated with behaviour in custody. At the individual-level, there were discrepancies between custodial officers' overall ratings and offenders' behaviour reflected through misconducts. Custodial officers seldom rated offenders as 'Of Major Concern'. The results from Study 1 provided the basis for developing a behavioural checklist, which was implemented in Study 3 as one of four measures to evaluate change in sexual offenders' behaviour in a custody-based treatment program. The additional three measures were: pre- to post-treatment psychometric measures; Treatment Gain: Short Scale; and, the SBRG. The four measures included offender self-report and observations from custodial officers and therapists. Despite inconsistencies between group- and individual-level results, overall, changes in dynamic risk factors were observed using the different measurement approaches through treatment. Both participants and custodial officers reported an increase in positive behaviours across all

four RSVP domains. There was less consistency in the changes observed within negative behaviour (manifest dynamic risk factor) domains. The most common results from the psychometric measures revealed participants were either already in the functional range pre-treatment or remained in the dysfunctional range post-treatment. Treatment Gain scores varied across participants, while the majority of SBRG ratings were Satisfactory. Overall, there was limited consistency between the four change measures, which suggests diverse methods of measurement (e.g., offender self-report, custodial officer observations, and therapist perspectives) will likely produce different information and this needs to be considered when determining whether offenders have changed and whether dynamic risk factors persist. Overall, these three studies provide an in-depth understanding of behavioural monitoring and change measurement in custody and reveal the complexity of these tasks. The challenges associated with using multiple change measures were observed and the benefit of evaluating individual-level change was highlighted, through exploration of the discrepancies between group- and individual-level changes. Limitations of the current research are presented, in addition to future directions for research in this field.

Acknowledgements

This project has formed a central component of my life over the past six years and represents the most focused and challenging long-term project I have ever completed. It would not have been possible without the invaluable assistance and support of my supervisor, Professor Michael Daffern. Michael, your advice, academic and emotional support, knowledge, and sense of humour, were qualities I was incredibly lucky to experience. Thank you for encouraging me to make sure my PhD did not become my entire life and instead have balance where I could. I am also very thankful to Professor Stuart Thomas. Despite becoming part of my supervisory team towards the latter part of my project, your feedback and ideas were paramount in assisting me create this thesis. Michael and Stuart, you are both incredibly inspiring as researchers, but also in your outlook on work and life.

I would also like to acknowledge Professor Richard Kemp, for your guidance during the initial stages of my research. I also gained insight into a form of research in which I did not previously have experience, from Narelle Warren. Narelle, thank you for your assistance in relation to the analysis of qualitative data and contribution to the development of the published article on the first study within my thesis. I would like to thank the panel members from my Confirmation of Candidature, Mid-Candidature Review and Pre-Submission Review, who provided constructive feedback, which assisted further development and consolidation of my research.

Data collection for my primary study was challenging in many ways. It would not have been possible without the assistance of custodial staff at Long Bay. Particular thanks to Richard; you contributed a huge amount, not only to the data but also to maintaining my sanity while collecting data.

The consistent support and advice I have received from Eden, my husband, has been a stabilising force, without which I would not have been able to complete this process. I cannot express how grateful I am for your encouragement and ability to deal with my emotional rollercoaster that was my PhD candidature. I am always in awe of the way you give me perspective when I feel completely overwhelmed. Thank you also to my parents, for your constant emotional and academic support. I'm so lucky to have you both and I admire you more than I can express. Thank you also to my sister, Naomi, for those initial stages of SPSS help.

Declarations

In accordance with the Swinburne University of Technology Statement of Practice for the completion of a Higher Degree by Research, the following declarations are made:

I, Tamara Sweller, hereby declare that this thesis, titled 'How do risk factors for sexual and non-sexually violent offenders manifest and change in custody?', contains no material which has been accepted for the award to myself, the candidate, of any other degree or diploma, except where due reference is made in the text of the examinable outcome. I declare that, to the best of my knowledge, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the examinable outcome. Where the work is based on joint research or publications, the thesis discloses the relative contributions of the respective workers or authors.

Signature of candidate:  _____

Date: 2/05/2018

Table of Contents

Abstract	i
Acknowledgements	iii
Declarations.....	iv
Table of Contents	v
Index of Tables.....	xi
Index of Figures	xiv
List of Abbreviations.....	xv
Chapter 1: Literature Review	1
1.1.1 Overview	1
1.1.2 Incidence and Prevalence of Sexual Offending	2
1.1.3 Impact of Sexual Offending	4
1.1.4 Role of Treatment	7
1.1.5 Treatment Effectiveness.....	8
1.1.6 Theories of Offender Rehabilitation	14
1.1.6.1 Risk-Need-Responsivity (RNR) principles.....	15
1.1.7 Risk Factors.....	16
1.1.7.1 Risk factors for sexual offending.	23
1.1.8 Risk Assessment	24
1.1.8.1 Types of risk assessment.	24
1.1.9 Protective Factors.....	27
1.1.9.1 Protective factor assessment.	30
1.1.10 Behavioural Monitoring in Custody.....	31
1.1.10.1 Behavioural consistency.....	31
1.1.10.2 Behavioural manifestations of risk factors.....	34
1.1.10.3 Behavioural manifestations of protective factors.....	42
1.1.10.4 Practical issues related to behavioural observation.....	43
1.1.11 Behavioural Monitoring in a Therapeutic Community	46
1.1.11.1 Therapeutic community environment.	46
1.1.11.2 Behavioural assessment.	46
1.1.11.3 Challenges with observation.	47
1.1.12 Evaluating Change in Offenders	48
1.1.12.1 Psychometric measures through self-report.	49
1.1.12.2 Structured clinical evaluation.....	52
1.1.12.3 Number of time points measured.	54

1.1.12.4 Change findings.	55
1.1.12.5 Determination of change.	57
1.1.13 Individual-Level Change.....	58
1.1.14 Summary	60
1.1.15 Current Research.....	61
1.1.15.1 Study 1.	61
1.1.15.2 Study 2.	61
1.1.15.3 Study 3.	62
Chapter 2: Study 1.....	64
2.1.1 Sexual Offender Risk Assessment	64
2.1.1.1 Behavioural manifestations of risk factors.....	64
2.1.2 Behavioural Monitoring in Custody.....	68
2.1.3 Current Study	70
2.2 Method	71
2.2.1 Research Design.....	71
2.2.2 Participants.....	71
2.2.3 Materials.....	72
2.2.4 Procedure.....	72
2.2.5 Data Analysis	73
2.3 Results	74
2.3.1 Descriptive Statistics	74
2.4 Discussion	86
2.4.1 Overview of Results.....	86
2.4.2 Exploration of Results.....	86
2.4.2.1 Risk vs prosocial behaviours.....	86
2.4.2.2 Variation between risk factors.	88
2.4.2.3 Overlap between risk factors and behaviours.	88
2.4.2.4 OPB vs OAB.....	89
2.4.3 Limitations and Future Research Directions	90
2.4.4 Conclusions	92
Chapter 3: Study 2.....	94
3.1.1 Use of Institutional Behaviour	94
3.1.2 Change Measurement.....	96
3.1.2.1 Importance of change measurement.....	96
3.1.2.2 Institutional misconducts.	97
3.1.2.3 Dynamic risk assessment.	98
3.1.2.4 Behavioural manifestations of risk factors.....	98

3.1.2.5 Staff observations.....	100
3.1.2.6 Limitations of using official records.....	101
3.1.3 Parole System Review	101
3.1.4 Relevance of Institutional Behaviour.....	103
3.1.5 Satisfactory Behaviour Rating Guide.....	103
3.1.6 Current Study	104
3.2 Method	104
3.2.1 Participants.....	104
3.2.2 Materials.....	105
3.2.3 Procedure.....	106
3.2.3.1 Violent behaviour.....	106
3.2.3.2 Involvement in substance abuse or like behaviour.	109
3.2.3.3 Attitude to employment, education and/or rehabilitation.	112
3.2.3.4 Response to direction and supervision.	113
3.2.4 Data Analysis	114
3.3 Results.....	115
3.3.1 Descriptive Statistics.....	115
3.3.2 Correlation Between Overall Rating and Incidents	115
3.3.3 Predicting the Overall Rating.....	116
3.3.3.1 Group-level analysis.	116
3.3.3.2 Individual-level analysis.	117
3.4 Discussion	118
3.4.1 Overview of Results.....	118
3.4.2 Custodial Officers' Ratings.....	119
3.4.3 Relation Between Behavioural Categories and Overall Ratings.....	120
3.4.4 Discrepancies Between Group- and Individual-Level	122
3.4.5 Behavioural Manifestations of Risk Factors	123
3.4.6 Use of Staff Ratings	124
3.4.7 Limitations	125
3.4.8 Future Directions.....	126
3.4.9 Conclusions.....	127
Chapter 4: Study 3.....	129
4.1.1 Behavioural Manifestations of Risk Factors	129
4.1.2 Guidance in Behavioural Observation	130
4.1.3 Change Measurement.....	132
4.1.3.1 Self-report.	132
4.1.3.2 Observer ratings.	133

4.1.4 Current Study	135
4.1.4.1 Behavioural checklists.	136
4.1.4.2 Multiple change measures.....	137
4.1.4.3 Aims and hypotheses.	138
4.2 Method	139
4.2.1 Participants.....	139
4.2.1.1 Custody-Based Intensive Treatment (CUBIT) program.	139
4.2.1.2 Participant selection.	140
4.2.1.3 Participant information.	141
4.2.1.4 Data availability.	142
4.2.2 Materials.....	147
4.2.2.1 Behavioural checklists.	147
4.2.2.2 Psychometric measures.	148
4.2.2.3 Treatment Gain: Short Scale.	154
4.2.2.4 Satisfactory Behaviour Rating Guide (SBRG). The development of the SBRG (Daffern et al., 2014) was described in Chapter 3.	154
4.2.3 Procedure.....	155
4.2.4 Ethical Considerations	159
4.2.4.1 Conflict of interest.	159
4.2.4.2 Consent.....	160
4.2.4.3 Confidentiality.	160
4.2.5 Data Analysis	160
4.2.5.1 Behavioural checklists.	162
4.2.5.2 Psychometric measures.	170
4.2.5.3 Clinically significant change and Treatment Gain scale.....	179
4.2.5.4 Overall change.	179
4.3 Results	179
4.3.1 Behavioural Checklists.....	180
4.3.1.1 Comparison between reports of behaviours.....	180
4.3.1.2 Group-level participant self-reported change.....	186
4.3.1.3 Group-level custodial officer observed change.....	189
4.3.1.4 Participant self-report versus custodial officer observations.	194
4.3.1.5 Individual-level self-reported and observed change.	195
4.3.2 Psychometric Measures.....	202
4.3.2.1 Descriptive statistics.....	202
4.3.2.2 Group-level program change.	204
4.3.2.3 Individual-level change.....	206

4.3.3 Treatment Gain: Short Scale and Relation to Psychometric Measures	215
4.3.4 Correspondence Between Change Measures	217
4.3.4.1 Participant 1.	217
4.3.4.2 Participant 2.	218
4.3.4.3 Participant 3.	218
4.3.4.4 Participant 4.	219
4.3.4.5 Participant 5.	220
4.3.4.6 Participant 6.	220
4.3.4.7 Participant 7.	221
4.3.4.8 Participant 8.	221
4.3.4.9 Participant 9.	222
4.3.4.10 Participant 10.	222
4.3.4.11 Participant 11.	223
4.3.4.12 Participant 12.	224
4.3.4.13 Participant 13.	225
4.3.4.14 Participant 14.	225
4.3.4.15 Participant 15.	226
4.3.4.16 Participant 16.	226
4.3.4.17 Participant 17.	227
4.3.4.18 Participant 18.	228
4.3.4.19 Participant 19.	228
4.4 Discussion	229
4.4.1 Overview of Results.....	230
4.4.1.1 Group-level change.	230
4.4.1.2 Individual-level change.	230
4.4.2 Behavioural Checklists.....	231
4.4.2.1 Comparison between reports of behaviours.	231
4.4.2.2 Domains related to behavioural checklists.....	234
4.4.2.3 Individual-level self-reported and observed change.	240
4.4.3 Psychometric Measures.....	240
4.4.3.1 Group-level change.	241
4.4.3.2 Individual-level change.	242
4.4.3.3 Domains in relation to psychometric measures.	244
4.4.4 Integrating Change Measure Results	248
4.4.4.1 Individual-level change.	253
4.4.5 Case Study Outlining a Clinician's Decisions	259
4.4.6 Limitations and Challenges.....	260

4.4.6.1 Limitations.	261
4.4.6.2 Practical issues and implications.....	261
4.4.7 Recommendations for Measuring Change.....	271
4.4.8 Future Research.....	273
4.4.9 Conclusions.....	276
Chapter 5: General Discussion.....	277
5.1.1 Overview of Research.....	277
5.1.1.1 Study 1.	277
5.1.1.2 Study 2.	278
5.1.1.3 Study 3.	279
5.1.2 Comparison Between Study Results.....	281
5.1.2.1 Differences between risk factors.	282
5.1.2.2 Observation of negative and positive behaviours.	283
5.1.2.3 SBRG results.....	284
5.1.3 Integration of Results.....	285
5.1.3.1 Psychological domains.....	285
5.1.3.2 Observer involvement.	288
5.1.4 Clinical and Policy Implications.....	291
5.1.4.1 Behavioural manifestations of risk factors.....	291
5.1.4.2 Change measurement.	292
5.1.5 Challenges and Limitations.....	297
5.1.5.1 Challenges within current research.	297
5.1.5.2 Limitations of current research.	297
5.1.6 Future Research.....	299
5.1.7 Summary and Conclusions.....	300
References.....	303
Appendices.....	327
Appendix A: Published Journal Article.....	328
Appendix B: Psychologists' Survey.....	345
Appendix C: Behavioural Checklist.....	349
Appendix D: Ethics Approval.....	352
Appendix E: Participant Information Sheet and Consent Form.....	358

Index of Tables

Chapter 2:

Table 1.	<i>Risk Factors and Themes of Negative Behavioural Manifestation Examples, with Examples of Behaviours within Themes.....</i>	76
Table 2.	<i>Risk Factors and Themes of Positive Behavioural Manifestation Examples, with Examples of Behaviours within Themes.....</i>	80

Chapter 3:

Table 3.	<i>Number of Participants from Each Prison.....</i>	105
Table 4.	<i>Frequencies of Violence Subcategories.....</i>	107
Table 5.	<i>Total Violent Incidents.....</i>	108
Table 6.	<i>Frequency of Total Violent Incidents Directed Towards Each Victim Type.....</i>	109
Table 7.	<i>Frequency of Total Substance Abuse/Use Incidents in Each Subcategory.....</i>	111
Table 8.	<i>Frequency of Total Substance Use/Abuse Incidents.....</i>	112
Table 9.	<i>Average Pay Frequency.....</i>	113
Table 10.	<i>Frequency of Times Banned from Employment.....</i>	113
Table 11.	<i>Frequency of Response to Direction and Supervision Incidents.....</i>	114
Table 12.	<i>Overall SBRG Rating Frequencies.....</i>	115
Table 13.	<i>Prevalence of Behaviour Categories.....</i>	115
Table 14.	<i>Association between Overall SBRG Rating and Variable Categories.....</i>	116

Chapter 4:

Table 15.	<i>Participant Risk Level and Offence Type.....</i>	141
Table 16.	<i>Number of Participants with Available Data for Each Change Measure.....</i>	142
Table 17.	<i>Behavioural Checklist Descriptive Information.....</i>	143

Table 18.	<i>Problems with Substance Abuse Contingency Table.....</i>	163
Table 19.	<i>Risk Factors Categorised within Psychological Domains.....</i>	165
Table 20.	<i>Data Distribution for Psychological Domains.....</i>	166
Table 21.	<i>Data Distribution for Psychometric Measures using Original Dataset.....</i>	171
Table 22.	<i>Skewness Index for Psychometric Measures using Multiple Imputation Datasets.....</i>	172
Table 23.	<i>Kurtosis Index for Psychometric Measures using Multiple Imputation Datasets.....</i>	173
Table 24.	<i>Classification of Psychometric Measures into Psychological Domains.....</i>	175
Table 25.	<i>CSC Calculations Including Means, Standard Deviations, and Cut-Off Scores Used for Each Measure.....</i>	177
Table 26.	<i>Post-Treatment Change Categories.....</i>	178
Table 27.	<i>Comparison between Participants' and Custodial Officers' Reports of Each Risk Factor.....</i>	181
Table 28.	<i>Comparison between Two Behaviours of Each Risk Factor.....</i>	184
Table 29.	<i>Participant Self-Report of Group-Level Change in Each Domain throughout Treatment.....</i>	186
Table 30.	<i>Comparison between Custodial Officer Observations throughout Treatment.....</i>	189
Table 31.	<i>Custodial Officer Observations of Group-Level Change throughout Treatment.....</i>	192
Table 32.	<i>Comparison between Participant Self-Report and Custodial Officer Observations of Behaviour.....</i>	194
Table 33.	<i>Self-Reported Individual-Level Change in Each Negative Domain throughout Treatment.....</i>	196
Table 34.	<i>Self-Reported Individual-Level Change in Each Positive Domain throughout Treatment.....</i>	197
Table 35.	<i>Custodial Officer Observations of Individual-Level Change in Each Negative Domain throughout Treatment.....</i>	199
Table 36.	<i>Custodial Officer Observations of Individual-Level Change in Each Positive Domain throughout Treatment.....</i>	200

Table 37.	<i>Descriptive Statistics for Pre- and Post-Treatment Psychometric Measures.....</i>	202
Table 38.	<i>Group-Level Change through Pre- to Post-Treatment Psychometric Measures.....</i>	204
Table 39.	<i>Group-Level Change through Pre- to Post-Treatment Psychometric Measures.....</i>	205
Table 40.	<i>Reliable Change and Clinically Significant Change for Each Participant.....</i>	207
Table 41.	<i>Number of Participants who Made Reliable Changes per Psychometric Measure.....</i>	210
Table 42.	<i>Number of Participants in Each Clinically Significant Change Category per Psychometric Measure.....</i>	213
Table 43.	<i>Number of Psychometric Measures in Each Clinically Significant Change Category and Treatment Gains Score per Participant.....</i>	215

Index of Figures

Chapter 3:

<i>Figure 1.</i>	Participant self-report of group-level change in each negative domain throughout treatment.....	187
<i>Figure 2.</i>	Participant self-report of group-level change in each positive domain throughout treatment.....	187
<i>Figure 3.</i>	Custodial officer observations of group-level change in each negative domain throughout treatment.....	190
<i>Figure 4.</i>	Custodial officer observations of group-level change in each positive domain throughout treatment.....	191

List of Abbreviations

BMS	Bumby Molest Scale
BRS	Bumby Rape Scale
CI	Confidence Interval
CISS	Coping Inventory for Stressful Situations
CSC	Clinically Significant Change
CSNSW	Corrective Services New South Wales
CUBIT	Custody-Based Intensive Treatment Unit
CUSI	Coping Using Sex Inventory
df	Degrees of Freedom
ERD	Earliest Release Date
HCR-20	Historical Clinical Risk Management
HCR-20v3	Historical Clinical Risk Management version 3
HLM	Hierarchical Linear Modeling
LS-UCLA	Revised University of California Loneliness Scale
M	Mean
MC-SDS	Marlowe-Crowne Social Desirability Scale
MDN	Mental Disorder Negative
MDP	Mental Disorder Positive
MNN	Manageability Negative
MNP	Manageability Positive
OAB	Offence Analogue Behaviour
OPB	Offence Paralleling Behaviour
ORB	Offence Reduction Behaviour
PAB	Prosocial Alternative Behaviour
PAN	Psychological Adjustment Negative
PAP	Psychological Adjustment Positive
PCC	Psychology of Criminal Conduct
PDS	Paulhus Deception Scale
PIMS	Prisoner Information Management System
RC	Reliable Change
RCI	Reliable Change Index

RNR	Risk-Need-Responsivity principles
RSVP	Risk for Sexual Violence Protocol
SAN	Social Adjustment Negative
SAP	Social Adjustment Positive
SAPROF	Structured Assessment of Protective Factors for Violent Risk
SBRG	Satisfactory Behaviour Rating Guide
SD	Standard Deviation
SE	Standard Error
SFQ	Sexual Fantasy Questionnaire
SIS	Social Intimacy Scale
SPIn	Service Planning Instrument
SSEI	Social Self Esteem Inventory
START	Short-Term Assessment of Risk and Treatability
SVR-20	Sexual Violence Risk scale
TRRG:SV	Treatment Readiness Responsivity Gain Scale: Short Version
VRS	Violence Risk Scale
VRS:SO	Violence Risk Scale: Sexual Offender version
WSFQ	Wilson Sexual Fantasy Questionnaire

Chapter 1: Literature Review

The focus of this thesis is the observation and assessment of sexual and non-sexually violent offenders' dynamic risk factors in custody, and the methods by which change in behavioural manifestations of these risk factors can be measured. The current chapter will provide a review of the relevant extant international literature, primarily concentrating on sexual offenders.

1.1.1 Overview

This chapter will commence with an outline of the incidence and prevalence of sexual offending, and its widespread impact. The far-reaching impact of sexual offending necessitates the implementation of interventions to reduce its prevalence. Custody-based intervention for sexual offenders will be discussed, including the effectiveness of treatment programs and the difficulties faced in this evaluation process. The development of treatment programs requires an understanding of theories of offender rehabilitation. A key theory, which forms the foundation of significant research and clinical practice to date, is the Psychology of Criminal Conduct (Andrews & Bonta, 2010). This theory underpins the Risk-Need-Responsivity principles (Andrews, Bonta, & Hoge, 1990), which are crucial in the understanding of risk assessment and change measurement. The review will subsequently focus on risk factors, including both static and dynamic factors. The assessment of dynamic risk factors has received much attention; in order to effectively assess these factors, it is crucial to have an understanding of what exactly is being assessed. There is ongoing debate amongst researchers and clinicians in relation to the conceptualisation of dynamic risk factors (Beech & Ward, 2004; Douglas & Skeem, 2005; Heffernan & Ward, 2015; Mann, Hanson, & Thornton, 2010; Ward, 2016; Ward & Beech, 2006). More recently, there has been an increase in the focus on protective factors in addition to risk factors, such that a more holistic assessment can be made in relation to offenders' changing risk of recidivism over time (Lösel & Farrington, 2012; Monahan & Skeem, 2016; Polaschek, 2017; Serin, Chadwick, et al., 2016).

It is important to assess the presence of, and changes in risk factors while offenders are incarcerated, to gain a thorough understanding of offenders' risk of recidivism. This is of critical importance for offenders being considered for release into the community. However, this assessment process is challenging in the custodial context. The custodial environment is more structured than in the community and offenders are faced with different situations, interactions and triggers; additionally,

there may be an absence of offending targets. The expression of incarcerated offenders' risk and protective factors may be altered as a consequence of this environment; therefore, the relevance of the expression of these factors in custody, for the assessment of their level of risk in the community, must be determined. Further, behavioural change over time in custody must be observed and assessed to assist in the evaluation of change in risk of recidivism. Processes involved in monitoring and measuring behaviours in a custodial context will be discussed, focusing on the behavioural manifestations of dynamic risk factors and the ways in which behavioural change can be effectively measured. These issues are important for the identification of offenders' ongoing areas of need, and for prediction of future criminal behaviour. Ultimately, these assessments should be considered when determining an offender's likelihood for future violence and thus suitability for release to parole. At the conclusion of this chapter, the aims and hypotheses for the three studies within the current research will be described.

While part of this thesis relates to non-sexually violent offending, the focus is predominantly on sexual offending. Therefore, this chapter presents a review of the literature pertaining to sexual offenders. A brief review of non-sexual offender literature will be provided.

1.1.2 Incidence and Prevalence of Sexual Offending

There are different methods by which to calculate the extent to which sexual offending is present in the community; for instance, the number of reported offences, the number of victims reporting to police (which could include multiple offences), the number of offenders charged (which could similarly include multiple offences), and the number of offenders whose charges resulted in convictions. The terms incidence and prevalence are at times used interchangeably; however, their definitions are distinct. Incidence refers to the number of separate victimisations, or incidents, perpetrated against people within a group during a specific time period. Prevalence refers to the number of people within a group who are victimised during a specific time period, such as the person's lifetime or the previous 12 months (Wiseman, 2015).

It can be difficult to measure and compare the prevalence and incidence of sexual offending between jurisdictions, due to the differences between methods of calculation as described above. Further, the definition of sexual offending differs between jurisdictions (Wiseman, 2015). In 2013, the definition was modified in the United States, such that the term "forcible" was removed from the offence name of rape. Across jurisdictions in the United States, 160 per 100,000 individuals aged 12 and older

reported rape/sexual assault to the National Crime Victimization Survey in 2015 (Truman & Morgan, 2016). Based on data from the Federal Bureau of Investigation (FBI), the rate of rapes known to law enforcement, using the revised definition, was 40.4 per 100,000 individuals in 2016 (FBI, Crime in the United States, 2016). In the United Kingdom, prevalence data are also collected; however, it is reported in a different manner, which limits direct comparisons. The most recent estimates from the self-report module on intimate violence in the Crime Survey for England and Wales revealed that in the year ending March 2016, 2.0% of males and females aged 16 to 59 reported experiencing sexual assault (Office for National Statistics, 2017). In this period, there were 106,098 police recorded sexual offences (Office for National Statistics, 2017). However, there was no available information regarding the rate, which this number represented.

There are a number of challenges associated with the different calculation methods in relation to ascertaining the true prevalence of sexual offending. Not all victims of sexual offences report these offences to police. For instance, according to the Australian Bureau of Statistics' Personal Safety Survey in 2016, 87% of women who reported in the survey that they were the victim of a sexual assault by a male in the preceding year, did not report these offences to police. Further, some sexual offences do not have identifiable victims (e.g., child abuse material offences); therefore, the number of victims cannot be accurately determined.

Of the victims who do report offences to police, not all offences result in charges against the alleged offender; in addition, offenders who are charged do not all receive convictions. The New South Wales Bureau of Crime Statistics and Research reported on the progress of sexual offences through the criminal justice system in 2015. The statistics were separated into offences against children aged 0 to 15 and offences against victims over the age of 16. Of the 6571 sexual offence incidents involving a victim between 0 and 15, which were reported to New South Wales police in 2015, 35.1% were dealt with within 180 days of reporting. Criminal proceedings were commenced in 40.7% of these incidents. Out of the offenders who were charged and appeared in court, 62.8% were found guilty of at least one sexual offence. The parallel statistics for victims aged 16 and over are as follows. Of the 4373 sexual offence incidents reported to New South Wales police, 35.1% were dealt with within 180 days of reporting. Criminal proceedings were commenced in 51.9% of these incidents. Out of the offenders who were charged and appeared in court, 53.5% were found guilty of at least

one sexual offence. These statistics highlight the potential for the underrepresentation of the incidence and prevalence of sexual offending.

The most recent Australian statistics available at the time of writing in relation to the number of victims of sexual offences, were from police records between January 1 and December 31 2016. These statistics were reported by the Australian Bureau of Statistics for each Australian state and territory. Although the most recent available information was from different time periods in Australia compared with internationally, on average in Australia, the prevalence of sexual offending is lower than that reported internationally. The Australian statistics are described below.

In New South Wales, there was a total of 113.8 per 100,000 individuals who reported that they were victimised in 2016. This statistic includes both male and female victims of all ages. The rates across other Australian states and territories ranged from 64.6 per 100,000 individuals (Australian Capital Territory) to 154.2 per 100,000 individuals (Northern Territory). Another difficulty with the comparison between rates of reported victimisation and offenders' contact with police is the different time periods during which the data are gained. Whereas the victimisation data were provided within the calendar year, the police action data were gained across the financial year. There were 30.7 per 100,000 individuals in New South Wales who were dealt with by police between 1 July 2015 and 30 June 2016 (i.e., the offender was cautioned or charged during this reference period, while the offence might have been committed earlier), in relation to sexual offending. The range across all states and territories was between 21.8 per 100,000 (Tasmania) and 74.9 per 100,000 (Northern Territory). Putting this statistic in context of offending overall in Australia, there were 1918.1 per 100,000 offenders dealt with by police in New South Wales during this time period.

1.1.3 Impact of Sexual Offending

Despite the difficulties measuring the incidence and prevalence of sexual offending, and its relatively low base rate compared with general offending, it is clear that globally, sexual offending has far-reaching consequences and is of particular concern to the general public, media and for crime policy development. The impact of sexual offending on both child and adult victims can be experienced through short- and long-term psychological, emotional and physical effects. The following information about the short- and long-term effects of sexual abuse was accessed from the Rape, Abuse and Incest National Network, and the Canadian Resource Centre for Victims of Crime, with empirical data to follow. The short-term effects for both children and adults

can include: feelings of shame or guilt; denial; minimisation; poor understanding of boundaries; mistrust of others; isolation; somatic problems; anger or sadness; inability to stop thinking about the assaults; nightmares; self-harm or suicidal ideation; substance abuse; mood disorders; and, Post Traumatic Stress Disorder. The long-term effects can include: ongoing anxiety; poor physical health; sense of helplessness; persistent fear; depression; sleep disturbance; flashbacks; withdrawal/isolation; relationship difficulties; and, paranoia.

The majority of children who are sexually abused will be moderately to severely symptomatic at some point in their life (Hornor, 2010). Short-term effects of child sexual abuse have included the development of inappropriate sexual behaviour (Beitchman, Zucker, Hood, DaCosta, & Akman, 1991). More recent research has also indicated that sexualised behaviour was persistent two years subsequent to detection in 48.4% of victims, such that childhood sexual abuse increased the relative risk of persistent sexualised behaviour problems by 3.29 times (Ensink et al., 2018). Further, sexually abused children have demonstrated higher levels of peer victimisation (Hébert, Langevin, & Daigneault, 2016); however, causality could not be inferred from the data as it was cross-sectional rather than longitudinal.

Prior research has been conducted in relation to the consequences of childhood sexual abuse for adults, with results indicating that it is related to negative outcomes in various domains of adult functioning, such as education, employment, offending, relationships, and parenting (de Jong, Alink, Bijleveld, Finkenauer, & Hendriks, 2015). Research has suggested that victims of childhood sexual abuse were more likely to experience cumulative adverse psychiatric and behavioural problems compared with individuals who were not victims of abuse (Papalia, Luebbers, Ogloff, Cutajar, & Mullen, 2017). Various studies have been conducted using a sample of 2759 victims of childhood sexual abuse in Victoria, Australia (Cutajar et al., 2010a, 2010b; Ogloff, Cutajar, Mann, & Mullen, 2012; Papalia et al., 2017). Papalia et al. (2017) demonstrated that more than half (52%) of victims experienced longer-term problems in one or more of the following domains: mental health problems; offending; further victimisation; or, fatal self-harm. The rate of psychotic disorders was significantly higher among victims of childhood sexual abuse compared with their peers (Cutajar et al., 2010a). Further, a lifetime record of contact with public mental health services was found in 23.3% of victims compared with 7.7% of control participants, with childhood sexual abuse estimated to account for approximately 7.8% of mental health contact (Cutajar et al.,

2010b). Ogloff et al. (2012) reported that 23.6% of the childhood sexual abuse victims had a recorded offence as an adult, compared with 5.9% of participants in the control group. Although causality cannot be concluded, Papalia et al. (2017) suggested that the results indicated sexual abuse was a particular risk factor for co-occurring adverse experiences if it occurred during adolescence, defined as ages 12 to 16 years.

Not only do the victims suffer as a result of sexual offending; victims' family and friends may also experience similar reactions such as fear, guilt, self-blame and anger. In addition, the community in which the victim resides might be affected (e.g., schools, workplaces, neighbourhoods, campuses, and cultural or religious communities), evidenced through fear, anger or disbelief. There may also be financial costs to the community through medical services, criminal justice expenses, crisis and mental health services fees, and the lost contributions of individuals affected by sexual violence. More generally, "sexual violence endangers critical societal structures through climates of violence and fear" (Impact of Sexual Violence Fact Sheet, National Sexual Violence Resource Center, 2010).

Public perceptions of sexual offenders are generally poor; it has been argued that perceptions are exacerbated by 'sensationalistic journalism'. For instance, Levenson, Brannon, Fortney, and Baker (2007) reported that participants in their study, who were members of the general public, appeared sceptical about the value of treatment in preventing recidivism, were concerned that sexual offence rates were rising, and perceived sexual offenders as more likely to reoffend compared with other types of offenders. Participants held punitive attitudes towards sexual offenders such that they recommended an average of 39 years in custody and 42 years on probation; the mode was 99 years, which the authors noted was the highest number that would fit in the spaces provided in the survey. The public tends to support laws and policies implemented with the intention of protecting the public, despite research suggesting such policies have limited effect on sexual offender recidivism (King & Roberts, 2017). These authors suggested that if public opinion about sexual offending is informed by misconceptions, there should be processes in place to educate the public and policy makers on the reality of sexual offending and victimisation. They reflected that more evidence-based policies might result in more informed decisions regarding risk of sexual victimisation.

1.1.4 Role of Treatment

Due to the widespread adverse consequences of sexual offending, there is strong social and political pressure to help victims and prevent sexual recidivism (Dennis et al., 2012). In England and Wales from 2012 to 2017, the number of reported sexual offences increased by approximately 162% (Office for National Statistics, January 2018). This upwards trend has also been observed in Australia. According to the Bureau of Crime Statistics and Research, in New South Wales, over the five years to June 2017, the incidence of recorded sexual offences increased by approximately 3.45% (New South Wales Recorded Crime Statistics Quarterly Update, June 2017). Although it is challenging to make a direct comparison between jurisdictions due to the differences in measurement described previously, and the lack of clarity surrounding some of the statistics provided internationally, this widespread increase in reported incidents of sexual offending highlights the importance of ongoing research into the implementation of effective intervention and assessment of risk for sexual reoffending.

The role of intervention is to prevent sexual recidivism by known sexual offenders, in order to maintain community safety. There are various forms of interventions, which can be implemented to reduce recidivism. Some of these interventions are rehabilitative (psychological treatment programs), while others represent risk management strategies seeking to decrease or eliminate sexual desire and performance, such as chemical castration (Dennis et al., 2012). There is a widespread expectation within society that sexual offenders should be incapacitated but that they should also be treated; this is despite the common perception that psychological therapy has limited effect on reducing recidivism. These expectations appear somewhat inconsistent with each other. Levenson et al. (2007) suggest a potential explanation; that is, although the public wants sexual offenders to be incarcerated for lengthy periods, there is also an attitude that it will not be harmful to provide treatment even if it is unlikely to benefit the community.

It is commonly thought that risk for sexual offending is dynamic (i.e., changeable) and, therefore, this risk is amenable to change through treatment (Olver & Wong, 2013). Contemporary treatment programs are commonly based on cognitive behavioural therapy, such that they involve the implementation of various strategies, which focus on changing inappropriate cognitions, emotions, and behaviours, and replacing them with skills that maintain prosocial beliefs and behaviours (Kim, Benekos, & Merlo, 2016). These treatment programs have generally demonstrated reductions in sexual recidivism

(Hanson et al., 2002). Further research on treatment effectiveness will be presented in the following subsection.

In order to continue working towards effective intervention in this area, the processes by which sexual offenders make changes in their attitudes and behaviour, and the methods through which these changes can be measured, must similarly be further investigated. Measuring offenders' change through treatment using various sources of information can both improve the evaluative component of treatment efficacy, and provide an early indicator of the level of risk an offender might pose to the public upon release (Friendship, Falshaw, & Beech, 2003). Therefore, change measurement is crucial at the policy-level to determine whether treatment effects change in offenders' attitudes and behaviours. It also facilitates an updated evaluation of offenders' risk of recidivism and the identification of outstanding risk-related needs, which can continue to be targeted.

1.1.5 Treatment Effectiveness

Given the importance of rehabilitation to reduce the impact of sexual offending, it is vital that the efficacy of treatment programs be evaluated. While some researchers have concluded that there is minimal evidence for treatment efficacy (Hoberman, 2016), others have indicated that treatment programs have demonstrated some effect in reducing sexual offenders' risk of recidivism (Olver & Wong, 2013). In an early review of treatment effectiveness, Marshall, Jones, Ward, Johnston, and Barbaree (1991) concluded that sexual offenders can be successfully treated to reduce subsequent recidivism, despite not all programs demonstrating success and not all offenders deriving benefit from treatment. A recent meta-analysis has revealed a positive treatment effect overall, indicating up to 25% lower recidivism rates in treatment compared with control groups (Lösel & Schmucker, 2017).

There are various methodological approaches to the evaluation of treatment effectiveness. The ideal is to have at least two comparison groups, one of which receives the treatment to be measured, and the other receives either no treatment or a purportedly less effective form of treatment (Marshall et al., 1991). Alternatives to providing an untreated control group are to compare the treatment group with an estimate of the likely untreated recidivism rate, for instance the recidivism rate for the STATIC-99's normative sample (Woodrow & Bright, 2011), or to compare treatment completers with treatment non-completers such as those who refused or dropped out of treatment (Seager, Jellicoe, & Dhaliwal, 2004).

As Carter and Mann (2016) noted, it is widely accepted that the effectiveness of sexual offender treatment is difficult to evaluate. Various issues have been identified, which complicate the process of determining the characteristics of effective intervention. These issues include: sexual offending is a heterogeneous category containing several forms of sexually abusive behaviour; sexual offenders have heterogeneous characteristics; despite the extensive research on risk factors related to recidivism and structured assessment instruments, knowledge about causal mechanisms remains unclear; treatment approaches are heterogeneous and cover a wide spectrum from psychosocial interventions to organic interventions; the active role offenders play in the change process and thus in the effectiveness of treatment; and, sound treatment evaluation is not possible due to ethical considerations related to untreated control groups, in addition to the relatively low base rate of sexual recidivism and the need for longer follow-up periods compared with other fields within correctional intervention (Carter & Mann, 2016; Schmucker & Lösel, 2015).

Carter and Mann (2016) indicated that the diversity of sexual offenders is evident through factors such as: the type and detail of the sexual offence committed; the individual's aetiology; the motivation for offending; co-morbidity issues; and, the individual's risk level. Therefore, the authors suggested that in order to effectively evaluate treatment programs, it might be necessary to consider multicomponent programs with flexible delivery schedules, which account for the individual differences between treatment participants. However, it was noted that this type of program is difficult to evaluate. It is more difficult to draw conclusions about the reasons leading to the program's effectiveness or lack of effectiveness.

The issues described above have led to relatively infrequent controlled evaluations of sexual offender treatment programs. As a result, there has been controversy surrounding the effectiveness of such programs, in particular with respect to methodological issues in research (see Schmucker & Lösel, 2015). Hoberman (2016) indicated that if only quasi-experimental research studies are considered, there is evidence for a relatively small decrease in recidivism for low- to moderate-risk sexual offenders; whereas if only methodologically rigorous research studies (such as randomised controlled trials) are considered, there is no definitive evidence for treatment leading to a "substantive" reduction in sexual recidivism.

Quinsey, Harris, Rice, and Lalumière (1993) noted that when reviewing outcome studies, decisions are required in relation to which studies should be included, and

which should be excluded due to methodological issues. They provided a critique of Marshall et al.'s (1991) research and noted that within this research, important methodological weaknesses were identified in almost every study reviewed. However, there are significant challenges associated with conducting methodologically sound research in this area. Therefore, Quinsey et al. (1993) suggested there are consequences for using such stringent rules to exclude studies from reviews, such as inability to provide satisfactory answers to questions regarding treatment efficacy. Due to these difficulties, Stinson, Becker, and McVay (2017) noted that overall, little is currently understood about specific treatment processes, which may facilitate sexual offenders' decreased risk.

In response to the methodological limitations of research into treatment effectiveness, the Collaborative Outcome Data Committee (2007) developed a set of standards to guide researchers in the evaluation process, to facilitate increased confidence in research results. This committee was established in Canada, with the goal of advancing outcome research on sexual offenders. Members of the committee were selected based on their expertise in sexual offender research evaluation and their ability to provide different perspectives through their divergent opinions regarding treatment effectiveness. Together, the committee outlined seven primary areas to be considered when evaluating a treatment program: administrative control of the independent variable; experimenter expectations; sample size; attrition; equivalence of groups; outcome variables; and, correct comparison conducted. However, the majority of available research conducted subsequent to the development of these standards has not explicitly reported the use of these guidelines.

Grady, Edwards Jr, and Pettus-Davis (2017) reported implementing these guidelines in a recent study investigating the longitudinal outcomes of sexual offenders who participated in the custody-based Sex Offender Accountability and Rehabilitation Program in North Carolina. Within this study, there were 256 offenders in the treatment group and 256 in a matched sample. The findings reflected that participation in the treatment program did not significantly decrease the rates of sexual or non-sexually violent recidivism. However, there was a significant reduction in the recidivism rate for non-violent offences by 34% for these participants. Of note was that prior to treatment, the participants in this study were low risk of sexual recidivism, which meant any additional impact the program might have had on decreasing recidivism would likely be difficult to detect.

An evaluation was conducted on the Custody-Based Intensive Treatment (CUBIT) program in Sydney, Australia, in which Study 3 of the current research was based. The sample within this evaluation consisted of 386 sexual offenders, who were found suitable for CUBIT between 2000 and 2010. There was a five-year follow-up period post-release, such that all court appearances up to June 30 2015 were used. Results revealed that treated sexual offenders were less likely to commit further general offences, with a recidivism rate of 40.9% compared with 56.5% of untreated sexual offenders; however, there was no significant difference in sexual (11.7% of treated vs 12.1% of untreated) or non-sexually violent (26.5% of treated vs 34.0% of untreated) offending (Halstead, 2016). One methodological limitation within this study was that during the period from which the sample was gained, there was a significant shift in the program implementation. This shift was in both the format by which the program was executed (e.g., the number of days each week the group sessions were run and the content of the sessions), and the clinicians' overarching approach (e.g., it was confrontational during the earlier years, but was subsequently transformed to a more collaborative approach). The impact of these changes was not considered in the research design. It is possible that these unexplored issues contributed to the non-significant findings.

Similarly, a sexual offender treatment program was recently evaluated in the United Kingdom. Participants included 2562 treated sexual offenders and 13219 untreated sexual offenders. After an average 8.2-year follow-up period post-release (between 2002 and 2012), the 'sexual reoffending' rate (all sexual offences except breaches) was 2.0% higher for the treated group than for the matched comparison group, and the 'child image reoffending' rate was 1.6% higher for the treated group, using binary reoffending (at least one reoffence) as the outcome measure (Mews, Di Bella, & Purver, 2017). In relation to the frequency of reoffending, the treated group had 0.15 more sexual reoffences per offender compared with the comparison group. Overall, the treated group had 0.27 fewer non-sexual and non-sexual non-violent reoffences per offender than the matched comparison group. This result was consistent with Grady et al.'s (2017) and Halstead's (2016) studies, which also demonstrated a decrease in general reoffending.

Consistent with research suggesting that treatment programs implemented in the community and in forensic hospitals, delivered in a partially individualised format for medium- to high-risk offenders, are the most promising (Lösel & Schmucker, 2017), a

recent study within an inpatient psychiatric hospital demonstrated improvements in aggression, sexual offending, and indicators of treatment compliance and change (Stinson et al., 2017). Despite research suggesting that interventions implemented in a group format have the weakest effects (Lösel & Schmucker, 2005), this format tends to be the most common form of treatment and is often recommended by researchers and clinicians. Ware, Mann, and Wakeling (2009) noted that the group process will likely benefit sexual offenders due to their common deficits (e.g., in relation to social skills), and it is likely to have financial advantages.

It has been suggested that reviews of treatment effectiveness studies in the form of quantitative meta-analyses can provide greater understanding of the cumulative effects of treatment outcome studies (Yates & Kingston, 2016). A meta-analysis based on 23 recidivism outcome studies demonstrated a treatment effect (Hanson, Bourgon, Helmus, & Hodgson, 2009). Findings reflected that the sexual recidivism rate for treated sexual offenders was lower than for comparison groups (10.9% vs 19.2%). Further, general recidivism was also lower for treated offenders (31.8%) than for comparison groups (48.3%).

The most recent Cochrane review, focusing on randomised controlled trials, in relation to sexual offender interventions, included 10 studies representing 944 sexual offenders (Dennis et al., 2012). The authors concluded that the evidence for the effectiveness of sexual offender treatment is weak. They advocated for additional randomised controlled trials, with an emphasis on methodologies that minimise risk of bias.

More recently, in Schmucker and Lösel's (2015) meta-analysis, which included 29 studies, ranging from 16 to 2557 participants, the results revealed that treatment can effectively reduce sexual recidivism. This review was not focused exclusively on randomised controlled trials. The mean follow-up period was 5.9 years. The mean rate of sexual recidivism was 10.1% for treated offenders and 13.7% for untreated offenders. The mean rate of general recidivism was 32.6% for treated offenders and 41.2% for untreated offenders.

Despite these reductions in recidivism rates for treated sexual offenders, Schmucker and Lösel (2015) indicated that such results cannot be generalised due to the heterogeneity in the results of the primary studies; therefore, there is no definitive trend within the most methodologically sound studies available. While some of the primary studies demonstrated treatment efficacy, others demonstrated minimal to no treatment

effects. For instance, Marques, Wiederanders, Day, Nelson, and van Ommeren (2005) conducted a randomised controlled trial for incarcerated sexual offenders in California. The findings reflected no differences between the treated and untreated groups in their rates of sexual or violent recidivism. Notably, the treatment program evaluated in this study targeted few empirically-based criminogenic needs, and it was based on a relapse prevention model. Although widely used in the past, more recently the relapse prevention model has been reported to have limited efficacy for reasons such as: it does not represent the dynamics of sexual offending; it is focused on avoidance rather than approach goals; it uses constructs that are not applicable to many sexual offenders; and, it has limited empirical support (Yates & Kingston, 2016).

Kim et al. (2016) reviewed a series of 11 meta-analytic studies focused on sexual offender treatment effectiveness. The most recent meta-analysis included in this research was from 2009. This research built on a prior review of six meta-analyses (Craig, Browne, & Stringer, 2003b). These meta-analyses included an array of intervention techniques, including psychological treatment such as cognitive behavioural therapy, and medical treatment such as surgical castration and hormonal medication. Kim et al. indicated that their review, which included more recent meta-analyses than did Craig et al.'s (2003b) review, demonstrated a larger and more robust sexual offender treatment effect in reducing recidivism. The additional meta-analyses suggested a 22% reduction in recidivism subsequent to treatment, compared with the 10% reduction in recidivism ascertained through Craig et al.'s (2003b) review.

In summary, treatment programs have demonstrated varying levels of efficacy. Some research has provided greater support for treatment decreasing offenders' risk of general reoffending, than risk of sexual reoffending. Overall, literature on treatment efficacy reflects a complex array of theoretical underpinnings (e.g., relapse prevention or strengths based), program content (e.g., targeting criminogenic or non-criminogenic needs), client characteristics (e.g., risk level, personality disordered), and research methodologies (e.g., randomised controlled trials or quasi-experimental). Due to these complexities, the evaluation of the effectiveness of treatment remains a challenge. However, the most effective treatment programs have been identified as those implemented in the community and in forensic hospitals, as opposed to the custodial setting. Although group treatment is the most cost effective, programs with at least partially individualised components are more beneficial. Further, programs targeted towards moderate- to high-risk offenders are more effective than those targeting low-

risk offenders. In order for treatment programs to evidence a reduction in recidivism, they must engage the most robust theoretical underpinnings (for instance, primarily have an approach- rather than avoidance-focus) and target empirically-based criminogenic needs. When using the most stringent methodological characteristics to measure treatment effectiveness, there is weaker evidence for sexual offender treatment. However, overall, there is evidence for the efficacy of sexual offender treatment programs. It must be noted that individuals' needs vary and while some forms of intervention are more effective than others, even effective treatment tends not to facilitate change in all offenders.

1.1.6 Theories of Offender Rehabilitation

In order to implement effective treatment programs, there is a need to understand the processes that lead individuals to sexually offend. The Psychology of Criminal Conduct (PCC) is a theory taking a holistic approach, with the objective to understand and explain variation in the criminal behaviour of individuals (Andrews & Bonta, 2016). The PCC seeks to understand the causes of individuals' criminal behaviour, with an empirical focus on the variation. That is, people differ in the number, type, and variety of antisocial incidents in which they engage, and they also differ in the timing and circumstances related to their harmful behaviour. The understanding of criminal behaviour sought by the PCC is empirical, theoretical, and practical; it seeks explanations of criminal conduct, which are consistent with the findings of systematic observation, rationally organised, and useful to people with practical interests in criminal behaviour. It acknowledges the contributions of social context, biology, and psychopathology. The understanding gained from the PCC can be used in the prediction of criminal behaviour. According to Andrews and Bonta (2016), the prediction of criminal behaviour may be one of the key activities of the criminal justice system, thereby leading to community safety, prevention, treatment and justice. They noted that predicting who might reoffend can guide police officers, judges, custodial staff, and parole boards in their decision-making.

The underlying psychology behind the PCC is a combination of social learning, cognitive behavioural, and social cognition theories. The PCC attributes the cause of antisocial behaviours to a combination of "personal control through antisocial attitudes, interpersonal control, through social support for crime provided by antisocial associates, non-mediated control established by a history of reinforcement of criminal behaviour, and/or personal predispositions" (Andrews & Bonta, 2010, p. 10). It provides a basis for

assessing risk of recidivism and for rehabilitation planning (Ogloff & Davis, 2004). The sources of variation in criminal recidivism are found through analyses of the effects of: pre-service characteristics of offenders; characteristics of custodial staff; specifics of the content and process of services planned and delivered; and, intermediate changes in the person and circumstances of individual offenders (Andrews et al., 1990). There were eight major risk/need factors identified within this theory, which were suggested as targets for reduced recidivism: history of antisocial behaviour; antisocial personality pattern; antisocial cognition; antisocial associates; family/marital; school/work; leisure/recreation; and, substance abuse (Andrews & Bonta, 2016).

The PCC provides focus for the assessment and treatment of offenders. The Risk-Need-Responsivity (RNR) principles (Andrews et al., 1990) are based on the PCC. The RNR principles relate to the general principles of effective offender rehabilitation and will now be described.

1.1.6.1 Risk-Need-Responsivity (RNR) principles. The risk principle refers to matching the level of recidivism risk to the intensity of the intervention. It is necessary to assess and predict each individual's level of risk for recidivism, such that the intervention intensity can be matched to this level. In order to predict criminality, risk factors that are empirically associated with offending must be identified (Andrews & Bonta, 2010). In custody, the risk principle is concerned with two types of risk: the risk that an offender poses to engage in institutional misconduct; and the risk of recidivism in the community, such that resources can be allocated appropriately in custody (Makarios & Latessa, 2013).

Criminogenic needs are the specific dynamic risk factors that relate directly to risk for recidivism (Andrews & Bonta, 2016). The RNR's need principle refers to the process of identifying specific criminogenic needs that are related to each individual's offending and subsequently targeting these needs within the intervention. In targeting treatment to focus on these criminogenic needs, the goal is to reduce offenders' risk of recidivism.

Although it is important to focus on criminogenic needs during intervention processes, there are many preconditions that must be satisfied prior to working on the criminogenic needs of an offender. For instance, treatment providers must build on strengths and remove barriers to effective participation in intervention. Addressing non-criminogenic needs initially might facilitate offender motivation and create a more effective therapeutic environment for the offender (Andrews, Bonta, & Wormith 2011).

This is also related to the responsivity principle, which relates to matching the style and mode of the intervention to the offender's learning style and abilities. Andrews and Bonta (2016) differentiated between general and specific responsivity. General responsivity relates to: the mode of treatment from which offenders tend to benefit most (i.e., those taking a cognitive behavioural approach); the importance of the therapeutic relationship between the therapist and offender; and, the use of prosocial modelling, reinforcement, and other appropriate methods to facilitate change. Specific responsivity relates to the recognition of the individual needs of the treatment participant, such as their intellectual ability, cultural background and personal strengths.

Another key component of the RNR model is to build on strengths and reward non-criminal alternatives to the risk factors that favour offending behaviour (Andrews et al., 2011). Not only is it important to focus on reducing the presence of risk factors, but these risk factors must be replaced with protective factors. It is important to identify an individual's strengths in order to further build on them and advocate a prosocial orientation (Ogloff & Davis, 2004).

1.1.7 Risk Factors

The effectiveness of the RNR principles for intervention relies on the identification of risk factors and developing measures with sufficient predictive validity (Ward & Stewart, 2003). Risk factors can be viewed as psychological and social processes (i.e., those associated with goals, plans, strategies, and action implementation), which impair normal functioning and disrupt an individual's internal and external relationships to the social, cultural, and physical environments (Heffernan & Ward, 2015). In determining risk of recidivism, there are two categories of risk factors: static and dynamic (Andrews et al., 1990). Static factors are unchangeable (e.g., historical factors such as offence or victim type). Due to their unchanging nature, they are not suitable as targets for intervention (Mann, Hanson, et al., 2010). Dynamic factors are those risk factors that can be changed. When changed, these factors are associated with subsequent changes in the likelihood of offending behaviour (Andrews et al., 1990). The concept of dynamic risk factors remains the centre of ongoing debate.

Douglas and Skeem (2005) discussed the concept of risk state, as an offender's propensity to become involved in offence-related behaviour at a particular time, based on changes in biological, psychological, and social variables in his or her life. They noted that static risk factors describe an individual's risk status, whereas a combination of static and dynamic factors describes an individual's risk state. In order for dynamic

risk factors to be relevant within intervention, they must be causally related to change in risk state (Douglas & Skeem, 2005). Labelled ‘causal dynamic risk factors’, Douglas and Skeem indicated that these variables must be shown to: precede and increase the likelihood of violence (i.e., be a risk factor); change spontaneously or through intervention (i.e., be a dynamic factor); and, predict changes in the likelihood of violence when altered (i.e., be a causal dynamic risk factor). It is crucial to identify causal dynamic risk factors in order to develop targeted interventions to reduce risk.

Ward and Beech (2006) proposed that static risk factors gain their predictive value because they reflect the past action of enduring psychological risk factors. Ward (2016) suggested that dynamic risk factors are similar to static risk factors in that they are best conceptualised as proxies or markers for the causes of offending; it was argued that due to their apparent conceptual incoherence, they may not plausibly be considered causal factors. Mann, Hanson, et al. (2010) endorsed and expanded upon Ward and Beech’s idea through their suggestion that the static/dynamic distinction should be abandoned. They suggested that risk factors should instead be conceptualised in terms of psychologically meaningful risk factors, which may manifest, and be measured, in a variety of ways.

Dynamic risk factors differ in the speed with which they change (Douglas & Skeem, 2005). These factors have been further categorised into stable and acute variables. Stable dynamic risk factors are more constant over time, while remaining changeable. In contrast to stable dynamic factors, acute dynamic risk factors are not primarily related to long-term risk of recidivism. Rather, they are used to predict when an offender might be likely to reoffend (Craissati & Beech, 2003). Therefore, they are useful in identifying the context in which sexual offenders are most likely to reoffend (Hanson & Harris, 2001). These risk factors are identified such that they can be addressed in treatment.

In contrast, similar to their view of static and dynamic risk factors, Mann, Hanson, et al. (2010) reported that to date, there has been limited empirical support for the distinction between stable and acute risk factors. For instance, prior research demonstrated that acute risk factors did not provide information about imminent reoffending; rather, they appeared to represent ongoing, current expressions of longer-term problems (Hanson, Harris, Scott, & Helmus, 2007). These findings suggested that the average of successive assessments of ‘acute’ risk factors was more predictive than the most recent assessment.

Heffernan and Ward (2015) suggested that stable and acute risk factors can also be reconceptualised as traits and states, respectively. Beech and Ward (2004) discussed these concepts such that some items identified as acute risk factors would be better conceptualised as triggering events or contextual risk factors, which interact with vulnerability (trait) factors to generate states likely to produce sexual offending. They asserted that learning events may result in the acquisition of psychological vulnerabilities, which in risk assessment are operationalised as historical factors and stable dynamic risk factors. Contextual factors, such as relationship, conflict, or access to a victim, then interact with the vulnerabilities to produce acute mental states, such as deviant sexual desires or intense anger (i.e., acute dynamic risk factors).

Building further on this concept, risk factors have been conceptualised as propensities and manifestations (Mann, Thornton, Wakama, Dyson, & Atkinson, 2010). Mann, Hanson, et al. (2010) conceptualised dynamic risk factors as propensities to offend, which are triggered in certain contexts, suggesting that they are vulnerabilities, which may or may not cause a sexual offence. Within this theory, a propensity is an enduring vulnerability, which may or may not manifest during any particular time period. Similar to the concept of a trait, propensities lead to predictable expressions of certain thoughts, feelings and behaviours. Active underlying propensities are evidenced through manifestations, which are observable behaviours.

The term propensity emphasises that the problematic behaviour of interest arises through interactions with the environment. For instance, aggressive offenders are not aggressive all the time; rather, they become aggressive in response to certain interpretations of their environment (Mann, Hanson, et al., 2010). Whether or not propensities are manifested is assumed to depend on an interaction between the ongoing environment and the strength of the propensity (Thornton, 2016). Therefore, it could be conceptualised such that a propensity is the risk factor and the ‘problematic behaviour of interest’ is the behavioural expression, or manifestation, of the risk factor.

As Thornton (2016) suggested, the absence of expression of the propensity/vulnerability in a particular setting might reflect that the relevant environmental triggers were not present, or that the environment produced little opportunity for its external expression. However, when an individual returns to a setting that provides the relevant triggers and allows related behaviour to be displayed, the vulnerability may re-activate and drive behaviour. The behavioural manifestations would subsequently be observed.

In the language of stable and acute dynamic risk factors, the underlying propensity can be conceptualised as a stable dynamic risk factor, while the manifestations are the acute dynamic risk factors relevant for a particular offender and can be targeted in treatment. This idea is consistent with Beech and Ward's (2004) suggestion that stable dynamic factors should be understood as traits while acute risk factors would be the state expression of these traits released by triggering/contextual factors.

Mann and colleagues (2010) proposed that various criteria must be met in order for a propensity to be considered a risk factor: there must be robust empirical evidence that it predicts recidivism; there should be a theoretically plausible justification that it could be a cause of sexual offending; it should be amenable to change; and, a change in this propensity would reduce the likelihood of recidivism. Although the authors suggested that the most useful propensities are those that are amenable to change, in contrast to Douglas and Skeem's (2005) suggestion that dynamic risk factors must be changeable, Mann and colleagues argued that it is not necessary for dynamic risk factors to be completely changeable in order to be psychologically meaningful. That is, even if the risk factor cannot be changed, it is possible that it can be 'neutralised' through external management or development of prosocial alternative behaviours to assist with management. For example, there is ongoing debate regarding whether deviant sexual interests can be changed through intervention; however, there is widespread agreement that it is a risk-relevant propensity. This propensity can be managed through the development of adaptive skills.

Within the RNR model, an intervention's focus on the need principle would motivate the largest changes for sexual offenders (Hanson et al., 2009). For a factor to be considered a criminogenic need, or a dynamic risk factor, an empirical association with recidivism is required. However, it has been argued that rather than dynamic risk factors representing the causal factors for offending, these are in fact the observable behaviours, or the 'front end' for complex constructs containing the causal elements (Ward & Beech, 2014). Furthering this argument would suggest that dynamic risk factors and propensities are the same constructs as the manifestations; that is, the dynamic risk factor is the observable behaviour related to complex composite constructs.

In relation to the components of dynamic risk factors, Klepfisz, Daffern, and Day (2016) suggested that one possible method by which to bridge the gap in evidence

between the influence of single causal dynamic risk factors and aggregated scores is to conceptualise dynamic risk factors in terms of broad levels, or domains, which subsume a number of more specific risk factors. This process in part relates to Thornton's (2002) discussion surrounding domains for sexual offending. Thornton (2013) asserted that a limitation of Mann, Hanson, et al.'s (2010) study in which they identified psychologically meaningful risk factors, was that they considered the risk factors individually without integrating them into broader categories, or domains, which might reflect more general patterns in the data.

This idea of categorising individual risk factors is consistent with Ward and Beech's (2015) argument that dynamic risk factors are composite constructs composed of multiple variables. They outlined a challenge in conceptualising dynamic risk factors, which is a function of the composite constructs; that is, these constructs have not been adequately explored and explained. Ward and Beech noted that a result of the complex composition of dynamic risk factors is that they therefore include causal processes, aetiological factors, and symptoms or clinical phenomena. Consequently, it is difficult to know exactly what is being referred to in explanatory, assessment, and treatment contexts (Ward, 2016). Similarly, Heffernan and Ward (2015) suggested that the contemporary view of risk does not include motivational and normative features. They indicated that the relationship between dynamic risk factors and sexual offending is primarily one of association and does not imply direct causation.

Ward (2016) noted that dynamic risk factors vary in their level of abstraction. For instance, at their most abstract, they consist of general domain names such as 'relational style' and at a more concrete level they are unpacked into factors such as 'emotional congruence with children'. Similarly, as argued by Ward and Beech (2014), risk factors identified in the literature contain both causal and descriptive elements, which must be teased apart in order to avoid combining distinct constructs. For instance, despite both being included in the dynamic risk category of interpersonal problems, lacking emotionally close adult relationships is an observable problem, while emotionally identifying with children may be more appropriately viewed as one of its possible causes.

Ward (2016) indicated that dynamic risk factors cannot at this stage be considered explanatory concepts but rather, they are predictive constructs; therefore, he argued that they should not be used in intervention planning. He concluded that dynamic risk factors are best understood as 'general problem indicators' rather than providing

‘detailed blueprints for change’ (p. 15). That is, they should be conceptualised as clusters of clinical features or ‘symptoms’ generated by underlying causal mechanisms, rather than being causes themselves (Ward & Beech, 2014).

Dynamic risk factors are composite constructs in at least three distinct senses (Heffernan & Ward, 2015). These constructs have different levels, with an overall set of domains, the domains themselves, and the particular dynamic risk factors contained within each of the domain categories. A further problem related to the composite nature of dynamic risk factors is that each domain is typically separated into additional features, some of which causally exclude each other. The overall domains often include qualitatively different variables, which may refer to distinct causal processes and their associated problems. Finally, the description of dynamic risk factors is vague and appears to include both dispositional and state aspects. For example, the stable dynamic factor of general self-regulation includes negative emotionality (a mental state) and poor problem solving (an enduring psychological feature). Risk factors are largely a combination of both long-term vulnerabilities (as evidenced by past offences) and their manifestations in certain contexts. Therefore, risk categories such as intimacy deficits may be more accurately viewed as composite constructs, which contain causal, descriptive, developmental, and contextual elements. Thus, taking into account the heterogeneous nature of dynamic risk factors, Heffernan and Ward (2015) proposed that these composite constructs require further analysis to disentangle their various inter-related components.

Klepysz et al. (2016) suggested that some items included in established risk assessments such as the Violence Risk Scale (VRS; Wong & Gordon, 2000) are not true dynamic risk factors, for instance, ‘violence during institutionalisation’ and ‘recent problems with treatment and/or supervision’. Rather, violence during institutionalisation might be more effectively defined as a manifestation of other dynamic risk factors. This variable could be considered a proxy variable or a risk marker for pervasive criminal attitudes or involvement with an antisocial peer group, which underlie and contribute to the offending behaviour. That is, there is a need to differentiate between ‘proxy’ (risk marker) and causal dynamic variables as defined by Douglas and Skeem (2005). To do so, Klepysz et al. suggested considering the treatment that would be required to reduce these areas of risk. Using ‘violence during institutionalisation’ as an example, treatment would more likely target the variables that contribute to this behaviour, rather than treating ‘violence during institutionalisation’ in itself.

Further, clinicians have a responsibility to determine whether particular risk factors are functionally related to the propensity (e.g., violence) for the individual being assessed, and consequently differentiate between the variables that have a causal relationship with the propensity, and also differentiate these variables from other risk markers (Klepfisz et al., 2016). Similarly, van den Berg et al. (2017) suggested that uncovering causal pathways between risk factors and sexual offending behaviour could provide greater insight into the constructs underlying the risk factors assessed by static and dynamic instruments, which may lead to the inclusion of more psychologically meaningful risk factors in these assessments. It may be valuable to further investigate how dynamic risk factors interrelate and which are the central factors in the reduction of recidivism risk.

In summary, the conceptualisation of dynamic risk factors is complex and there is disagreement within the field. The traditional view of dynamic risk factors is consistent with the RNR principles and the PCC, and within assessment, dynamic risk factors are delineated by established risk assessment instruments based on these definitional properties. In order for a factor to be considered a criminogenic need (dynamic risk factor), an empirical association with recidivism is required. As previously described, researchers have presented various alternative views. For instance, dynamic factors have been identified as proxies, or markers for the causes of offending, which cannot be considered causal factors due to their conceptual incoherence. The static and dynamic factor distinction has been called into question, with an alternative concept suggested through psychologically meaningful risk factors, which consist of propensities, vulnerabilities and manifestations. Risk factors have been labelled as traits and states, with the idea that rather than items identified as acute risk factors, these factors should be conceptualised as triggering events or contextual risk factors, which interact with trait factors (vulnerabilities) to generate states likely to produce sexual offending. In these definitions, static and stable dynamic factors are referred to as vulnerabilities, which are acquired as a result of learning events; contextual factors subsequently interact with these vulnerabilities. Researchers have asserted the need for differentiation between risk markers and causal dynamic factors such as those defined by Douglas and Skeem (2005) and which include the traditionally defined dynamic risk factors. Further, rather than dynamic risk factors representing discrete variables, they have been described as composite constructs composed of multiple variables including causal processes, aetiological factors, and symptoms or clinical phenomena.

Within the current research, the traditional view of dynamic risk factors has been maintained. A key component of this conceptualisation of dynamic risk factors is that they are able to change. According to this view, dynamic risk factors are important because they should be a focus of intervention and assessment of sexual recidivism risk. Established risk factors form the focus of contemporary assessments that are used to assess offenders' change in risk level. All three studies within the current research are based on established dynamic risk factors (those that exist in widely used risk assessment measures, as described below) and their manifestation within the custodial setting.

1.1.7.1 Risk factors for sexual offending. A substantial amount of research has been conducted to determine the risk factors that are related to sexual reoffending. Although there has been much debate, there is general consensus regarding several key risk factors. Core dynamic variable domains have been evidenced thus far through research, which contribute to assessing risk of sexual recidivism. These domains include intimacy deficits, which might manifest in several ways; for example, child sexual offenders might fear rejection, which leads to an avoidance of adult intimacy, and adult sexual offenders might lack empathy for women, which leads to multiple casual sexual encounters. Social influences constitute another predictor of risk, such that prosocial supports are indicative of a lower risk of reoffending. Further examples of criminogenic needs, which are empirically linked to sexual recidivism, are sexual deviancy, sexual preoccupation, low self-control, and grievance thinking (Hanson et al., 2009). Prior research revealed that sexual recidivism was best predicted by measures of sexual deviancy (Hanson & Bussière, 1998).

Mann, Hanson, et al. (2010) identified a list of psychologically meaningful risk factors, which were plausible causes of offending. The empirically supported risk factors they identified are: sexual preoccupation; any deviant sexual interest; offence-supportive attitudes; emotional congruence with children; lack of emotionally intimate relationships with adults; lifestyle impulsivity; general self-regulation problems; poor cognitive problem solving; resistance to rules and supervision; grievance/hostility; and, negative social influences. However, they concluded that none of the psychological risk factors identified to date had a strong relationship to sexual offending. They suggested some implications for this finding. For instance, clinicians should refrain from focusing on the presence of any single risk factor, which may appear to strongly manifest, as this might unduly bias assessment. Further, if there are a large number of risk factors, which

each make only a small contribution to prediction, it is likely that only relatively comprehensive assessment of a range of these risk factors will make it possible for risk assessment to have useful predictive power, with this assessment likely to benefit from mechanical integration of the risk factors rather than reliance on human judgement alone (Mann, Hanson, et al., 2010). It is also important to consider each relevant risk factor for the individual offender concerned, to form a judgement as a whole, rather than focusing on individual risk factors, which might contribute only a small amount to the prediction. The process of risk assessment will now be explored.

1.1.8 Risk Assessment

Risk assessment forms a crucial part of offender management and intervention (Andrews, Bonta, & Wormith, 2006) and the assessment of recidivism risk has been widely researched. The effectiveness of policies and interventions targeted at the reduction of recidivism depends on clinicians' ability to predict who is most at risk of reoffending (Berg et al., 2017). Based on the RNR's risk principle, in which the level of intervention an offender receives is matched to the level of risk, it is important for clinicians to implement an assessment process to determine the appropriate form of intervention (Andrews & Bonta, 2010). Risk assessment focused on dynamic risk factors also allows clinicians to determine whether changes were made subsequent to intervention. Risk assessment tools are continually undergoing improvements in order to assist in the protection of the public from harm by offenders released from periods of imprisonment (Jones, 2010).

1.1.8.1 Types of risk assessment. There have been several generations of risk assessment; the most recent form is the fourth generation (Andrews et al., 2006). The first generation consisted of unstructured professional judgements rating the likelihood with which an individual would engage in offending behaviour. In this procedure, neither the risk factors nor the method of forming the overall evaluation were specified in advance. The development of modern types of risk assessments was motivated by evidence that unguided clinical judgements generally offered poor predictive validity compared with more structured methods (Beggs & Grace, 2010).

Second generation assessments were empirically based risk instruments; however, they were not theory-based and primarily consisted of static items (actuarial instruments). These approaches tended to focus the assessment on a limited number of factors and ignore potentially crucial case-specific, idiosyncratic factors (Doyle & Dolan, 2002). For instance, the STATIC-99R is a widely used actuarial tool (Hanson &

Thornton, 2000). Due to the heterogeneity of sexual offenders, questions have been raised regarding the validity of applying one tool to predict reoffending (Barnett, Wakeling, & Howard, 2010). Although actuarial tools differ greatly in their predictive accuracy, ranging from $r = 0.09$ to $r = 0.45$ (Craig, Browne, & Stringer, 2003a), these tools are considered by many researchers to be the best predictors of sexual offender recidivism risk (Casey, 2016).

There are two limitations of exclusive use of static risk assessments (Craissati & Beech, 2003). The first limitation is that a classification of risk, which is associated with a probability of sexual recidivism, does not provide a determination of whether an individual offender is likely to belong to the reoffending category associated with his level of risk. The second limitation is that due to the unchanging nature of static risk factors, it is difficult to develop risk management strategies that target the potential for change and a subsequent reduction in risk of recidivism.

Third generation assessments were also empirically based. They included static risk factors in addition to a wider sampling of dynamic risk items (criminogenic needs) and tended to be theoretically informed. The items are rated but are not summed to generate a total score. The clinician can evaluate patterns in the item ratings and uses professional judgement to generate a summary risk rating. These tools also included the assessment of change over time. An example of a structured risk tool for sexual offenders is the Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003). The term structured professional judgement stems from the first generation approach. It reflects a decision based on a review of specified items but without a validated mechanical system that links scores to decisions. Research on these measures is still sufficiently underdeveloped that important questions remain concerning their conceptual foundations, whether they target the most relevant factors and the extent to which it is possible to associate recidivism rates with specific scores (Mann, Hanson, et al., 2010).

Adjusted actuarial assessments represent a combination of actuarial (second generation) and structured professional judgement (third generation) assessments. They include static and dynamic items derived from theory and research. Items are assigned point ratings and a total score is generated, which fits within a risk category. An example of such an assessment is the Violence Risk Scale: Sexual Offender version (VRS:SO; Wong, Olver, Nicholaichuk, & Gordon, 2003). In a validation study of the VRS:SO, the Dynamic scale was found to be a significant predictor of sexual recidivism after controlling for static measures, which provided further evidence that dynamic

factors can make independent contributions to risk predictions among sexual offenders beyond that predicted by static factors (Beggs & Grace, 2010).

On the contrary, it has been noted that there is a lack of clarity surrounding the predictive power of dynamic variables (see Wakeling, Freemantle, Beech, & Elliott, 2011). Wakeling et al. (2011) found that the static variables in their study were more powerful predictors of sexual and violent recidivism than any of the dynamic variables measured via self-report assessments (i.e., psychometrics). A suggested cause of this result was that dynamic risk domain scores and the static variables used in the study were 'highly related', such that they explain much of the same variance in outcome.

Proponents of assessments that are purely actuarial argue that accurate risk assessment requires the use of statistically based models, which omit clinical judgement. These risk assessments would use an instrument that statistically identifies relevant factors, producing a score, which translates into risk categories. These risk estimates are based on specific, objective information as opposed to clinical opinion. However, the absence of any clinical input in risk assessment can significantly impact the results (Craig et al., 2003a). For instance, critical clinical variables such as sexual deviance are often poorly represented in actuarial instruments. Further, actuarial models tend not to measure dynamic change based on motivation, insight, or intervention. These factors emphasise the value in structured professional judgement models.

The fourth generation of risk assessment includes service and supervision from intake through to case closure and the intention is to facilitate clinical supervision (Andrews et al., 2006). These measures not only include risk-need assessment; they also integrate the assessment with a case management plan (e.g., the assessment of strengths). This integration ensures that criminogenic needs, which are specifically linked with recidivism, are targeted within intervention. That is, these assessments provide structured intervention plans for supervising officers to use with offenders. These assessments facilitate post-closure follow-up, during which outcome can be linked with intake assessments or risk, strengths, need, and responsivity, with reassessments, and with service plans, service delivery, and intermediate outcomes. A key goal of the fourth generation assessments is to strengthen adherence with the principles of effective treatment and to facilitate clinical supervision focused on enhancement of public protection from offenders' recidivism.

van den Berg et al. (2017) conducted a meta-analysis of 52 research studies focused on dynamic risk assessment instruments, published between 2001 and 2015.

The aims of this review were: (1) to examine the predictive properties of dynamic risk assessment instruments designed to assess recidivism risk; and, (2) to examine the incremental validity of these risk assessments above and beyond that of static risk assessment instruments. The findings reflected that dynamic risk assessments, which were designed to measure sexual recidivism, significantly contributed to the prediction of sexual, violent, and any recidivism, with the largest effect sizes for sexual recidivism. The incremental validity of dynamic over static risk assessments was established for all outcome measures; however, the small effect sizes may have suggested that static and dynamic assessments overlap in their predictive value. van den Berg et al. noted that this outcome reinforces the question Ward and Beech (2015) raised regarding whether dynamic risk factors measure correlates of underlying propensities, as static risk factors do, but in alternate ways. That is, dynamic risk factors might measure clinical features associated with the psychological propensities that actually cause recidivism, rather than assessing these propensities themselves. This possibility highlights the need for ongoing research into the conceptualisation of risk factors.

Overall, the STATIC-99 and STATIC-99R are the most widely used sexual offender risk assessment measure. The predictive accuracy of this actuarial tool outperforms other risk assessment measures, including those that target dynamic risk factors (Hanson & Morton-Bourgon, 2009). However, structured professional judgement measures continue to be used in the clinical context. Although many of the available sexual offender risk assessment measures that incorporate static or both static and dynamic risk factors have similar and acceptable predictive accuracy, some structured professional judgement measures go beyond risk prediction to risk reduction through treatment and rehabilitation (Olver & Wong, 2016). Contemporary research continues to focus on these measures. There is no published validation for some measures (e.g., RSVP), while other measures have received some support (e.g., SVR-20, STABLE-2007 and VRS:SO). Olver and Wong (2016) noted that only the VRS:SO presented empirical evidence that satisfies the requirements of causal risk factors. The VRS:SO validation work remains one of the few studies examining the dynamic nature of putatively dynamic risk factors.

1.1.9 Protective Factors

Historically, discussion around the concept of offenders' change processes focused on change in dynamic risk; that is, the focus was on deficits or needs (Serin, Chadwick, et al., 2016), and the decrease in these deficits or management of

criminogenic needs. However, more recently, the process of behavioural change inevitably includes debate surrounding protective factors. Although there is more focus on these factors, Serin, Chadwick, et al. (2016) identified that confusion appears to remain amongst researchers and clinicians in relation to the operational definition of protective factors, and the ways in which they are conceptually distinct from risk factors, in addition to the ways in which they may influence client outcome. The confusion surrounding the definitional properties of protective factors and the comparatively limited research in this area compared with risk factors may be exacerbated by the limited use of strengths within risk assessment settings and correctional decision-making processes.

Researchers continue to present varying perspectives of protective factors. For instance, protective factors have been described as: buffers, which mitigate risk; factors that decrease risk; or, factors that function independently of risk. Some professionals maintain that protective factors are merely the opposite of risk factors; however, others assert that a protective factor can co-exist with risk factors and account for variability in outcome in a high-risk group of offenders (Serin, Chadwick, et al., 2016). Polaschek (2017) suggests that a protective factor predicts a decrease in offence-related behaviour, which can be considered to be directly protective since the relationship is independent of other factors (Lösel & Farrington, 2012). These factors have been referred to as promotive (Loeber, Pardini, Stouthamer-Loeber, & Raine, 2007). Monahan and Skeem (2016) suggested that protective and promotive factors are often confused with each other. They defined the difference between these two factors: promotive factors act in the opposite direction of risk factors, whereas protective factors moderate the impact of risk factors. This definition was further clarified such that promotive factors reduce the probability of recidivism, whereas protective factors reduce the probability of recidivism among individuals who are exposed to risk factors.

Protective factors can be further divided into static and dynamic factors in the same way that risk factors can be divided, such that static factors do not change over time while dynamic factors are changeable and can be influenced through treatment. Further, in line with Mann, Hanson, et al.'s (2010) theory regarding risk factors, de Vries Robbé et al. (2015) distinguished between protective factors that are underlying propensities and the observable manifestations of the propensity. This observable manifestation is the factor that can be directly measured.

Some risk factors are bipolar rather than unipolar, such that there is a

corresponding risk factor in opposition to the protective factor (Polaschek, 2017). Jones, Brown, Robinson, and Frey (2015) applied the label 'strength' as an umbrella term to broadly identify a positive or prosocial part of an individual's life that is "intuitively apt to buffer risk of criminal outcome" (p. 323). They reported that risk and strength scores significantly interacted. This interaction suggested that high strength scores were particularly effective in decreasing recidivism among higher risk offenders. It is possible that this result portrays individual differences in the impact of both behavioural manifestations of risk factors and more adaptive behaviours, such that the change process is affected by overall risk level. The complex relation between risk factors and protective factors might lead to a more complex view of offenders' change processes and the differential impact of risk and protective factors on post-treatment outcomes.

Ward (2017) discussed a perspective based on research by MacDonald (2016) in which, while protective factors reduce the impact of dynamic risk factors because interventions target and replace them, they do not causally influence the risk factors. An example was provided, such that emotional identification with children is replaced by a preference for emotional intimacy with adults, with no interaction between the two. However, this concept does not necessarily exclude the possibility that protective factors can at times be causally related to risk factors.

There might be a variety of protective factors, which are developed either as a consequence of intervention or independently (e.g., through circumstance or maturation). As a result, there might be differing levels and patterns of change observed between individual offenders, depending on the interaction between risk factors and protective factors. This concept is consistent with Lösel and Farrington's (2012) proposal regarding the cumulative effects of both risk and buffering protective factors. They suggested that direct and buffering protective factors belong to the same pool of variables as risk factors, such that they have effects at one of their 'poles', which reflect both an absence of risk and actively increase the likelihood of a desirable outcome. Jones et al. (2015) reported on the findings of Lloyd (2007) from an unpublished thesis. Criminogenic needs, which were linked to antisocial behaviour, had diminished predictive power over an offender's life course, with the extinction of risk factors only partially accounting for ultimate desistance. This outcome provides evidence for the added value of strength factors above and beyond that of the absence of risk factors.

Serin, Chadwick, et al. (2016) proposed that the presence of a strength factor should not necessarily be assumed to have a protective effect in an individual, because it

is the individual's responsibility to actively choose an adaptive response within any given situation. Hence, the presence of behavioural manifestations of protective or strength factors could be used as indicators that the individual is implementing appropriate strategies.

1.1.9.1 Protective factor assessment. To complement risk focused dynamic assessment tools, a dynamic structured professional judgement tool was developed specifically for the assessment of protective factors: Structured Assessment of Protective Factors for Violent Risk (SAPROF; de Vogel, de Ruiter, Bouman, & de Vries Robbé, 2009). The SAPROF was the first adult assessment tool to solely focus on protective factors for violence. It was developed for use with violent as well as sexually violent offenders (de Vries Robbé, de Vogel, Koster, & Bogaerts, 2015).

de Vries Robbé et al. (2015) indicated that most of the protective factors in the SAPROF are potentially changeable and aim to provide opportunities for positive intervention and risk management. They conducted a retrospective validation study using 83 sexual offenders in two forensic psychiatric hospitals, using three assessment measures: SAPROF; HCR-20; and, SVR-20. Both the HCR-20 and the SVR-20 are risk-focused. The results revealed that the combination of the SAPROF and each of these two risk assessment measures had greater predictive validity for sexual and non-sexually violent recidivism than either of the HCR-20 or SVR-20 alone. It was suggested that the results provided further evidence for the increased predictive accuracy of risk assessments that include protective factors.

The Short-Term Assessment of Risk and Treatability (START; Webster, Martin, Brink, Nicholls, & Middleton, 2004) is a risk assessment instrument developed to assess 'strengths' (protective factors) as the opposite of 'vulnerabilities' (risk factors). It measures both ends of each domain simultaneously. Prior research has revealed that the vulnerability and strength total scores improved the prediction of aggression towards others (Desmarais, Nicholls, Wilson, & Brink, 2012). Further, the total strength score added predictive validity over the total vulnerability score in addition to significantly improving the model fit for the prediction of aggression (O'Shea, Picchioni, & Dickers, 2016).

Jones et al. (2015) suggested that quantifying a strengths-based approach might increase the predictive accuracy of risk assessment measures. They tested the Pre-Screen version of the Service Planning Instrument (SPIn), a risk assessment tool incorporating risks, needs, and strengths. Most strength items are scored using Likert-

type scales, which can be summed to produce domain totals and an overall strength score to assist in determining an offender's classification in custody. In their research, Jones et al. found that both the SPIn Pre-Screen's aggregate risk score and aggregate strength score independently predicted recidivism, with the aggregate strength score providing incremental validity in excess of the risk score. In addition, high strength scores had a greater effect on reducing recidivism among high risk offenders compared with low risk offenders.

Therefore, there is some evidence for the use of protective factor assessment as an improvement in accuracy of risk estimates for physical aggression toward others, and the benefit in assessing specific dynamic factors through a protective lens when a factor is considered from both a risk and protection perspective (Klepfisz, Daffern, & Day, 2017). These authors noted that the assessment measures including protective factors comprise a narrow subset of risk and protective factors, most of which were not developed from overarching desistance theories. Therefore, the complex interaction of psychological, biological and situational factors relevant for this process are not considered. They concluded that there remains a lack of empirical evidence to support the identification of specific protective factors; however, comprehensive risk assessments should allow for the evaluation of both risk and protection.

1.1.10 Behavioural Monitoring in Custody

Monitoring dynamic risk factors and protective factors is crucial for the effective implementation of both assessment and treatment. As Mann, Hanson, et al. (2010) discussed, further research is required concerning the measurement of these factors in offenders; the manifestation of risk factors and the point at which they become problematic in an individual remains to be determined. There is an emphasis on observable variables (Craissati & Beech, 2003), such that the manifestation of risk and protective factors can be monitored and assessed.

1.1.10.1 Behavioural consistency. There has been ongoing discussion about the use of behavioural observations in custody to fill gaps in risk assessment, based on the idea that observed behaviour may be a manifestation of an offender's criminogenic needs or dynamic risk factors related to their violence, for instance, antisocial lifestyle or sexual deviancy (Pearson & McDougall, 2017). However, observation of behaviour in custodial settings may not apply to risk assessment in the community (McDougall, Clarke, & Woodward, 1995). Although Harris and Rice (2003) argued that sexual aggression is determined by enduring physiological and genetic traits, which persist

throughout the lifespan, there is a need to reveal how these traits manifest within institutions such that they can be monitored (Daffern, Howells, Stacey, Hogue, & Mooney, 2008). A central problem for assessments in custody is that the context of assessment and intervention is generally significantly different from the context of the behaviour being examined (Jones, 2010). For valid assessments to be conducted, there must be evidence of consistency between behaviours exhibited in custody and those exhibited in the community. Similarly, it is necessary to determine if there is personal consistency between the ways an offender behaves during the commission of an offence and the ways in which the offender behaves in everyday life (Canter, 2000).

According to Dvoskin and Heilbrun (2001), risk-related behaviour cannot be monitored in populations in which institutional behaviour is less relevant to specialised outcomes, such as offenders who commit sexual offences against children. However, due to the persistence of relevant human behaviour patterns, offenders often find themselves in similar situations in custody as in the community and respond in similar ways to these situations (McDougall et al., 1995); for example, an offender who tends to have conflict within his romantic relationships due to a sense of entitlement and poor emotional regulation, may similarly have interpersonal conflict in custody in response to stressors. Various studies have provided support for behavioural consistency in criminal behaviour between custody and the community (Zamble & Porporino, 1990). For instance, a relationship has been found between sexual misconduct in custody and re-arrest for violent offences (Heil, Harrison, English, & Ahlmeyer, 2009).

Clarke, Fisher, and McDougall (1993) found that 60% of offenders' index behaviours identified from police and court documents could be independently identified as similar to their observed offence-related behaviours in custody, compared with 20% when offence behaviour was compared with a random set of behaviours in custody. While some behaviours incarcerated offenders engage in may be similar to their index offence, there might be some behaviours they exhibit, which are not parallels to their index offence but are still related to their criminogenic needs and potential future offending (McDougall, Pearson, Willoughby, & Bowles, 2013). Therefore, there is a wide range of potentially relevant behaviour, which could be observed in custody.

McDougall et al. (2013) conducted a study to determine whether there was evidence that high risk sexual and non-sexually violent offenders' behaviour within different situations in custody would be reflective of both type and frequency of possible behaviours in the community. Staff in the prison were consulted in the

identification of potential behaviours in custody, which could indicate ongoing positive attitudes towards offending. Custodial officers and other staff in contact with a group of high risk offenders were asked to record behaviours “causing concern” or “positive behaviours” related to offending. The negative behaviours were not necessarily linked to an index offence. The behavioural observations were conducted in various contexts within custody, such as education, workshops and visits rooms.

Results indicated that the frequencies of negative behaviours in custody and in the community, were strongly correlated and that reoffence could be predicted by the frequency of negative behaviours in custody (McDougall et al., 2013). The frequency of incidents was more predictive of cross-situational behavioural consistency than analysis of individual instances, such that it was the aggregation of behaviours that indicated consistency. Of note is that low-level behaviours (i.e., persistent patterns of antisocial behaviour that were learned to deal with social situations, such as verbal aggression) were important in predicting future risk. Pearson and McDougall (2017) suggested that observation of behaviour in custody including lower-level coping behaviours might be beneficial because these behaviours are indicative of community behaviour (e.g., insults, threats, bullying, refusals, defiance, and generally disruptive aggressive behaviours). However, despite their evident importance, these behaviours are often not documented in official records in custody or in the community, as their perceived significance may be limited.

Further, there was a significant correlation between positive behaviour in custody and the community, despite staff members in custody generally focusing more on negative rather than positive behaviours. McDougall et al. (2013) suggested that this result might have been due to custodial staff being asked to monitor low-level behaviours, which could facilitate recognition of positive behaviours. It is possible that in situations where few or no negative behaviours have been recorded, the negative behaviours might have been missed or concealed; however, the presence of positive behaviours, which appear incompatible with reoffending, could provide a more balanced view of risk. This concept is related to the assessment of protective factors and the combined measurement of both risk and protective factors, as Klepfisz et al. (2017) discussed. It also relates to the view that improvement should only be concluded if there is an increase in positive behaviours, rather than just a decrease in negative behaviours (Daffern et al., 2007). McDougall et al.’s results reflect the potential predictive accuracy of monitoring both negative and positive behaviours in custody.

Within McDougall et al.'s (2013) study, there was a similarity between identified custody- and community-based behaviours, with 80% of behaviours observed in the community being assessed by three independent researchers as 'similar' or 'very similar' to those observed in the custodial context. Therefore, based on this premise that patterns of behaviour are similar across contexts, there is scope to gain information about relevant behaviour in custody within even specialised offender populations.

1.1.10.2 Behavioural manifestations of risk factors. Recent research has provided greater understanding of the contribution that offence-related behaviour of offenders in custody might offer for risk prediction in individual cases (McDougall et al., 2013). In order to appropriately assess risk and needs, and manage offenders in a custodial setting, it is important to understand the characteristics and behavioural manifestations of relevant risk and protective factors (Simourd & Hoge, 2000). Pearson and McDougall (2017) indicated that monitoring dynamic behaviours can serve to determine an offender's current level of functioning, through monitoring the evolution of behaviour (e.g., over time in treatment). Additionally, progress in risk reduction can be monitored through information obtained from sources such as employment and behaviour in the wing (Dvoskin & Heilbrun, 2001).

The concept of observing behaviour in custody within specialised offender populations can be further understood through Mann, Thornton, et al.'s (2010) theory regarding the manifestations of underlying propensities. The propensities remain present in an offender regardless of the specific environment. The behaviours identified within McDougall et al.'s (2013) study are equivalent to this concept of ongoing propensities' manifestations. Consistent with Gordon and Wong's (2010) suggestion that antisocial behaviours that are observable in the community may manifest in different ways in custodial settings, and Mann, Thornton, et al.'s discussion regarding the cross-context manifestation of underlying propensities, McDougall et al.'s findings indicated that the custodial environment does not necessarily suppress offence-related behaviours, although their expression might change due to situational circumstances. In their study, an example of behavioural differences between contexts was an offender who got into debt in the community and absconded from his hostel; in custody, he similarly got into debt but due to the context, rather than absconding, he sought a wing change to avoid the consequences of his behaviour. Additionally, although in the study there was no similar custody-based behaviour identified, a negative behaviour in the

community was ‘loitering around a school’; this situation is impossible in custody, but equivalent manifestations are possible such as watching children on visits.

It is also possible that as Mann, Thornton, et al. (2010) suggested, environments differ in the strength with which they trigger propensities. Therefore, the strength of the manifestation depends both on the strength of the underlying propensity and the extent to which the environment has triggered the propensity. In a custodial setting, the environment might produce fewer triggers for an underlying propensity as compared with the community. For instance, an offender who has sexual thoughts about children may be more successful inhibiting the thoughts in custody due to fewer triggers. However, the thoughts may be triggered in situations relevant to the environment; for instance, seeing images of children in the newspaper, or on visits when children are present. Subsequent manifestations of this propensity could include the offender collecting images of children from newspapers or magazines. Alternatively, the offender might be observed to watch or attempt to engage with children on visits.

Further, an offender who has an underlying propensity related to problems with intimate relationships, might continue to have difficulties with intimate relationships but the manifestation might be reduced in custody due to his physical distance from his partner. However, this propensity might be triggered in situations such as on visits or during phone calls. Manifestations might include the offender: engaging in verbal aggression towards his partner; becoming passive aggressive (e.g., use of sarcasm or raising issues from the past); or, having a sense of entitlement reflected through comments about how his partner should be treating him. Similar manifestations might be observed within relationships the offender develops in custody, thus indicating ongoing relevance of the propensity.

Within the examples above, the behavioural manifestations might be less likely in custody compared with the community and those that are present might be weaker or less easily observable (for example, those related to sexual deviance). Although the underlying propensity has not necessarily decreased in intensity, the strength of the manifestations decreased as a function of the environment. However, the relevance of the propensity can be determined through ongoing monitoring of the manifestations.

On the other hand, there might be propensities that are weaker in the community, which are strongly triggered in custody, and thus lead to more frequently observed manifestations. For example, in the community, an offender might not hold particularly strong grievance attitudes towards those in authority (Mann, Thornton, et al., 2010).

However, in custody, there is often an attitude amongst offenders that they are in opposition to those in authority. Further, there are often situations in which offenders feel victimised by authority figures, which may exacerbate feelings of hostility within the relationship. As a result, although the propensity in the community for hostility towards authority might be comparatively weaker, it may manifest to a greater extent in custody due to the environment. Manifestations such as verbal or physical aggression might be more likely in this environment. The level at which the propensities continue to be active in custody can be determined through observation and monitoring of the manifestations, provided the potential triggers and behavioural manifestations are sufficiently understood.

When monitoring behaviours exhibited in custody, it is important to discriminate between behaviours that are risk-related and those that are merely a function of the environment. This process is complex but important because contextual factors within the institutional environment might either inhibit or activate certain behaviour such as aggression (Daffern, Ferguson, Ogloff, Thomson, & Howells, 2007). That is, due to the context of the institutional environment, some behaviour might emerge despite its inconsistency with behaviour in the community; alternatively, the function of the emergent behaviour might be inconsistent (Daffern, 2010). It may be important to determine if these inconsistent behaviours remain relevant for observation in custody (see McDougall et al., 2013). Discrimination between the underlying factors leading to certain behaviours can be achieved through conducting a functional analysis of the behaviours to determine their potential cause as either offence-related or context-specific (Clarke et al., 1993).

Due to the differences in context between custodial and community settings, in order to measure risk factors in custody, it is important to identify the ways in which these risk factors might manifest in custody. Once these behavioural manifestations have been identified, they can be monitored through observation, and they can be more effectively targeted through intervention (Gordon & Wong, 2010).

It has also been suggested that dynamic risk factors predictive of recidivism may not be predictive of risk in the custodial context, for example, risk for institutional misconduct, because there might be differences in the causes of misconduct in custody compared with offending in the community. As a result, the assessment of dynamic risk factors in custody may require context-specific measures that operationalise factors related to poor institutional adjustment (Makarios & Latessa, 2013).

Turning to non-sexually violent offenders, the behavioural manifestations of risk factors related to non-sexually violent offending may be more readily observable in the custodial environment (e.g., physical aggression, emotional outbursts) than some behaviours exclusively relevant to sexual offenders. Daffern, Howells, Mannion, & Tonkin (2009) developed structured methodology to assist professionals compare the function of behaviours between contexts in a violent offender population. In order to compare an index act with in-hospital aggressive behaviours, Daffern et al. developed a structured aggressive behaviour analysis schedule assessing the following factors: victim characteristics; schema; affective antecedents; physiological activation; environmental context; disinhibitors (e.g., drug and/or alcohol use); opportunity factors; weapon use; and, function of behaviour. Their findings revealed that aggressive behaviours occurring within hospital were no more similar to a patient's index behaviour than randomly selected aggressive behaviours perpetrated by a different patient. However, aggressive behaviours that contained four or more similar elements to the index act were significantly more similar than randomly selected in-patient aggressive behaviours (Daffern et al., 2009). One suggested explanation for these findings was that some individuals might not display entrenched or consistent aggressive behaviour, whereas some individuals might consistently display certain offence-related behaviour.

As previously noted, a connection has been established between sexual deviance and sexual reoffending (Hanson & Morton-Bourgon, 2005). In practice, professionals assess sexual deviance to identify the potential for future sexually deviant behaviours. Due to the challenges associated with the assessment of sexual deviance based on observable behaviour, there is some reliance on self-report. However, in their research (which will be further described in section 1.1.12.1, p. 50), Seifert, Boulas, Huss, and Scalora (2017) found no relationship between self-report measures of sexual fantasies and actual sexually deviant behaviours (e.g., sexual compulsivity and sexual sensation seeking). That is, none of the behavioural predictors of sexual deviance were significant predictors of scores on the self-report measures of sexual fantasy. Seifert et al. suggested that the lack of temporal proximity between the measurements of the variables may have accounted for the non-significant findings; that is, participants responded to self-report measures on admission to treatment, which was several years after they had engaged in the sexually deviant behaviours. Sexually deviant behaviours occurring while incarcerated were not available in the study. Nevertheless, these results

revealed additional challenges associated with the measurement of some risk factors predictive of sexual recidivism. It further highlights the importance of gaining information about offenders' behaviour in custody to assist with behavioural monitoring and the interpretation of self-report measures of change.

As previously noted within this chapter, it is important to assess offenders' risk of harmful behaviour. While in custody, a determination of risk level is difficult due to the context of the environment. As a result, a means by which to assess risk based on observed behaviour in this environment is vital to the determination of risk level. Dynamic risk factors are associated with behavioural manifestations, which can be triggered in custody. Researchers and clinicians have grappled with this process. Various theories related to the operationalisation of risk factor manifestations have been suggested. Two of these theories will now be described: the offence paralleling behaviour (OPB) framework and the offence analogue behaviour (OAB) framework.

1.1.10.2.1 Offence paralleling behaviour. Through developing an algorithm to measure the behavioural 'similarity' of any two cases, behavioural consistency has previously been identified in serial sexual offenders (Grubin, Kelly, & Brunsdon, 2001). However, within custodial settings, it is difficult to gauge whether an offender's capacity to offend continues over time, if there are no opportunities or characteristic activating factors (Daffern et al., 2008). Jones (1997) operationalised the concept of the behavioural manifestations of risk factors within institutions through the OPB framework. The definition of OPB has been refined over time and in order to limit the potential misapplication of this framework, Daffern et al. (2007) defined it as "...a behavioural sequence incorporating overt behaviours (that may be muted by environmental factors), appraisals, expectations, beliefs, affects, goals and behavioural scripts, all of which may be influenced by the patient's mental disorder, that is functionally similar to behavioural sequences involved in previous criminal acts" (p. 267). There are two assumptions within the framework, which are relevant to offence-related behaviour such as sexual and non-sexual aggression: (1) behaviour within institutions is similar to past aggressive (offence-related) behaviour in another environment; and, (2) it is possible to reliably identify OPB.

Although there remains conjecture about their precise definition, dynamic risk factors are thought to have a primary role in offending and as such, they represent intermediate targets for treatment (Andrews et al., 2011). These dynamic risk factors manifest through sequences of OPB. Therefore, this framework can assist clinicians to

assess an offender's current risk based on behaviours observed in custody. The OPB framework was developed to be used in tandem with structured and actuarial risk assessment methods and treatment programs (Daffern, 2010). It was developed as a method by which idiosyncratic case formulations could be identified and used to predict relevant behaviours in custody, which could then inform risk assessment and treatment planning. As Mooney and Daffern (2011) noted, an individualised approach informed by the OPB framework may facilitate the risk assessment and treatment of incarcerated offenders through the identification of idiographic offence-related behaviours, which may manifest in custody. Monitoring these behavioural patterns in the custodial context could provide a more accurate representation of an offender's risk for future offending. According to Daffern (2010), the OPB framework emphasises the social reinforcing and therapeutic role that all staff can have in the treatment of violent offenders in custody. It offers a way for wing staff to understand and integrate the informal treatment work conducted between formal program sessions, through the observation and monitoring of offenders' behaviour.

A central component of the OPB framework is that for a risk assessment to be conducted using a comparison of behaviour (e.g., sexual behaviour) across settings, comprehensive assessment of the index offence is required. This assessment is referred to as the reference formulation. It is important for this assessment to include the contribution and relevance of dynamic risk factors related to the offending, the actual characteristics of the abusive behaviour, and the function of the behaviour; these characteristics would then be compared with the behavioural sequences in the institution to determine whether the components, function and behavioural sequence are equivalent. Daffern et al. (2008) emphasised the importance of remaining cautious in the implementation of the OPB framework; that is, it should not be assumed that behaviours are necessarily manifestations of offending behaviour indicative of future risk of offending when there is only topographical similarity. For instance, low-level sexually abusive behaviours (e.g., sexually suggestive comments) in an institution may not be manifestations of sexual deviance. Rather, they may be the expression of anger; therefore, the function of the behaviour should be determined prior to concluding whether it is offence paralleling or not. Specifically, the functions of aggressive behaviour in institutions may be different from the functions of aggressive behaviour in the community (Daffern & Howells, 2009).

The OPB framework has been used to examine similarity in aggressive behaviours across situations (e.g., to determine whether behaviours offenders engage in within custody are offence paralleling). For instance, preliminary research has been conducted to assess the validity of the OPB framework for incarcerated violent offenders (Daffern et al., 2009), which was described earlier in this chapter (section 1.1.10.2, p. 36). One methodological limitation identified in Daffern et al.'s (2009) research, was that only topographically similar behaviours were assessed for similarity. Behaviours can be classified as offence paralleling if they show functional equivalence despite topographical dissimilarity. In addition, due to the potential evolving nature of an individual's aggressive behaviour, the index act might not represent an individual's prototypical offending patterns.

Using the terminology of offenders' vulnerabilities and manifestations of these vulnerabilities, Daffern (2010) noted that persistent vulnerabilities do not necessarily manifest in overt aggression in institutions. Therefore, the task for clinicians is to determine the ways in which persistent vulnerabilities or dynamic risk factors manifest in behaviour within custody. A necessary requirement for aggression observed in an institution to be considered offence paralleling is that equivalent psychological features (e.g., schemas of abuse and mistrust) had a functional role in the initiation of prior aggressive behaviour and maintain causality in current aggressive behaviour. For instance, an offender's behaviour may be triggered by the activation of schemas through perceived rejection of his partner in the community and similarly triggered by perceived rejection by a therapist in custody (Daffern, 2010).

Daffern (2010) indicated that the OPB framework promotes comparison of two or more sequences of behaviour to determine similarity; this similarity is concluded by comparing functional analyses to determine whether psychopathology related to past or future offending is maintained. Importantly, comparison of behaviour in an institution with a single index act is problematic because this single act might not be representative of the offender's entire repertoire or their typical aggressive behaviour. Therefore, Daffern identified that it is more important to compare the offender's entire repertoire and their trajectory of behaviour prior to incarceration, with their current behaviour.

One difficulty in the task of determining whether behaviour is offence paralleling or not, is that there might be some, but not total, similarity. The inclusion of high base-rate components of an offending-type behaviour may result in an over identification and mislabelling of aggressive behaviours as offence paralleling (Daffern, 2010). For

instance, the custodial environment might trigger antisocial behaviours (e.g., argumentative behaviour towards staff) that appear to be offence paralleling but are in fact only a product of the environment. As a result, staff must be mindful that they do not label behaviours as offence paralleling when they are not (Mann, Thornton, et al., 2010) and ensure that it is consistency in the sequence and function of behaviours rather than one topographically similar behavioural element.

1.1.10.2.2 Offence analogue behaviour. Gordon and Wong (2010) use the label OAB to describe the “here-and-now markers for the individual’s criminogenic needs, that is, the current manifestations of the individual’s problem areas within custodial settings” (p. 172). In general, these OABs are idiosyncratic to the individual, linked to the individual’s criminogenic needs, and result from an interaction between these criminogenic needs and the immediate environment (Gordon & Wong, 2010).

One key difference between the OPB and OAB definitions is that whereas OPBs refer to a behavioural sequence that mirrors an individual’s offending behaviour in either topography or function, OABs refer to behaviours, which may be discrete instances, mirroring criminogenic needs (dynamic risk factors), which are not necessarily specific to the offender’s index offence. These OABs are the manifest propensities as described by Mann, Thornton, et al. (2010).

1.1.10.2.3 OPB vs OAB. McDougall et al. (2013) noted that a subset of offenders’ offence-related behaviour in custody might be offence paralleling; however, there might be other important offence-related behaviours, which, although not reflecting parallels with the index offence, are also important and should be examined. These behaviours include negative low-level behaviours, which might be indicative of the potential for more severe behaviour such as offending. Therefore, in their study, McDougall et al. examined all behaviours that might be classified as offence-related, or negative, in addition to positive behaviours, without engaging in functional analysis.

With high-risk but infrequent behaviours (e.g., murder, in the absence of prior aggressive behaviours), it may be difficult to develop a meaningful and observable parallel, which subsequently limits assessors’ ability to measure relevant behaviour in custody (Davies et al., 2010). This difficulty may promote an advantage of the OAB framework, in which behavioural manifestations of dynamic risk factors are identified rather than limiting the assessment process to development of parallel sequences to offending behaviour. Furthermore, the OAB framework may more easily facilitate the assessment of the meaning and significance of observed changes. That is, Davies et al.

(2010) noted that if the OPB has considerable functional equivalence to the offending behaviours, the tendency may be for clinicians to confidently conclude that change has been achieved in the risk of engaging in offending behaviours. However, the change in OPB might not necessarily remain when the offender is faced with different contexts, resources, demands and stressors, such that it might not be generalisable to contexts in which an offence may be imminent.

1.1.10.3 Behavioural manifestations of protective factors. Andrews et al.

(2011) asserted that alternative ways of thinking, feeling and acting can be developed to counteract an offender's risk factors. A key part of rehabilitation is the development of skills to replace previously learned maladaptive behaviour patterns (Baglivio, Wolff, Piquero, Howell, & Greenwald, 2017). A central feature of risk assessment is the measurement of changes in criminogenic factors in order to determine whether these factors have moved away from risk, towards becoming strengths (Hanson et al., 2009). This idea is relevant to both the OPB and OAB frameworks.

1.1.10.3.1 Prosocial alternative behaviour. With the implementation of effective intervention for OPB, it is suggested that adaptive, prosocial behaviours with similar functions to the OPB may emerge and replace OPB (Daffern et al., 2007). These behaviours have been termed prosocial alternative behaviours (PAB) and are observed through skills acquisition or adoption of prosocial behaviours (e.g., assertive communication) in contexts that would be expected to elicit problematic behaviours (e.g., violence). Therefore, progress is not only measured by a decrease in OPB, but also by an increase in PAB (Daffern et al., 2007). In the risk assessment process, observation of the changes in OPB towards PAB is important in determining change in risk level.

1.1.10.3.2 Offence replacement behaviour. Similarly, Gordon & Wong (2010) identified offence replacement behaviours (ORB) as more 'helpful behaviours' that are developed to counteract OABs. Consistent with Daffern et al.'s (2007) suggestion, they noted that while it is positive for offenders to engage in fewer OABs, it is not sufficient; individuals must also engage in a greater number of appropriate behaviours.

1.1.10.3.3 Protective factors. Returning to the concept of protective factors, it could be argued that the acquisition of prosocial behaviours in the form of PAB and ORB is equivalent to the development of protective factors. As noted earlier, there remains debate about the most valid method by which to operationalise protective factors. These factors may contain both state and trait constructs, as could risk factors (Klepfisz et al., 2017). However, based on the perspective that protective factors are

equivalent to the underlying propensity in the same way that risk factors are equivalent to the underlying propensity (de Vries Robbe et al., 2015), it can similarly be argued that the behaviours exhibited as a reflection of these protective factors are in fact the observable behavioural manifestations, which are the PABs/ORBs.

Polaschek (2017) noted that a factor is protective if the individual uses it when given the opportunity to do so in the face of offending opportunities. This idea is consistent with that proposed within the OPB framework, in which PABs are defined as the positive behaviours an offender engages in within a situation that could trigger offending-related behaviour, instead of engaging in the pattern of problematic behaviour he/she would have engaged in previously. Daffern et al. (2007) emphasised the importance of determining whether the observed positive behaviours are related to positive skill development or unhelpful factors. For instance, they identified three potential scenarios in which an offender might appear to be managing his or her relevant risk factors. One is situational muting, which might be a result of either environmental or other individual factors (e.g., potentially restrained or secluded, or lacking certain mental state factors that might lead to risky behaviour). These factors might mute the expression of an offender's behaviour such that the result was less extreme; however, it does not necessarily indicate that the offender has made positive changes and has acquired behaviours to successfully manage his offending propensities. Second, problematic behaviour might appear to decrease as a result of detection evasion skills, in which the behaviour persists but skills are acquired such that the offender can continue to engage in it without detection. Finally, the apparent change might be due to more positive psychological processes developed, which lead to use of a more prosocial behaviour in response to triggers.

1.1.10.4 Practical issues related to behavioural observation. In order to monitor the presence of dynamic risk through behavioural manifestations, it is important to conduct systematic and objective assessments to reduce subjectivity, observer bias and the reliance on accidental observations (Daffern et al., 2007). For instance, a potential limitation to the assessment of risk based on observations of behaviour in custody is that the presence of a risk factor may depend on the degree to which behaviour is monitored by staff as well as how it is interpreted. Therefore, observer biases can compromise the assessment and classification of behaviour. For example, observers may attribute the causes of an offender's institutional behaviour to characteristics of the person, rather than considering temporary factors within the

offender or characteristics of the environment (Daffern, 2010); the behaviour may thus be incorrectly attributed to underlying vulnerabilities. Further, Serin, Chadwick, et al. (2016) highlighted the potential for observer bias, such that ratings of change might be affected by the relationship the assessor has with the offender and their personal opinion of this individual, as well as the importance they place on various aspects of change.

This problem could be solved for commonly occurring risk factors by developing checklists with the most usual types of institutional behaviour associated with each risk factor (McDougall et al., 1995). Predictions could therefore be made about institutional behaviour that would indicate whether or not the risk factors were being demonstrated through the analysis of each offender's case (McDougall et al., 1995). McDougall et al. (2013) further suggested there may be benefits from using a behavioural monitoring form such as that developed within their research, in combination with a range of actuarial and structured clinical assessment tools available.

In addition, Daffern et al. (2007) suggested that it would be beneficial for clinicians to predict the behavioural sequences that are likely to manifest, by analysing the individual offence patterns and systematic observation of behaviours in custody. Risk assessments such as the VRS have been used for this process in incarcerated violent offenders. Gordon and Wong (2010) suggested that OAB and ORB should be monitored and that the list of these behaviours can be amended throughout treatment, through observation by various staff members in different contexts, including both the treatment sessions and daily functioning (e.g., at work or in the prison unit).

Gordon and Wong (2010) encouraged the use of behavioural observations for offenders over time and place, rather than restricting assessment to time spent in 'formal' treatment contexts. If observations are restricted to treatment contexts, they might be biased because offenders spend relatively small amounts of time in these contexts and often attempt to present themselves in an overly positive light. Additionally, offenders may use detection evasion skills in environments in which they feel more scrutinised, such as within the therapy room. Thus, positive impression management within treatment settings might lead to staff members forming unrealistically positive views of offenders' behavioural changes. Offenders' application of these adaptive behaviours might be less frequent outside of this scrutinised environment. Therefore, through observation outside of treatment, this discrepancy between contexts can be gauged.

There are more opportunities for offenders to practise developing skills learned within the treatment context and for a range of staff members to observe these behaviours in alternative contexts, if behavioural monitoring occurs in offenders' daily living situations (e.g., custodial officers' observations). Further, if psychologists use custodial officers' reports of offenders' behaviour, assessment methods may be enhanced by incorporating a psychological direction to observations regularly made in the custodial environment (McDougall et al., 1995). This variety of contexts in which to monitor offenders' behaviour is similar to the method McDougall et al. (2013) used, in which custodial officers' observations were made across various settings.

Important additional information can be gained regarding offenders' behaviour through use of custodial officers' observations. However, custodial officers might be less inclined to report some behaviours. Atkinson and Mann's (2012) study included exploration of custodial officers' decisions about whether or not to report offenders' behaviour. Focus groups were conducted with custodial officers. Some reasons custodial officers provided for not reporting offenders' behaviour included there being an accepted level of poor behaviour in custody, and relatedly, the behaviour not considered concerning enough to justify reporting it. However, as evidenced through the work cited in McDougall et al. (2013), low-level behaviours can be important in determining offenders' risk.

The model implemented in McDougall et al.'s (2013) study was designed to involve custodial officers in the monitoring process as part of their daily work. Specifically, custodial officers' observations of offenders' day-to-day activities in custody may provide greater information about offenders' behaviour outside a direct intervention context (e.g., education, employment, social activities). One issue that might increase the challenges associated with this process is its reliance on custodial officers' regular interactions with offenders and active engagement in behavioural observation. That is, some custodial officers' interactions with offenders may be limited to that required for the maintenance of the safety and security of the prison rather than related to more generalised situations. Therefore, information might not be gained regarding offenders' risk-related behaviour. Nevertheless, if several methods are used to monitor behaviours, they can be compared with each other to determine consistency. For instance, in addition to using custodial officers' observations, clinicians' observations of offenders' behaviour could similarly be gained. Further, often custodial officers are unclear about the relevance of observed behaviour to risk of reoffending,

particularly when the custodial behaviour is not topographically similar to prior offending. Therefore, it might be helpful to provide them with potential risk-related behaviours to focus their monitoring.

1.1.11 Behavioural Monitoring in a Therapeutic Community

1.1.11.1 Therapeutic community environment. Many treatment programs in custody are implemented within the context of a therapeutic community. Forensic Therapeutic Communities are residential units within an institutional rehabilitative context. They allow every event and any relationship within the environment to be considered a learning opportunity, which can assist offenders maximise therapeutic gain. They provide an environment where treatment gains and new learning from group therapy sessions can be rehearsed outside the group setting, that is, within the therapeutic community environment. Therefore, therapeutic communities provide offenders with an opportunity to increase the intensity of the treatment experience beyond the group treatment context. The culture has a positive and rehabilitative focus, which is developed and maintained with the active participation of both staff and offenders (Ware, Frost, & Hoy, 2010).

In these settings, open communication is encouraged between offenders and staff as well as between multi-disciplinary staff members. As a result, all staff members are involved in the implementation of the community environment. For instance, the role of custodial staff extends from the standard provision of humane, secure and safe containment to actively promoting the change process (Ware et al., 2010). Custodial staff monitor offenders' positive and negative behaviour in the wing and engage in day-to-day communication with therapeutic staff regarding offenders' progress in working towards their goals. For therapists, it is vital to maintain communication with custodial staff regarding clients' behaviour and to draw upon their observations and impressions to help gauge treatment progress and to assess risk of reoffending. This communication between staff members also serves to strengthen the treatment process for clients.

1.1.11.2 Behavioural assessment. Prior research has suggested that pre-treatment psychometric measures were more predictive of reconviction than were post-treatment psychometric measures (Barnett, Wakeling, Mandeville-Norden, & Rakestrow, 2012). Although these research findings were explained through the potential impact of social desirability, there are possibly differences between the context of the community and custody on offenders' demonstration of positive impression management. For instance, in a custodial context, the impact of social desirability might be countered by other

factors, such as those inherent in a therapeutic community. Initially, this environment might elicit reactions based on prior experiences. Offenders might experience feelings of threat and fear as a consequence of prior experiences in and out of custody, in particular due to the stigma surrounding sexual offenders. As a result, they might be more likely to present themselves in an overly positive light in response to other offenders and staff members, for physical and emotional self-preservation. As they progress through treatment, their new experiences might lead to changes in their attitudes such that they feel safer to disclose information that might present them in a negative light.

Blagden, Winder, and Hames (2016) noted that the findings of their research supported the idea that offenders found the custodial context in which they engaged in treatment, to be safe and constructive, which contributed to their development of positive cognitions around the experience. In tandem with this process, offenders will likely gain further insight into their attitudes and behaviour and learn alternative ways by which to manage difficulties, through both the treatment process itself and dynamic interpersonal learning (Yalom, 1985), such that staff will observe behavioural improvements. Therefore, offenders' post-treatment psychometric measure responses might provide more accurate measures of their attitudes and beliefs. Based on these factors, it is possible that in combination with staff observations, psychometric measures might provide useful information regarding assessment of risk at the completion of treatment. Further, self-report measures may be valuable in combination with staff observation to facilitate greater opportunity for a wider spectrum of behaviours to be monitored.

1.1.11.3 Challenges with observation. When assessing risk, one advantage of therapist judgements compared with offender self-report is that it avoids any potential social desirability bias (Beggs & Grace, 2010). However, due to the nature of residential environments, the range of social and physical activities will likely be restricted, which might reduce the observable manifestation of risk factors (e.g., alcohol dependence may not produce easily recognisable manifestations within custody, whereas anger dysregulation might). Further, a difficulty in relation to observation of offenders' behaviour in custody is that some behaviours are more easily observable (e.g., self-regulation deficits) than other behaviours (e.g., deviant sexual fantasies). As a result, there might be a higher rate of observation of these easily observed risk factor manifestations (Hanson & Harris, 2001).

Sexual deviance might be particularly difficult to measure through observation. Not only are the behavioural manifestations more likely than many other behaviours to be concealed, but many manifestations of this risk factor are likely to be in the form of attitudes and beliefs. However, it can be difficult to identify and measure the extent to which attitudes or beliefs are present. For sexual offenders, the presence of offence-supportive beliefs is often inferred from the statements offenders make about their offending (Mann, Hanson, et al., 2010). There might be a limit to the extent to which staff within the custodial environment can observe risk that is manifested through thoughts and beliefs. The use of inference rather than purely direct observation might allow a greater degree of subjectivity and bias from staff, which likely impacts on the reliability and validity of observations.

Brown et al. (2009) noted that additional research was required to determine the most effective method by which to assess real-world dynamic risk factors prior to release from a secure environment. They suggested that future studies should investigate methods by which to reliably assess behaviour in custody, which most closely represents real life situations, such as: how the offender spends leisure time in custody; debt acquisition in custody; and, employment performance in custody. It should be emphasised that as previously indicated, prior research has demonstrated the similarity between behaviours in custody and those in the community (Zamble & Porporino, 1990). Behavioural manifestations of risk factors in custody likely approximate relevant behaviours in the community. This process is vital in order to assess ongoing risk in the custodial context prior to offenders' release; however, it is a process that is inherently challenging, in part due to the difficulties related to observing dynamic risk factors in this context.

1.1.12 Evaluating Change in Offenders

A crucial component of dynamic risk assessment is the assessment of change over time, to determine whether offenders' changes through treatment have corresponding effects on changes in outcome such as recidivism (Olver & Wong, 2013). A challenge associated with change assessment is determining salient risk-related changes in a population of men residing in a controlled institutional environment, and for whom there are generally few readily available opportunities for offending or access to potential victims (Gordon & Wong, 2010). It is important to note that the processes by which sexual offenders make changes are not well known (Hanson et al., 2009). Results of various studies have suggested that improvements on factors presumed to be

criminogenic, in fact have no effect on sexual recidivism rates (e.g., hostility towards women; Quinsey, Khanna, & Malcolm, 1998). These results highlight the lack of clarity surrounding the relation between change in dynamic risk factors and recidivism.

It is worth noting that in order to effectively measure change in these behaviours and attitudes, the relevant dynamic risk factors and their prosocial alternatives must be identified and defined. There are various methods by which offenders' change can be assessed. For instance, psychometric tools can be used to assess theoretically important psychological and/or risk constructs relevant for sexual offender risk and wellbeing (Olver & Wong, 2013). These tools tend to be used in the form of pre- to post-treatment self-report measures, and can provide information about an offender's changes in attitudes and beliefs in addition to behaviour. Further, clinicians' ratings of change can be used in the form of pre- to post-treatment structured clinical risk evaluations (e.g., VRS:SO). Changes in behaviour can be gauged through staff observations, or alternatively, methods such as institutional misconduct over a specified time period (e.g., throughout an offender's incarceration). There are differences between change measures in relation to the number of time points used to assess change. The various methods of measuring change and the challenges surrounding this process will now be further outlined.

1.1.12.1 Psychometric measures through self-report. Willis, Yates, Gannon, and Ward (2013) emphasised the benefits of using self-report processes within assessment and intervention planning. Self-report can serve to minimise the challenges associated with behavioural observation, as previously discussed. This idea was highlighted through the Good Lives Model, in which self-report is recommended to aid clinicians within assessment and intervention planning. For instance, according to Willis et al. (2013), use of an assessment feedback session could serve to ensure the clinician has an accurate understanding of an offender's values, their relationship to offending behaviour and life problems, and offender strengths. Self-report can also be used in relation to change measurement, such as through use of psychometric measures.

Self-report psychometric measures are used to gain insight into unobservable (latent) variables, and measure personality traits, abilities, attitudes and knowledge (Beech, Wakeling, Szumski, & Freemantle, 2016). There has been ongoing discussion about the advantages and disadvantages of using psychometric measures to assess change in attitudes and behaviours over time, largely due to the potential impact of positive impression management in self-report measures. However, the advantages of

using psychometric tools include their convenience and their lessening of assessor biases (Wakeling et al., 2011). In addition, prior research findings have provided evidence to suggest that psychometric assessment can be used to measure dynamic risk of sexual offending and that this form of assessment adds to the predictive power of static risk assessments (Barnett et al., 2012).

In Barnett et al.'s (2012) study, dynamic risk-domain scores were calculated for all offenders, to represent a measure of the level of problematic behaviour in each risk domain for each offender. This measure was based on the standardisation of psychometric assessment scores. A key finding was that post-treatment scores on psychometric measures were less discriminative and predictive of reconviction than were pre-treatment scores. In contrast to the changes in the effects of social desirability over time within a custody-based therapeutic community, one potential explanation the authors provided was that post-treatment scores might be more likely affected by social desirability. The authors suggested that as a result, clinicians should be wary in using these measures as a basis for risk assessment.

It is possible that the authors' explanation was related to the context of the study; that is, the sample consisted of sexual offenders attending a community-based probation service-run treatment program. Offenders under community supervision may be more cognisant of the potential for breaching their order and subsequent incarceration, should they perform poorly in treatment. This hyper vigilance might increase the likelihood of socially desirable responding.

Research has been conducted to assess the role of response bias on self-report measures of sexual fantasies (Seifert et al., 2017). Results examining the association between the Marlowe-Crowne Social Desirability Scale (MC-SDS; Crowne & Marlowe, 1960), and the Wilson Sexual Fantasy Questionnaire (WSFQ; Wilson, 1988) and the Sexual Fantasy Questionnaire (SFQ; O'Donohue, Letourneau, & Dowling, 1997) indicated that both the measures of sexual fantasies were influenced by socially desirable responding. That is, higher scores on the MC-SDS were associated with lower scores on the sexual fantasy measures, which suggested that participants attempted to present themselves as less deviant in their self-reports of sexual fantasies. Results also suggested that although responses on these measures were influenced by social desirability, the level of social desirability did not suppress the robustness of the significant associations between sexual fantasies and self-report measures of sexual

deviance. Seifert et al. (2017) concluded that offenders' level of socially desirable responding does not invalidate responses on measures of sexual fantasy.

Further, Stevens, Tan, and Grace (2016) investigated the effect of controlling for the variance associated with socially desirable responding as measured by the MC-SDS, in the self-reports of sexual offenders in a custody-based treatment program within a therapeutic community. While MC-SDS scores were negatively correlated with psychometric variables associated with risk (social inadequacy, sexual interests, anger/hostility, and pro-offending attitudes), results reflected that correcting for socially desirable responding had little impact on predictive validity for recidivism. Correlations of individual variables and dynamic risk factor scores with recidivism were not significantly altered when socially desirable responding was removed, such that there was no evidence for the relationship between dynamic risk and sexual recidivism being dependent on socially desirable responding. Stevens et al. indicated that these results are consistent with prior research suggesting that this form of responding should be regarded as a personality characteristic or trait, rather than a response bias that when removed increases the accuracy of risk assessment. They also reported finding that it was unrelated to recidivism and static risk although still correlated with dynamic risk measures. They noted that their results suggested such responding did not pose a threat to the predictive validity of dynamic risk assessment through self-report.

Many offenders with a history of sexual offences will respond well to exploration of their sexuality and will acknowledge persistent or sporadic deviant sexual interests (Craissati & Beech, 2003). However, change over the course of treatment might be more difficult to ascertain reliably. Due to the potentially significant consequences of those in authority perceiving an offender to either have made sufficient changes or not, offenders may seek to present themselves in a positive light to researchers and treatment providers, such that they report a reduction in their potential risk for recidivism. Despite this potential for positive impression management through self-report measures, prior research has provided evidence that staff perceptions of offenders are in agreement with the offenders' self-perception and expressed beliefs (McDougall et al., 1995). However, clinicians must remain mindful that staff perceptions of offenders might be influenced by offenders' expressed beliefs.

Walters (2006) reported finding that psychometric scores could predict outcome with the same accuracy as other risk assessment tools, but only if the psychometric measures were designed to measure constructs that were empirically related to risk of

reoffending. One important point in relation to using psychometric measures is that more deviant scores and lower levels of change should theoretically be linked with higher rates of recidivism; however, individuals with more deviant scores have the opportunity to attain higher change scores. Therefore, use of raw change scores within these measures has limitations that must be countered. Use of clinically significant change (CSC; Jacobson & Truax, 1991) manages this limitation through evaluating post-treatment scores against non-deviant norms to determine the practical value of the individual's changes. Therefore, note that this form of analysis is used to determine individual-level change, rather than the group-level change.

1.1.12.2 Structured clinical evaluation. In the past, risk assessment measures tended to focus on the prediction of recidivism; more recently, greater attention has been provided to the assessment of change and incorporating information about intermediate change (e.g., that achieved through treatment) into sexual offender risk appraisals (Olver, Christofferson, Grace, & Wong, 2014). Central to the use of these assessment measures is the idea that risk is dynamic, it can be accurately operationalised and measured with current dynamic tools, and that risk-relevant changes should be related to changes in recidivism outcomes (Sowden & Olver, 2016). Measurement of behaviours in custody, and further, measurement of changes in these behaviours, may allow more effective prediction of offenders' potential behaviour upon release. For instance, by assessing change in an offender's OABs within treatment, an indication can be gained regarding treatment progress. As treatment progresses, and offenders make improvements, the intensity and frequency of OAB occurrences may decrease (Gordon & Wong, 2010). This progress may also be indicative of changes in risk level if treatment is effective in targeting relevant risk areas.

Olver, Wong, Nicholaichuk, and Gordon (2007) suggested that some dynamic risk factors may be more resistant to change processes than others. Therefore, it is also possible that the changes in intensity and frequency of behavioural manifestations of risk factors might differ. This difference in the ability of various risk factors to undergo change over time might be relevant for clinicians' attempts to evaluate clients' changes subsequent to intervention or prior to release from custody.

The VRS:SO was developed to integrate risk assessment and treatment planning, in addition to change measurement, in a single instrument (Olver et al., 2007). It includes both static and dynamic variables to assess risk for sexual recidivism, with the dynamic variables additionally used to identify treatment targets and to measure

changes in risk pre- to post-treatment. The measurement of change is based on the application of the transtheoretical model of change (Prochaska, DiClemente, & Norcross, 1992), which suggests individuals progress through five stages of cognitive, experiential and behavioural changes. Behaviours are described in the VRS:SO and changes are assessed and quantified through the application of a modified version of the transtheoretical model of change. The VRS:SO has been validated as a change measurement instrument (Olver et al., 2014; Olver et al., 2007; Sowden & Olver, 2016).

Wakeling et al. (2011) suggested that based on their research findings, the most useful psychometric variable to examine as an indicator of future recidivism, may be a sexual offender's overall level of deviancy pre-treatment, using a composite of all of the psychometric measures. However, when comparing predictors of sexual recidivism, Beggs and Grace (2010) reported that although there was no statistically significant difference, VRS:SO scores were better predictors of sexual recidivism than the Deviance scores derived from psychometric tests. They concluded that it may suggest structured therapist judgement is a more effective methodology for risk assessment than self-report.

The inconsistency in research findings related to measurement of sexual deviance in offenders might reflect the challenges associated with this process. Since this risk factor is likely the most difficult to measure through methods other than offender self-report, a combination of self-report and staff observations may provide a greater opportunity for all behaviours to be monitored.

Sowden and Olver (2016) examined the assessment of change in sexual violence risk using two risk assessment tools, the VRS:SO and STABLE-2007 (Hanson et al., 2007). Both risk assessment tools demonstrated small-to-moderate prediction effect sizes for four recidivism outcomes: sexual, non-sexual violent, violent, and general recidivism. Only the VRS:SO predicted sexual recidivism in the sample. Both measures demonstrated participants' pre- to post-treatment change. On average, participants scored about three quarters of a standard deviation lower on each measure post-treatment, which is consistent with prior research (e.g., Nunes, Babchishin, & Cortoni, 2011). The VRS:SO change scores were not significantly associated with reductions in sexual recidivism specifically, while STABLE-2007 change scores were not significantly associated with any recidivism outcomes. It was noted that as the VRS:SO was intended to primarily assess risk for sexual violence, the observed associations of risk change with reductions in other recidivism outcomes is consistent with the

possibility that risk-relevant change may extend to reductions in the risk for outcomes beyond sexual violence.

1.1.12.3 Number of time points measured. Brown, St. Amand, and Zamble (2009) noted that one of the greatest challenges facing correctional research is establishing a ‘gold standard’ for the statistical analysis of change data in relation to recidivism, with no consensus regarding the most appropriate method. Prior research has examined risk factors that were described as dynamic, at one specified point of time. In effect, this process serves to treat a dynamic variable as static. It provides no information about change over time, or whether more proximal assessments may be more effective indicators of imminent risk (Serin & Lloyd, 2016).

The majority of prior research that has investigated changes in dynamic risk over time has done so through the use of two time points (Lewis, Olver, & Wong, 2013). These two time points have generally been pre- and post-treatment (Walters, 2006). Specifically, within-treatment change has been assessed through measurement of dynamic risk factors prior to treatment and reassessing these same factors following treatment completion. As described earlier, these assessments typically involve psychometric measures designed to measure variables that were empirically or theoretically associated with risk, were assumed to be dynamic, and were targeted for change in the treatment program (Beggs, 2010). This form of assessment is also the method used within the VRS:SO as described previously.

Other methods by which to analyse changes measured over time include the use of test-retest scores (Nunes et al., 2011). Hanson and Harris (2000) recommended the creation of an actual change variable, whereby the observed level of change is assigned a code (e.g., -1 = deterioration; 0 = no change; and 1 = improvement).

Walters (2006) noted that the assessment of dynamic risk factors at multiple time points during and after treatment appears to be an effective technique to improve change assessment. Similarly, Lewis et al. (2013) suggested that linking recidivism to dynamic variables assessed at multiple time points might be a more accurate assessment of the dynamic nature of the variables. Brown et al. (2009) suggested that the ideal study requires at least three distinct assessments of dynamic risk, as without multiple time points, researchers cannot differentiate measurement error from actual change. They proposed the use of a longitudinal prospective research design, incorporating information about risk factors that fluctuate over time. It is not clear how often various risk factors must be measured to capture their true rate or nature of change; however,

Douglas and Skeem (2005) noted that the longer the interval between assessments, the greater the risk that changes will be missed.

Douglas and Skeem (2005) suggested that “What is needed is prospective, repeated-measures studies of hypothetically dynamic risk factors... with enough observations to discern patterns or trajectories of change... as well as rapidity of change” (p. 371). Exploratory research has been conducted with violent offenders in New Zealand, in which multiple time points were used to assess change (Yesberg & Polaschek, 2014). Therapist-rated pre- and post-treatment scores were used, with an additional assessment conducted six to 12 months post-treatment while participants were still in custody. At the group-level, there was generally a reduction in risk for non-sexual violence at the completion of treatment and subsequent to treatment while still in custody. At the individual-level, it was evident that offenders have disparate patterns of change both during treatment and subsequent to treatment.

Yesberg and Polaschek (2014) noted that the results suggested the direction and volume of in-program change does not necessarily allow for prediction of post-program change. They indicated that their use of file information alone to rate the third assessment of change may have been less sensitive than using information from the offenders directly; therefore, they suggested use of offenders’ self-report to prospectively assess a third rating of change in future research. They provided a tentative conclusion that failure to account for patterns of change subsequent to the end of treatment may help explain prior research results in which treatment change has not been predictive of future outcome.

1.1.12.4 Change findings. Where more recent studies have commenced exploring change, findings have been mixed. Some research has reflected minimal pre- to post-treatment change (Hildebrand & de Ruiter, 2012), while in other studies, more robust change has been found (Hudson, Wales, Bakker, & Ward, 2002). These equivocal findings raise a problem that has been observed in the psychometric assessment of dynamic risk. It is generally difficult to determine whether null results about change across treatment are the result of a real lack of change, insensitive measurement, or the investigated risk factors in fact not being dynamic (Cording, Beggs Christofferson, & Grace, 2016). The conceptualisation of dynamic risk factors may also contribute to the mixed outcomes of extant change research. For instance, Heffernan and Ward (2015) noted that while dynamic risk assessment tools may be reasonably accurate in their predictions of reoffending, this accuracy does not necessarily mean that the risk factors

used in the measures are psychologically meaningful, or that they facilitate further understanding of the aetiology of, or desistance from, antisocial behaviour.

Additionally, the sensitivity to change might vary depending on whether the tool includes factors that are both modifiable and causally related to offending (Viljoen, Shaffer, Gray, & Douglas, 2017).

Despite some researchers attributing difficulties regarding change observation to the investigation of static factors, Brown et al. (2009) suggested that static factors should also be used in change measurement. They found that static factors, in particular, the number of prison misconducts in the 12 months prior to release, made significant contributions to individuals' survival time. However, it could be argued that behaviour lending itself towards gaining institutional misconduct charges is in fact dynamic (e.g., aggression, substance use). Since misconducts received over different time points during a period of incarceration can fluctuate, it suggests this factor is dynamic.

Within van den Berg et al.'s (2017) meta-analysis described previously (section 1.1.8.1, p. 26), one of the aims was to determine the predictive validity of change scores on dynamic risk assessment instruments. Significant effects were found for change scores, such that these scores predicted all three types of recidivism (sexual, non-sexually violent and general); that is, offenders who demonstrated larger positive changes such that there was a reduction in dynamic risk scores, reoffended at a lower rate than those who demonstrated smaller changes in dynamic risk factors over time. All effect sizes were relatively small, indicating that only a small part of change in recidivism was explained by changes in dynamic risk factors. This finding raises questions about the nature of the variables included in dynamic risk assessments. The authors noted that if dynamic risk factors are predominantly correlates of psychological propensities, which are causally related to sexual offending, then treatment of these dynamic risk factors would still only treat the symptoms rather than the causes of sexual offending. Change scores would thus be expected to achieve small effect sizes. Nevertheless, the positive change observed in dynamic risk suggests a lowering of risk level, which is important in this field.

An indication of change in risk might be gauged if recent events have activated relevant risk-related propensities, with the immediate situation facilitating either prosocial or antisocial opportunities (Thornton, 2016). This opportunity for the observation of change might be available in a custodial setting, with either risk-related behaviour or more adaptive behaviour resulting from a given situation. Within a

treatment program, there might be different events occurring, which trigger responses (e.g., interactions between offenders, or with staff members; or particular treatment components). It might be argued that observation of real change could be concealed depending on the environment and the triggering situations. Countering this argument is the possibility that the response observed is in fact a manifestation of an acute risk factor, whereas over time, there will be a general trend in the methods by which offenders choose to respond to situations, with these decisions based on more stable dynamic risk and development of alternative skills to manage this risk.

1.1.12.5 Determination of change. It is not only the presence of risk factors, which should be measured to assess change. Additionally, as Serin and Lloyd (2009) highlighted, change may be most effective when individuals concentrate on developing prosocial habits rather than focusing purely on changing infrequent antisocial behaviours. The RNR principles include the development of strengths (Andrews & Bonta, 2010, p. 22). As previously discussed, change should be evaluated through observation of both positive and negative behaviours. Assessments such as those described previously (e.g., SAPROF and START) may assist with this process.

The model on which the RNR is based recognises active, conscious, and deliberate self-regulation through self-monitoring, comparison of ongoing behaviour and outcomes with standards of conduct, and self-talk, self-imaging, and self-delivery of consequences to assist in aligning behaviour with an individual's self-identity (Andrews et al., 2011). Notably, Carter and Mann (2016) identified that one concern with using the term 'treatment' is that there is a risk of failing to acknowledge the active role of the treatment participant in the process of change. Within change assessment, it is important to ensure the offender's attempts to action this process are considered and evaluated. This concern is consistent with Serin, Chadwick, et al.'s (2016) suggestion identified earlier in this chapter, in which the mere presence of a strength factor should not be automatically assumed to have a protective effect on an offender; the offender must choose to engage in an adaptive response, just as he/she has an active role within the change process throughout treatment.

The ultimate indication that change has been achieved is through an individual's desistance from further offending. Serin and Lloyd (2009) suggested that desistance is directly connected to the psychological mechanisms, which drive changes in patterns of offending. They noted the importance of exploring the process by which offenders cease crime, and whether it is spontaneous or in response to intervention. At times, it may be

difficult to gain an accurate measurement of the extent to which an offender has changed, because change occurs over time, such that it is a gradual process, and can include both improvement and deterioration within a given period of time. Additionally, as the changes are observed through offenders' behaviour, perspective and attitude, it is important to implement multiple measures of change, which target different aspects of these internal and external processes. These different measures may provide greater depth of information about specific dynamic risk factors that have improved or deteriorated over time, which will in turn allow clinicians to evaluate offenders' relevant changes.

1.1.13 Individual-Level Change

It has been argued that dynamic risk must be disentangled from aggregate recidivism outcomes, because while correctional interventions are tested and accredited partly on their ability to demonstrate a recidivism risk reduction at the group-level, it remains unclear how these interventions are beneficial at the individual-level (Hannah-Moffat, 2016). Therefore, there have been arguments made for the measurement of change at the individual-level rather than purely focusing on overall changes measured within a group of offenders.

Serin, Lloyd, Helmus, Derkzen, and Luong (2013) asserted that criminal justice organisations regularly make highly consequential decisions about individual offenders; however, focusing on a group-level understanding of treatment change weakens the decision-making rationale used within this process. That is, the gains that offenders make through treatment are often evaluated through group-level analyses such as recidivism studies, which determine whether the group as a whole reoffend less frequently than an untreated group. This group-level analysis obscures the various reactions that offenders have to treatment; some improve, some remain the same, and some deteriorate. These inconsistent changes within groups highlight the value of measuring individual-level change.

Beggs (2010) similarly suggested that the fact that some offenders reoffend post-treatment indicates that there are individual differences in the degree of benefit gained from treatment. Yesberg and Polaschek's (2014) findings, in relation to using multiple time points to assess change, supported this suggestion. Baglivio et al. (2017) also emphasised that some offenders make far more change through treatment than others, regardless of the starting point; they noted that the amount of change is important in

determining outcomes because there are individual differences in the magnitude of changes and the speed with which those changes are made.

Davies, Jones, and Howells (2010) also discussed the importance of engaging in change assessment to determine the change that has occurred, and that which has been maintained, on an individual basis. Serin et al. (2013) noted that the association between individual-level changes and reduced recidivism risk remains relatively unexplored. Although research into offender treatment suggests individuals can make detectable gains, the specific individual-level gains responsible for this change process have not been identified. Serin and Lloyd (2009) suggested that a greater focus should be placed on the individual differences that initiate, sustain and characterise offenders' changes.

The significant differences in frequency of non-sexual reoffending within Mews et al.'s (2017) study, described previously in this chapter (section 1.1.5, p. 11), appeared driven by a minority of individuals in the matched comparison group who reoffended at a particularly high rate. This result emphasises the importance of determining the impact of individual differences on change processes; for instance, engagement in other programs, attitudes towards sexual and non-sexual offending, or levels of motivation. It is clear that within sexual offender treatment programs, individual offenders may make treatment gains while others derive limited benefit. In addition, one conclusion, which can be drawn from the variability of results from outcome studies, is that different forms of intervention may be more effective for some offenders than for others. Therefore, it may be beneficial to determine which individuals benefit from which interventions, in order to implement more effective interventions (Nunes et al., 2011).

Beggs (2010) suggested that traditional treatment outcome measures such as recidivism should be separated from more proximal outcomes, with investigation into methods by which to assess the benefits specific offenders have gained from treatment. That is, there are important applications for within-treatment outcome assessment, such as using information about treatment gains within post-treatment risk assessments. Therefore, as suggested, change on treatment targets during the course of intervention should also be determined at an individual-level. Davies et al. (2010) cautioned that clinicians might provide a biased response in conducting change assessments, as they will likely have an interest in change having occurred. Therefore, there is a need for independent assessment of change. The potential value of including behavioural observations from other staff members throughout treatment is thus highlighted.

Further, Pearson and McDougall (2017) argued that the assessment of individual change is particularly important for serious offenders; that is, whether there is evidence of improvement or deterioration within the present case, prior to determining progression through custody and ultimately release. While actuarial risk assessment measures provide aggregate benefits, which are suitable for organisational strategy on structuring service delivery (for example, the initial use of the STATIC-99R within Corrective Services NSW Sex Offender Programs to assess treatment suitability), there are few empirically validated risk assessment measures evaluating individual change. Therefore, identification of effective methods by which to determine whether an individual has made relevant changes is crucial within the field, to assist professionals measure treatment outcome, update risk assessments, make decisions relating to offenders' progression through the custodial classification system and finally the granting of release.

1.1.14 Summary

The overall rate of sexual recidivism is low in comparison with other offender populations. Nevertheless, due to the significant impact this type of offending has on victims and society as a whole, it is imperative that sexual offenders receive effective treatment such that they can work towards making changes relevant to their risk and needs. In order to provide this treatment, there is a requirement for the assessment of risk and needs; however, there remains debate regarding the conceptualisation of dynamic risk factors. Further, there has been limited research to date concerning the behaviour of sexual offenders in custody that may be relevant to recidivism and the manifestation of dynamic risk factors in this context. In order to monitor risk-relevant behaviours in a custodial context, a greater understanding of potential manifestations of risk factors is required.

Not only is risk assessment vital in the ongoing treatment and management of sexual offenders; change measurement is also important in revising risk assessments and determining requirements for ongoing intervention. To date, the focus has tended to be on group-level change assessment. However, there is also a need for the determination of individual-level change. Conclusions regarding the most effective methods by which to measure individual changes are yet to be made. These issues are important within the realms of custody-based treatment programs and subsequent parole decision-making.

1.1.15 Current Research

There were five key aims and hypotheses within the current research. These aims and hypotheses extend the extant research base, in relation to the observation and measurement of behavioural manifestations of risk-related behaviours and their prosocial equivalents, in addition to the measurement of change over time. These aims and hypotheses were investigated through three studies.

1.1.15.1 Study 1. Prior research has demonstrated that value can be gained through using checklists outlining relevant behaviours for observation and monitoring (e.g., McDougall et al., 2013). However, questions remain regarding the specific behavioural manifestations of risk factors relevant for sexual offending, which would be expected within a custodial setting. Based on prior research such as Atkinson and Mann's (2012) study demonstrating the potential use for custodial officers' observations of offenders' behaviour in custody, the purpose of Study 1 was to provide a basis for the development of a behavioural checklist, which could subsequently be implemented in a custody-based treatment program. This behavioural checklist could be used to monitor offenders' behaviours in custody and additionally, to measure changes over time.

Based on this rationale, the aim of Study 1 was to investigate how dynamic risk factors for sexual offending and prosocial equivalent behaviours might manifest in a custodial therapeutic community, through conducting a survey of professionals experienced in working with sexual offenders in custody. It was hypothesised that sexual offenders' manifest dynamic risk factors for sexual offending and prosocial equivalent behaviours would be recognisable to clinicians, who would be able to provide examples of expected behaviours.

1.1.15.2 Study 2. There remains a need for further investigation into effective methods by which to measure offenders' change over time. One measurement tool, which has been developed for use in custody to assist in the parole decision-making process in Victoria, Australia, is the Satisfactory Behaviour Rating Guide (SBRG; Daffern, Thomson, Ogloff, & Sweller, 2014).

The aim of Study 2 was to explore the validity of the SBRG through a retrospective archival analysis of serious sexual and non-sexually violent offenders who had applied for parole. It was hypothesised that custodial officers' final ratings on the SBRG would correspond with official records pertaining to the domains of: violent behaviour; involvement in substance abuse or like behaviour; attitude to employment, education and/or rehabilitation; and, response to direction and supervision.

1.1.15.3 Study 3. In order to increase the potential for the accurate assessment of a wide range of attitudes and behaviours relevant to sexual offending, Davies et al. (2010) recommended using a combination of psychometric measurement, self-report and staff observations. Each of these three measurement modes were implemented in Study 3. Staff observations were gained through the behavioural checklists, SBRG, and the Treatment Gain: Short Scale, while offender self-report was gained through the behavioural checklists and pre- and post-treatment psychometric measures.

Study 3 included three aims. The first two aims were related to the behavioural checklist, which was developed subsequent to Study 1, as a method by which offenders' behavioural manifestations of risk factors could be monitored in custody. The first aim was to determine whether dynamic risk factors and prosocial equivalent behaviours manifested in a custodial therapeutic community in a sample of sexual offenders, and whether they were evident to both offenders and custodial officers; in addition, to determine whether some behavioural manifestations were more evident than others. It was hypothesised that sexual offenders' dynamic risk factors for sexual offending and prosocial equivalent behaviours would manifest in a custodial therapeutic community and would be evident to custodial staff and offenders, with some behavioural manifestations observed with greater frequency than others.

The second aim was to determine whether the frequency of dynamic risk factors and prosocial equivalent behaviours changed over the course of treatment; in addition, to determine whether custodial staff and prisoner's perceptions of the presence and change in manifest dynamic risk factors and prosocial equivalent behaviours corresponded, and whether there was a change in their correspondence over the course of treatment. It was hypothesised that over time, there would be a reduction in dynamic risk factors and an increase in prosocial equivalent behaviours. Based on prior research related to the tendency for the group process (particularly within therapeutic communities) to increase offenders' self-awareness and motivation for increased openness (e.g., Corey, 2017; Ware et al., 2010), the subsequent impact of treatment on offenders' development of alternative strategies to manage risk-related behaviours (Andrews et al., 2011), and the importance of emphasising for observers the behaviours that are relevant (see McDougall et al., 1995); it was hypothesised that initially, custodial officers' observations of participants' behaviour would differ from participant self-report due to participants' tendency to minimise reporting of dynamic risk factors and emphasise prosocial equivalent behaviour. However, over time in treatment, the

two would converge as participants' insight increased through treatment and custodial staff observed more prosocial and less risk-related behaviour.

The final aim was to determine whether changes in these behaviours corresponded with other markers of change, such as reliable and clinically significant change in the psychometric measures used in pre- and post-treatment testing, and the Treatment Gain scale and the SBRG. It was hypothesised that these changes would correspond with the reliable and clinically significant change outcomes measured through pre- and post-treatment psychometric tests. They would also correspond with change as assessed using the Treatment Gain scale. Additionally, those participants who demonstrated positive change (reduced dynamic risk factors and increasing prosocial equivalent behaviour) would be more likely to be regarded as Satisfactory on the SBRG.

Chapter 2: Study 1

A central component of working with incarcerated offenders is the assessment of risk and change. One difficulty for assessors is determining how dynamic risk factors manifest in the restricted and intensely monitored custodial environment (Jones, 2010); this difficulty impacts the scoring of risk assessment measures and efforts to measure change. This study sought to elucidate the behavioural manifestations of sexual offenders' dynamic risk factors for incarcerated sexual offenders. These behaviours may be considered by assessors when scoring structured risk assessment instruments, and further, may focus custodial staff members' attention on risk-related behaviours.

Prior to outlining the current study, attention will be drawn to some of the key issues presented in Chapter 1, which are particularly relevant for this study. Risk assessment will be revisited, with a particular focus on the importance of developing an understanding of the ways in which risk factors manifest in custody. The OPB and OAB frameworks will be further discussed. Behavioural monitoring in custody will also be reviewed, including the use of custodial officers' observations.

2.1.1 Sexual Offender Risk Assessment

Several approaches to risk assessment have been designed to improve the predictive accuracy of an offender's risk level, including both static and dynamic risk factors. Examples of structured risk assessment tools for sexual offenders that measure both static and dynamic factors include the Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003) and the Violence Risk Scale: Sexual Offender version (VRS:SO; Wong et al., 2003). The VRS:SO also includes a measure of change.

2.1.1.1 Behavioural manifestations of risk factors. In order to observe, monitor and assess offenders' changes in risk over time in custody, there must be an understanding of the ways in which risk factors manifest in this environment. That is, the behavioural manifestations of empirically-derived risk factors must be identified. However, although risk assessments conducted in custody are based upon assessments of behavioural manifestations of risk factors, the way in which these dynamic risk factors manifest in custody has been the focus of limited research (Daffern & Ogloff, 2017).

Assessors must also be aware of the consistency with which relevant risk-related behaviours manifest across time and situations (Olver & Wong, 2011). However, one problem within the custodial setting is that the context in which offenders are assessed is significantly different from the context of the initial occurrence of the offending

behaviour and the potential recurrence of the behaviour (Jones, 2010); for instance, the preferred victim type may not be available to the offender in custody (McDougall et al., 1995). Additionally, problematic behaviours that are observable in the community may be considered absent in custody because they manifest in different ways in this alternative setting (Gordon & Wong, 2010). To avoid this potential for neglecting relevant behaviours in custody due to their alternative expression, assessors must have insight into the potential manifestations. According to Daffern (2010), an additional consideration is that the custodial environment might trigger antisocial behaviour that is interpreted to be a manifestation of risk, but is an artefact of the environment (e.g., anti-authoritarian attitudes and interpersonal violence may be highly valued in some custodial settings). Therefore, in addition to the potential for missing relevant behaviours, Mann, Thornton, et al. (2010) noted that it is important that assessors are aware of the potential for over-interpreting behaviours as indicators of risk.

Furthermore, some manifestations of dynamic risk factors are more easily observed by staff than are other behaviours; for instance, self-regulation deficits, which may manifest through behaviour such as verbal or physical aggression, and isolating from others. Therefore, observation of these behaviours might be more frequent than other, less observable, behavioural manifestations (Hanson & Harris, 2001). As described in Chapter 1, staff observations of sexual self-regulation, including sexual deviance and sexual preoccupation, might be less common due to related behaviours being less public. An additional complexity is that some rating criteria are based on behaviour that is more likely in the community; for instance, it might be difficult to assess sexual compulsivity when offenders have limited opportunity to engage in promiscuous behaviour in custody (Olver et al., 2007).

Further, manifestations of these risk factors are often evidenced through attitudes and beliefs rather than overt behaviours. It can be difficult to identify and measure the extent to which attitudes or beliefs are present unless offenders make statements that are truly indicative of these attitudes. Although behaviour is generally an outcome of related thoughts, offenders may make proclamations that are contrary to their beliefs due to a wish to impress in a socially desirable manner and they may behave in a manner that differs from these proclamations (Mann, Hanson, et al., 2010). Such measurement difficulties might decrease clinicians' ability to assess the presence or relevance of behavioural manifestations of certain risk factors (Webster, Müller-Isberner, & Fransson, 2002). If staff members (e.g., psychologists, custodial officers)

infer an offender's attitudes based on their comments, it might result in more subjective assessments, which could affect their reliability and validity.

Daffern and Ogloff (2017) noted that further research into the ways in which dynamic risk factors manifest in custody is crucial to facilitate valid risk assessments. There are various theories that consider this issue and attempt to provide insight into the conceptualisation of manifest risk. Two primary frameworks that have been developed are the offence paralleling behaviour and offence analogue behaviour frameworks. These frameworks were discussed in section 1.1.10.2 (p. 38) and will now be revisited.

2.1.1.1.1 *Offence paralleling behaviour.* As described in section 1.1.10.2.1 (p. 38), Jones (1997) conceptualised a methodology for identifying behaviours within custody that relate to offence processes, referring to these behavioural sequences as offence paralleling behaviour (OPB). The OPB framework has various theoretical underpinnings. It has a case formulation approach, and is consistent with theories such as: interpersonal circumplex (Blackburn, 1990; as cited by Jones, 2004); attachment theory framework; and, the typology of roles described by narratologists.

Within Daffern et al.'s (2007) definition, OPB referred to sequences of behaviour that include behaviours, beliefs, and affects, which are functionally similar to those sequences present in prior offences. It should be noted that OPB refers to both overt behaviour and internal processes such as attitudes and beliefs, all of which can be manifestations of dynamic risk factors. Importantly, Jones (2004) provided explanations of the OPB framework. For instance, the OPB framework is based on the notion that OPB refers to a culmination of a process or chain of events, rather than one single event. In order to develop a formulation of an individual's offence paralleling behaviour, one strategy is to first systematically develop a cognitive, affective and behavioural protocol. This protocol involves engaging in functional analysis to identify sets of behaviours that may have a similar function or developmental structure for an individual, using sequences of discrete episodes to determine the antecedents to the behaviour (i.e., to the sexual offence).

Jones (2004) identified that offence paralleling behaviour and offending processes often represent an escalating sequence of unsuccessful attempts to solve interpersonal problems, which leads to more extreme attempts to gain the desired outcome. Therefore, he suggested that if the function of the offending behaviour is to meet specific interpersonal needs, then behavioural observations within the custodial environment can

be based on insight into these same interpersonal needs, such that offence paralleling behaviour can be identified and understood.

2.1.1.1.2 Offence analogue behaviour. Similarly, Gordon and Wong (2010) use the term offence analogue behaviours (OAB) to represent behaviours in a custodial setting that indicate manifestations of dynamic risk factors. These OABs are generally idiosyncratic to the individual and result from an interaction between their criminogenic needs and the environment (Gordon & Wong, 2010).

As Gordon and Wong (2010) asserted, there are differences in the theories underlying the OPB and OAB frameworks. As noted previously, the OPB framework is based on the case formulation approach. On the other hand, the OAB framework is based on the RNR principles. Gordon and Wong stated that “OABs are explicitly anchored to the theoretical underpinning of the PCC [Psychology of Criminal Conduct] and the principles of effective correctional treatment, that is, the risk, need and responsivity principles” (p. 174). They noted that the PCC essentially provides a theoretical basis for the identification of common criminogenic factors that can be targeted within treatment. They further indicated that since they define OABs as the here-and-now manifestations of the offenders’ criminogenic needs, OABs should also be the “logical theoretical extensions of criminogenic needs” (p. 175) based on the RNR framework. Gordon and Wong highlighted that due to the large differences in the underlying theories behind the OPB and OAB frameworks, they prefer to use the term OAB rather than OPB to describe the RNR principles and risk reduction treatment-based conceptualisations of behaviours that are equivalent to the pattern of offending behaviours.

Validation research on the offence analogue and offence reduction behaviour rating guide (Gordon & Wong, 2009) has shown that behavioural manifestations of dynamic risk factors in incarcerated violent offenders are identifiable in a custodial context and that some are related to recidivism. Specifically, Mooney and Daffern (2013) highlighted the importance of such a behaviour guide to increase staff awareness of behaviours that might be indicative of reduced or ongoing risk of recidivism.

2.1.1.1.3 Positive behaviour. Of relevance to the OPB and OAB frameworks is Andrews et al.’s (2011) assertion that alternative cognitions, emotions and behaviours can counteract an individual’s risk factors. Rather than focusing only on manifestations of dynamic risk, the development and measurement of positive behaviour through

intervention is similarly important in determining an offender's current risk level (see Rogers, 2000).

As described in section 1.1.9 (p. 27), the definition, operationalisation and assessment of protective factors have begun to receive more attention in recent years. It has been argued that the clinical utility of some risk assessment measures is limited due to the low number of factors that reveal improvement during treatment (de Vries Robbé et al., 2015). There is evidence to suggest that the assessment of protective factors in combination with risk factors, provides greater depth and accuracy to risk and change assessment (de Vries Robbé et al., 2015; O'Shea et al., 2016).

The same challenge that continues to be tackled in relation to risk factors is similarly relevant for protective factors. That is, how these factors manifest in offenders such that the presence of positive attitudes and behaviours can be measured. Attempts to resolve this challenge can similarly be seen through the OPB and OAB frameworks. Specifically, these prosocial behavioural manifestations have been termed prosocial alternative behaviours (PAB) by Daffern and colleagues (2007), and offence reduction behaviours (ORB) by Gordon and Wong (2010). A decrease in risk level is not only measured by a decrease in negative behavioural manifestations of risk factors, but also by an increase in prosocial equivalent behaviours (Daffern et al., 2007). Using the OAB/ORB rating guide, Mooney and Daffern (2013) demonstrated that prosocial skills were of greater predictive value for violent recidivism than negative behaviours.

2.1.2 Behavioural Monitoring in Custody

Daffern and Ogloff (2017) asserted that since there has been limited research into how dynamic risk factors manifest in secure environments, there are questions about the validity of assessments of OPB and OAB, and the dynamic risk factors that are assessed within custodial risk assessments. They suggested that the limited research into these behavioural manifestations in custody highlights the importance of further research such that assessors can more confidently identify dynamic risk factors and more weight can be attributed to their assessments.

McDougall et al. (1995) noted that while ideally, psychologists might seek to conduct assessments through use of functional analyses, the small number of professionals and the large number of clients who require assessments, limits the capacity for such assessments. Therefore, it might instead be more practical to assess the presence of common risk-related behaviours, rather than those specifically identified from the functional analysis of an individual offender's behaviour. McDougall et al.

(1995) indicated that it might be easier for staff to identify frequent behaviours, whereas infrequent behaviour is more difficult to assess. They suggested that the first stage of assessment for frequent behaviour would be identification of the behaviours by operational staff (e.g., custodial officers). Infrequent behaviour may require greater understanding of the underlying patterns of risk-related behaviour.

While custodial officers may be in the best position to observe offenders' behaviour, their observations are not necessarily focused on behaviour that is directly relevant to risk (McDougall et al., 1995). They may require guidance in relation to the behaviours they should specifically monitor. Within McDougall et al.'s (1995) research, custodial officers who had previously been trained in using a risk assessment model that required behavioural monitoring, were asked to generate a list of behaviours that offenders might display, which may be targeted for change in intervention. They were asked to make general behavioural descriptions and to consider only behaviours that are easily monitored. Subsequently, psychologists' assistance was sought to refine the lists of identified behaviours. The behaviours most commonly suggested were retained. Some behaviours were similar and could be combined. Although the behavioural manifestations generated for these checklists were not identified as offence paralleling or offence analogue behaviours, they likely met the definition for OABs, as they were behavioural manifestations of risk factors. However, within the description of the research, it was not specified whether the behaviours were based on empirically-determined risk factors.

In summary, assessment of OPB and PAB, and OAB and ORB, involves conducting structured assessments to reduce observer bias and use of accidental observations (Daffern et al., 2007). In a custodial environment, custodial staff have the capacity to make important behavioural observations that contribute to the risk assessment process. Clarke et al. (1993) described a risk assessment methodology involving custodial officers' monitoring of offenders' behaviour patterns throughout their sentence. The research revealed that custodial officers could often recognise behaviours related to offending within the custodial context. Since custodial officers are not ordinarily trained in this complex assessment task, it may be useful for staff to focus their attention on a list of behaviours that more broadly indicate the presence of common dynamic risk factors (McDougall et al., 1995). The first step in the process of behavioural monitoring is to identify potential behavioural manifestations.

2.1.3 Current Study

Although research has been conducted in relation to the manifestation of violent offenders' criminogenic needs in custody (see Mooney and Daffern, 2013), there has been minimal research of a similar nature in relation to the behaviour of incarcerated sexual offenders and the manifestation of relevant dynamic risk factors in custody. The practice of risk assessment based on monitoring behavioural manifestations of dynamic risk factors in custody is therefore limited. In this study, psychologists with experience working with incarcerated sexual offenders were asked how risk factors derived from established structured risk assessment measures manifest in custody.

The method used in the current study resembled that described previously, which McDougall et al. (1995) used. In their study, the custodial officers who were asked to identify potential behaviours they could monitor in custody, had previously been trained in relation to behavioural manifestations of risk factors. Psychologists refined the lists, which resulted in the retention of common items, combining similar items, and redefining very specific behaviours such that they would be more applicable to all offenders. Within the current study, custodial officers did not have the background knowledge required to generate risk-related behaviours, as they had not received training in this area. Therefore, psychologists were asked to identify predicted behavioural manifestations of risk factors. These behaviours could then be refined through qualitative analyses for the development of a behavioural checklist, which was subsequently implemented in Study 3.

The aim of this study was to investigate how dynamic risk factors for sexual offending and prosocial equivalent behaviours might manifest in a custodial therapeutic community, through conducting a survey of psychologists experienced in working with incarcerated sexual offenders. It was hypothesised that sexual offenders' manifest dynamic risk factors for sexual offending and prosocial equivalent behaviours would be recognisable to psychologists, who would be able to provide examples of expected behaviours.

Note that within section 2.2, reference is made to OPB and PAB. The survey was developed with the intention that psychologists would identify behavioural manifestations of risk factors in the form of OPB and PAB. However, since the time this study was conducted and the related journal article was published (Sweller, Daffern, & Warren, 2016; Appendix A), further discussion has been entered into within the research field such that there is increased awareness of the differences between

behaviour labels and the importance of appropriately defining the behavioural manifestations of risk factors, and therefore, distinguishing between OPB and OAB.

As Daffern and Ogloff (2017) highlighted, assessors must understand ‘normal’ or ‘typical’ behaviour in custody, in addition to the typical manifestations of dynamic risk factors in this environment. It is apparent that the behavioural manifestations that psychologists identified within this study represent these ‘typical’ behavioural manifestations. As a result, the conceptual framework for this study has been re-defined. Rather than the behaviours that participants were asked to identify being considered OPB and PAB, they would be more appropriately considered to represent OAB and ORB. The primary difference is that the behaviours identified did not represent sequences of behaviour specifically related to offenders’ sexual offences. Rather, the behaviours reflected discrete behavioural manifestations of empirically-derived risk factors.

2.2 Method

2.2.1 Research Design

This qualitative study involved a structured survey completed by psychologists with experience working with sexual offenders. Participants were asked to identify whether they were currently working with sexual offenders or had previously worked with sexual offenders. They were also required to indicate the length of their experience. The survey comprised open-ended questions, in order to capture the diversity of responses. Survey responses were analysed using thematic analysis, a process by which qualitative data can be examined and organised into themes (Braun & Clarke, 2013).

2.2.2 Participants

In total, 120 participants commenced the survey. Of these, 86 were classified as “partial completers” because they completed less than 60% of the risk factor questions. One reason participants provided for withdrawing from the survey was the time it took them to provide responses. These participants were excluded from the data analysis. As a result, 34 participants were included in the sample. All participants were members of Australian and North American organisations or associations that require employees or members to have a minimum level of clinical experience or to be registered as a psychologist. Participants were contacted indirectly by email through their affiliation with national and international professional organisations. These organisations included:

Corrective Services New South Wales; Corrections Victoria; the Australian Psychological Society; and, the Association for the Treatment of Sexual Abusers. The introductory email outlined the purpose of the study and the inclusion criteria (i.e., only those with experience working with offenders convicted of a sexual offence in a custodial setting were able to participate). Of the 34 participants, 30 were working with sexual offenders at the time they completed the survey. There were 10 participants with over 10 years' experience. The majority had at least five years' experience ($n = 27$). The median level of experience was six years.

2.2.3 Materials

The survey was developed using an online survey tool, SurveyMethods. It consisted of 32 questions (see Appendix B). The introductory questions were related to participants' clinical background, i.e., whether they currently or in the past worked with offenders convicted of a sexual offence, and the depth of their knowledge of OPB. A brief explanation of OPB was provided, prior to the commencement of further questions. Each of the subsequent questions was based on an empirically derived risk factor predictive of sexual recidivism. These risk factors were drawn from two risk assessment tools: RSVP and VRS:SO. For each risk factor, participants were required to respond to two parts: (1) provide examples of behaviours they might expect to observe in offenders if the risk factors were active; and, (2) provide examples of behaviours they might expect to observe if these risk factors were no longer active and had been replaced with more adaptive prosocial behaviour.

An example of a survey question based on a risk factor is as follows:

8. Problems with stress or coping (i.e., extent to which the person's psychosocial adjustment is unstable or susceptible to external events and occurrences).

What are some examples of behaviours/attitudes you might observe if this risk factor manifests in custody?

What are some examples of behaviours/attitudes you might observe if the offender has made positive changes and no longer manifests this risk factor?

2.2.4 Procedure

An introductory email was sent to potential participants. The study was explained and the inclusion criteria were outlined. Participants were informed in the email that by clicking on the link to the survey, they were consenting to participate in the study. Participants could withdraw consent at any time until they submitted their responses. The survey was anonymous, so consent could not be withdrawn after completion.

2.2.5 Data Analysis

The survey results were imported into a qualitative data management program, QSR NVivo (Version 10). A database was created in this program and data were analysed using a six-step thematic analysis approach commonly used in psychological research (Braun & Clarke, 2006, 2013). Repeated readings of the participants' responses to the open-ended survey questions allowed initial and ongoing *data familiarisation* (step 1). For each participant's response for each survey question, potential emergent codes were noted. This involved assigning a representative or summative word or phrase (a 'code') to a passage of the data ('*forming codes*'; step 2). Responses were coded systematically and initial codes were *categorised* into potential themes and sub-themes on the basis of any patterns identified (step 3). A *theme was identified* for the analysis (step 4) by noting its frequency and/or its relevance for the research questions. In this way, themes were refined to ensure that each had sufficient and meaningful supporting data. This process was largely inductive, drawing on the responses of participants; however, deductive themes were also developed based on the candidate's knowledge of relevant risk factors and behaviour that sexual offenders tend to exhibit as they progress through a treatment program (see Markovic, 2006). The use of deduction was considered appropriate given the applied purpose of the research (Pope, Ziebland, & Mays, 2000).

Each behaviour provided in the survey responses could be coded into one or more of the themes. The number of themes created for each risk factor was flexible and was not pre-determined. Themes were collapsed into a manageable number while maintaining their discriminative validity. Each negative and positive behavioural manifestation of each risk factor comprised of between eight and 15 themes. Themes were then *defined and named* (step 5) in the context of the broader project (Braun & Clarke, 2006). At this point, frequencies were obtained for each theme. An inter-rater reliability check was conducted on this data. A researcher with no experience of forensic psychology or sexual offenders' behaviour undertook an independent coding of responses for 10% of the data (see Liamputtong, 2012). Where there were discrepancies in the coding between coders, discussion was undertaken and codes were redefined. This process was used to demonstrate that the findings extended beyond the subjective judgements of the researchers (Pope et al., 2000), ensuring the rigour of the analysis. The outcome reflected reliability within the themes, which are *reported* (step 6) in this chapter (Braun & Clarke, 2006, 2013).

2.3 Results

2.3.1 Descriptive Statistics

At the time of the survey, 85% ($n = 29$) of participants were currently working with sexual offenders and 15% ($n = 5$) had worked with sexual offenders previously. Levels of experience working with this client group varied from between one and 20 years. Eighty two per cent of participants ($n = 28$) reported that they were familiar with the concept of OPB.

Participants' responses for risk-related behaviours were generally more descriptive than responses for positive behaviours. For instance, when providing examples of behaviour that might indicate a risk factor is no longer problematic, there were suggestions of 'opposite of above' or 'absence of above'. The interpretation of these comments would have been subjective; therefore, they were coded into a separate theme.

As can be seen in Tables 1 and 2, there was some overlap between risk factors and examples of behavioural manifestations. That is, there were several behavioural manifestations that were provided, which were also risk factors. For example, Problems with Substance Abuse was itself a risk factor, but was also described as a manifestation of other risk factors (e.g., Problems with Stress or Coping). There were also behavioural manifestations that were suggested for several risk factors (e.g., 'sexual comments to staff' was suggested for risk factors including Sexual Compulsivity and Attitudes that Condone Sexual Violence).

For some risk factors, participants provided a greater number of relevant behavioural manifestations than for other risk factors. For instance, the overall number of predicted behavioural manifestations was higher for the risk factor Problems with Substance Abuse (100). Conversely, although the overall number of predicted behavioural manifestations for the risk factor Sexual Deviance was 97, of these responses, 21 were coded as 'broad or questionable relevance' (see Table 1). There was also a difference in the quality of the responses, such that for some risk factors, the examples provided were more specific and observable behaviours (e.g., Problems with Stress or Coping; Problems with Substance Abuse; and, Interpersonal Aggression). Responses that were less specific and observable were commonly reported for risk factors such as: Sexual Deviance; Attitudes that Condone Sexual Violence; and, Problems with Self-Awareness. Further, there were several behavioural manifestations that were uncommon and appeared irrelevant to the risk factor in question (e.g., one

participant identified a behavioural manifestation of substance use as ‘placement issues’). These responses were coded into a separate theme (i.e., ‘broad or questionable relevance’).

Risk factors related to sexual self-regulation (e.g., Sexual Compulsivity, Sexual Deviance) elicited responses that were commonly based on self-report (e.g., excessive masturbation). Observable behaviour was also described; however, this behaviour would likely be observed only in the most extreme cases (e.g., public masturbation). In addition, such extreme behaviour would often be classified as a sexual offence, rather than a risk factor.

Table 1. Risk Factors and Themes of Negative Behavioural Manifestation Examples, with Examples of Behaviours within Themes

Risk Factor	Number of different themes of behaviour identified	Most frequently reported themes	Example/s of behaviour in the theme	Second most frequently reported themes	Example/s of behaviour in the theme	Third most frequently reported themes	Example/s of behaviour in the theme
Attitudes that support sexual violence	10	Lack of empathy/victim blaming = Not accepting responsibility = Broad / questionable relevance	Joking about victims, blaming victims = Justifying behaviour = Poor attitude	Supporting others' unhelpful attitudes	Normalisation of sexual violence, colluding with others	Sexualised or sexually aggressive comments	Sexualised comments to staff
Problems related to child abuse	14	Fear or mistrust	Hyper-vigilant, refusal to discuss childhood	Attachment issues	Poor boundaries with others, keeps physical distance	Poor coping	Isolation, substance abuse, feelings of worthlessness
Problems with stress or coping	13	Physical or verbal aggression	Loss of temper, angry outbursts, irritability in interactions, fighting	Poor problem solving	Failing to keep appointments, catastrophic thinking	Broad / questionable relevance	Deterioration in behavioural case notes, control issues
Problems with substance abuse	15	Substance use	Glazed eyes, positive urinalysis	Drug seeking behaviour	Telling the Clinic they will self harm if they do not get medication	Emotional dysregulation = Unhelpful attitude to substance use	Calm one day and aggressive the next = Psychological arousal to

Problems with treatment	10	Disengagement	Sabotages process, refusal to complete treatment work	Not taking responsibility for treatment	Focusing on others rather than self	Therapy interfering behaviour	discussions about drugs Oppositional attitude, disruptive, distraction to avoid topics
Problems with planning	11	Poor release planning	Unrealistic pre-release plans	Poor problem solving	Impulsivity, leaving requests to the last minute	Difficulty engaging in routine	Late for muster, not doing homework
Problems with supervision	9	Not abiding by rules	Institutional misconducts, fighting against rules	Lack of regard for authority/system	Lies to staff, hostility towards staff	Not adhering to staff requests	Refusing orders
Problems with employment	9	Refusing to work or avoiding = Unable to maintain a job	Calling in sick, refusing to do training = Frequently loses job in custody	Lack of motivation	No goals, loses interest	Not taking responsibility	Sense of entitlement, saying the system or government owes them
Problems with intimate relationships	10	Communication difficulties	Aggressive phones calls with partner	Poor attachments	Superficial connections, denying the need for others	Avoidance	Isolation, withdrawal from others
Problems with non-intimate relationships	12	Withdrawal or isolation	Avoidance of supports	Difficulties forming and maintaining relationships	Not communicating with more than one other inmate	Aggression	Interpersonal conflict, fighting

Interpersonal aggression	10	Intimidation	Intensive eye contact	Verbal aggression	Raised voice to express hostility	Physical aggression	Violence
Offence planning	9	Manipulating others	Setting up relationships of obligation	Reporting plans to offend	Expressing offence fantasy	Victim contact	Contacts previous victims, makes connections with others to facilitate offending
Extreme minimisation or denial	13	Not accepting responsibility	Downplaying significance of harm to victims, saying they cannot remember	Resistance to treatment	Anger at discussing offence, refusal to participate in treatment	Blame shifting	Blaming others
Psychopathic personality disorder	14	Manipulation	Playing staff or inmates off against each other	Superficial	Appears to comply but undertones of non-compliance	Lack of empathy	Delight in others' discomfort in group
Mental illness	15	Psychotic symptoms	Talking to themselves	Broad / questionable relevance	Medical records, acting crazy	Mood states	Excessive crying
Non-sexual criminality	12	Institutional misconduct	Theft, rule breaking	Pro-criminal attitudes	Joking about criminality	Aggression	Fights, standovers
Physical coercion	11	Physical or verbal aggression	Direct threats to safety	Sexual aggression	Preying on younger inmates	Broad / questionable relevance	Creepy behaviour
Psychological coercion	7	Manipulating for favours	Buying others food in return	Stand overs / intimidation	Social shunning, standovers	Broad / questionable relevance	Sexual offences in custody, poor

Problems with self awareness	11	Broad / questionable relevance	for going two-out Use of drugs or alcohol, poor coping	Not accepting responsibility	Denial that behaviour requires addressing	Not understanding triggers	impulse control Talking about returning to old lifestyle
Sexual compulsivity	10	Sexualised comments	Sexual content in conversations, sexual comments to staff	Sexual behaviour in public	Being caught masturbating, talking and acting in sexually explicit ways	Excessive sexual behaviour	Daily masturbation, sex with other prisoners
Sexual deviance	12	Broad / questionable relevance	Watching sexual TV, excessive masturbation	Inappropriate images	Collecting pictures from children's clothing catalogues	Inappropriate sexual arousal	Watching afternoon children's TV programs
Violent or suicidal ideation	9	Threats or preparation	Weapon-creating	Emotional instability	Irritability, sudden change in presentation, tearfulness	No future focus = Broad / questionable relevance	Giving away possessions, aggressive acting out without regard for personal safety = Periods in segregated custody

Note. If more than one theme includes the same frequency of responses, it is denoted by an '='.

Table 2. Risk factors and Themes of Positive Behavioural Manifestation Examples, with Examples of Behaviours within Themes

Risk Factor	Number of different themes of behaviour identified	Most frequently reported themes	Example/s of behaviour in the theme	Second most frequently reported themes	Example/s of behaviour in the theme	Third most frequently reported themes	Example/s of behaviour in the theme
Attitudes that support sexual violence	11	Broad / questionable relevance	Using helpful coping strategies, maturity	Challenging unhelpful attitudes	Challenging others about sex offending, encouraging others to participate in treatment	Victim empathy	Recognition of harm to the victim
Problems related to child abuse	14	Appropriate attachments	Understanding / respect for personal boundaries	Greater self awareness	Identifying how abuse history contributed to misbehaviour	Engaging in intervention = Good coping	Willingness to address issues = Discuss emotional reactions to stimuli
Problems with stress or coping	11	Using supports	Seeking out assistance from others	Broad / questionable relevance = Good problem solving	Coping skills, consistency in behavioural case notes = Positive attitude to a challenge, calmer	Appropriate use of services = Manage unhelpful thoughts	Keeping appointments with psychology = Withdraw from anger-provoking situations, perspective taking

Problems with substance abuse	11	Engaging in intervention = Negative urinalysis	Compliance with methadone program	Helpful attitudes about drug use	Discusses problems caused in the past	Risk management	Strategies for managing cravings
Problems with treatment	9	Engagement	Listens and contributes	Takes responsibility for treatment	Addresses own needs, takes advice, active participation	Discusses relevant issues in depth	Gives answers based on their own circumstances, volunteers information
Problems with planning	12	Release planning	Communication with outside supports	Good problem solving	Identify thought pattern before behaviour	Time management	Planning activities, taking control of their own timetable
Problems with supervision	9	Compliance with rules	Complies with directions from staff	Engagement	Reliable with tasks, attendance at case meetings	Positive attitudes to supervision	Post release plans involve continued treatment and supervision
Problems with employment	10	Regular work attendance	Stable employment, establishes and maintains engagement	Positive feedback from work	Good work reports	Positive attitude to work = Seeking training	Pride in job = Active participation in vocational training and education
Problems with intimate relationships	10	Taking responsibility in relationships	Making contact with loved ones, engagement of partner in	Positive interactions and relationships = Stability in relationships	Positive interactions with family = Desire to maintain relationships	Broad / questionable relevance = Valuing relationships	Pattern of behaviour over time, reflection = creating an equal partnership

Problems with non-intimate relationships	9	Interacting with others	release planning Mixing with others, works in a team	Stable and meaningful relationships	Developing and maintaining relationships	Assertive communication with staff	Assertive communication with staff
Interpersonal aggression	11	Appropriate interactions	Calm and able to tolerate disagreement, listen calmly	Assertive communication = Good coping	Open, discussing concerns in a respectful way = When not get what they want, speak about it in a non-threatening way, self calming	Absence of behaviour	Reduction or absence of above, no fights or assaults
Offence planning	9	Accepts responsibility for offending	Acknowledges intent, implements relapse strategies	Appropriate interactions and relationships	Encouraging peers to be compliant with rules	Risk management	Challenging distortions, accepts urges may continue but will pass
Extreme minimisation or denial	11	Accepting responsibility for behaviour	Accepts own behaviour and feedback without defensiveness	Treatment engagement	Challenge of self and others through feedback, honesty	Empathy	Articulates how behaviour might have affected others
Psychopathic personality disorder	12	Absence of certain behaviours	Not callous or cynical	Appropriate relationships = Empathy =	Maintaining appropriate boundaries = Kindness	Self reflection and insight	Acknowledging emotional shallowness

			Treatment engagement	when not needed =	
Mental illness	12	Medication compliance	Regularly taking medication	Stable mood, more spontaneous in responding	Implement management strategies
Non-sexual criminality	10	Abide by rules = Absence of certain behaviours	Positive self regulation, following rules = Absence of incidents or accusations	Access services and planning for future without crime = Improved attitudes to law	Broad / questionable relevance More positive case notes
Physical coercion	15	Absence of behaviour = Respectful interactions	Absence of threats = Engaging respectfully with staff and inmates	Efforts made to genuinely help others = Talking about problem issues, challenge cognitive distortions	Broad / questionable relevance Engaging in work preparation
Psychological coercion	10	Respectful interactions	Humility and openness	Lack of manipulation	Assertive communication = Broad / questionable relevance Openness in communication = non-coercive

Problems with self awareness	12	Accepting responsibility for behaviour = Realistic risk management = Treatment engagement	Removal of self from risky situations, speaks about actions leading up to offence = Scenario-planning = Identification of risk factors and participation in treatment	Understanding triggers	Verbalizing own risk factors	Broad / questionable relevance	Good manners, improved coping
Sexual compulsivity	10	Absence of behaviour	No sexually inappropriate behaviour in custody	Control of sexual behaviour	Delay gratification, sitting with discomfort	Insight into behaviour and intervenes	Paying attention to thoughts, arousal management techniques
Sexual deviance	9	Treatment engagement	Openness to discuss difficulties in the area	Coping strategies to manage deviance	Open about sexual thoughts, replace deviant fantasy	Absence of behaviour	Reduction of these behaviour, absence of above
Violent or suicidal ideation	11	Reduction in threats or acts	Fewer attempts, no talk of violence	Engagement in intervention	Compliance with medication, distress tolerance, seeking help	Appropriate engagement with others = Future focused	Assertive communication, cooperative with others = Consequential thinking,

	talking of hopeful future
--	------------------------------

Note. If more than one theme includes the same frequency of responses, it is denoted by an ‘=’.

2.4 Discussion

The aim of this study was to identify predicted behavioural manifestations of sexual offenders' dynamic risk factors in a custodial environment. The identification of these behaviours may assist clinicians assess risk, monitor change during lengthy periods of incarceration and identify opportunities for timely intervention (i.e., intervening when a risk-related behaviour is present and active). Furthermore, the survey illuminated the wide range of factors that assessors focus on when appraising the presence of dynamic risk factors embedded in structured risk assessment instruments.

2.4.1 Overview of Results

The results not only provided potential behavioural manifestations of incarcerated sexual offenders' dynamic risk factors, but also revealed the potential challenges that might arise within the process of identifying these manifestations. Overall, there were some interesting results that emerged from this study, which will now be briefly identified prior to a more in-depth discussion. The first of these results was that participants appeared to identify problematic behaviours with greater ease than positive behaviours. Second, there were differences between risk factors in the numbers of behavioural manifestations participants could generate. Relatedly, within some risk factors, participants identified behavioural manifestations that appeared unrelated to the risk factor, which might have been a reflection of the difficulties experienced in predicting the behaviours. Finally, there was overlap between risk factors and behavioural manifestations, such that some behavioural manifestations participants identified were themselves risk factors. These primary results will now be further explored and interpreted.

2.4.2 Exploration of Results

The following subsections explore the results in more detail and provide potential explanations and interpretations.

2.4.2.1 Risk vs prosocial behaviours. Based on the responses provided, participants more readily identified manifestations of dynamic risk factors compared to examples of prosocial equivalent behaviour. Responses were generally more common, detailed and specific in relation to the manifest dynamic risk factors. This difficulty in identifying behaviour related to positive behaviour and behaviour change might reflect a range of issues: psychologists are trained in the use of risk assessment tools, which focus on risk-related factors (de Vries Robbé, de Vogel, & de Spa, 2011; Miller, 2006; Rogers, 2000; Sheldrick, 1999); the culture of a custodial environment is based on

punishment or consequences for antisocial behaviour; and, when clinicians discuss clients with custodial staff, more often than not, the focus is on observed difficulties or problematic behaviour. As a result, staff in this environment may focus on risk-related behaviour rather than positive behaviour.

This finding is consistent with prior research on the offence analogue and offence reduction behaviour rating guide, which has shown that staff rarely document prosocial behaviour (Mooney & Daffern, 2013). This emphasis on risk factors is likely to contribute to “professional negativism” and could result in a negative bias against clients (Rogers, 2000). Rather than limiting assessment of risk to a determination of the absence of problematic behaviour, it is important to consider behaviour improvement. That is, consistent with research showing the benefits of protective factors (e.g., de Vries Robbé et al., 2011), the development of prosocial skills that replace, or reflect improvement in a dynamic risk domain, may also be relevant to the assessment of risk for recidivism (Mooney & Daffern, 2013).

An alternative explanation for participants’ greater difficulty in identifying prosocial behavioural manifestations might be the wording of the survey question in this study. In relation to the positive behaviours, participants were asked to identify behaviours they might expect to observe if the risk factor was no longer active. As a result, participants might have focused on the elimination of risk rather than a decrease in risk. They may also have been more likely to focus on the absence of negative behaviour rather than the presence of positive behaviour because they may have interpreted the focus of the question to be on the risk factor’s absence. Future research could alter the wording of the survey questions such that participants focus on a decrease in the relevance of the risk factor rather than a complete eradication. The questions could also be rephrased to focus on the development of positive behaviour rather than the absence of negative behaviour. This idea has a foundation in the intervention context, in relation to the use of approach-oriented, rather than avoidance-oriented, intervention. Many psychologists are likely to be more familiar with avoidance-oriented intervention, as it is central to the relapse prevention model that was popular in the past within the psychological field (Mann, Webster, Schofield, & Marshall, 2004). An avoidance orientation is more focused on avoiding risk and implementing avoidance-based strategies to manage risk, than on working towards developing positive skills. It is possible that this focus on avoidance-based strategies

may lend itself towards psychologists' increased tendency to identify risk-related behaviours rather than positive behaviours.

2.4.2.2 Variation between risk factors. For both manifest dynamic risk factors and prosocial alternative behaviour, there was variation between risk factors in the frequency with which participants identified relevant behaviour. For instance, a greater number of relevant behavioural manifestations of Interpersonal Aggression were identified than Sexual Deviance. This variation might be related to the difference in visibility of behavioural manifestations of some dynamic risk factors (for instance Sexual Deviance) compared with others (Mooney & Daffern, 2013). Further, within the custodial context, behaviours related to Interpersonal Aggression are likely more common than Sexual Deviance. Although psychologists working with sexual offenders should be familiar with behavioural manifestations of a range of risk factors including sexual deviance, it is likely that it is the more overt behaviour that would form the focus of greater discussion between staff members (e.g., custodial officers and psychologists). This focus might also be related to some behaviours having a greater impact on others, including staff.

Not only was there a difference between risk factors in the frequency with which examples were provided, but further, some suggested examples of behavioural manifestations were uncommon and appeared unrelated to the risk factor in question. This finding may indicate that some psychologists are uncertain about the behaviours that sexual offenders might exhibit when a particular dynamic risk factor is active; alternatively, there may be idiosyncratic manifestations of these risk factors. Ultimately, careful scrutiny of the behaviour is required before determining that it is related to the risk factor and relevant to the individual's offending.

Risk factors that are more internal (i.e., related to thoughts and attitudes) produced more uncommon and seemingly irrelevant responses. This result suggests that when psychologists cannot directly observe the relevant behaviour they produce a broader range of potential risk-related behaviours. This issue has important implications for the scoring of structured risk assessment instruments; if psychologists have difficulty determining whether a behaviour is a valid manifestation of a risk factor, then the reliability with which risk assessment measures are scored may be decreased, increasing the risk of item drift (Webster et al., 2002).

2.4.2.3 Overlap between risk factors and behaviours. The overlap between risk factors and behavioural manifestations, such that behaviours identified as examples of

manifest dynamic risk factors and prosocial equivalents, were themselves risk factors in the risk assessment tools (e.g., substance use) was an unexpected finding. In addition, the same behavioural manifestations were identified within several risk factors. For example, 'not accepting responsibility for behaviour' was identified as relevant for the following risk factors: Problems with Treatment; Substance Use; Psychopathic Personality Disorder; Attitudes that Support or Condone Sexual Violence; Problems with Employment; Extreme Minimisation or Denial; Non-Sexual Criminality; Offence Planning; and, Problems with Self Awareness. These findings raise questions about the most appropriate way of scoring a risk assessment item; that is, how to determine which risk factors are relevant when a specific behaviour is observed in custody (e.g., if drug use is observed, is it best captured under the risk factor Substance Use or does it also infer Problems with Stress or Coping). Finally, this finding may highlight the strong inter-relationship between dynamic risk factors (Thornton, 2002), such that any behaviour may relate to various risk factors (e.g., using drugs in prison may be a consequence of Problems related to Child Abuse, and Problems with Stress or Coping).

These results highlight an issue that Ward has explored in relation to the definition and properties of dynamic risk factors (Ward, 2016; Ward & Beech, 2014). Ward (2016) questioned whether dynamic risk factors are predictor variables or psychological constructs that refer to causal processes. He indicated that there is confusion surrounding whether these factors refer to aetiological factors, causal processes, or offence-related phenomena. Further, the conflation between surface features, symptoms or problems evident in theories surrounding sexual offending highlights the need to clarify what theorists are attempting to explain (Ward & Beech, 2014). These challenges associated with the form and function of dynamic risk factors provide an explanation for the overlap between risk factors and behavioural manifestations, and between behavioural manifestations across several risk factors. The appropriate classification of the behaviour might depend on the role the risk factors takes (e.g., is 'substance use' the offence-related construct or a symptom).

2.4.2.4 OPB vs OAB. As previously indicated, subsequent to this survey's completion and publication, the ways in which behavioural manifestations of risk factors can be defined and operationalised were further investigated. Consequently, revision was made to the underlying framework of this study, such that the behavioural manifestations were considered and labelled OABs and ORBs rather than OPBs and PABs.

Despite the initial framework representing offence paralleling behaviour, it appeared that participants had difficulty applying the framework to the survey questions. The survey instructions requested participants to identify potential offence paralleling and prosocial alternative behaviours. The majority of participants indicated that they were aware of the OPB theoretical framework. Nevertheless, in contrast to the definitional properties inherent within the OPB framework, the behavioural manifestations that participants suggested tended to encompass individual behaviours, rather than behavioural sequences. The results within this study further highlight the importance of ongoing investigation into the definitional properties of behavioural manifestations of risk factors.

2.4.3 Limitations and Future Research Directions

The first limitation relates to the sample size. While a large number of psychologists partially completed the survey, the final sample was substantially smaller. This rate of partial completion may be an indication that the length and/or difficulty of the survey presented a barrier to its completion. If all 120 psychologists who commenced the survey had completed it, the sample may have been more representative of the population and may have provided the opportunity to compare and contrast the behavioural examples statistically. This process may have permitted exploration of associations between the items, and rationalisation of the inclusion of each item as well as the total number of items on the checklist.

Not all risk factors related to sexual offending are consistently and directly observable in a custodial environment. Therefore, it can be difficult to determine whether these risk factors are active. For instance, offence-supportive attitudes, sexual preoccupation, and sexual deviance, are all predictive of sexual offending (Mann, Hanson, et al., 2010), but tend to be related to thoughts or more covert behaviour. This research highlights the potential difficulties associated with determining whether risk related to sexual self-regulation is current (e.g., sexual deviance and sexual compulsivity). Behaviour related to this risk area is generally reliant on offender self-report. Although participants in the current study appeared to have some understanding of behavioural manifestations of this risk area, the more overt manifestations suggested are likely to occur only in individuals who have significant problems with sexual self-regulation (e.g., excessive masturbation). Participants in the current study were all psychologists. Psychologists may only see offenders in clinical contexts. This context may provide a somewhat limited or even biased account of offenders and their relevant

risk-related behaviours (e.g., offenders may be more careful about what they say, knowing that the psychologist may prepare a risk assessment report that could be used at a forthcoming parole board hearing). This limitation may have influenced psychologists' ability to generate potential observed behaviour in the custodial context. In the future, subject to ethics approval, the results from psychologists could be validated through asking staff from varying disciplines to comment on the suggested behaviours, based on their observations of sexual offenders in other custodial contexts. Triangulation could be used to validate the observational assessment reports made by psychologists.

The limited context in which psychologists interact with offenders has been discussed in previous research. For instance, there has been discussion about the benefits of custodial officers' observations being used to assist with behavioural monitoring and assessment in custody. Gordon and Wong (2010) suggested that behavioural manifestations of risk factors should be monitored by staff in different contexts rather than limiting it to the treatment group room. McDougall et al. (1995) also highlighted the use of custodial officers' observations through completion of checklists about offenders with whom they had sufficient contact. Atkinson and Mann (2012) discussed the behaviours custodial officers reported observing in custody and noted the potential use for this process. The potential benefits to engaging custodial officers within the process of behavioural monitoring provide further rationale for the use of the current study's results in the development of a behavioural checklist, which custodial officers could use. This idea will now be further explored.

Prior research has demonstrated that behavioural manifestations of dynamic risk factors and the opposing prosocial behaviour can be reliably identified in incarcerated violent offenders through the use of a rating guide that identifies relevant behaviour from the VRS (Mooney & Daffern, 2013). Similarly, it is important in a custodial environment for inexperienced staff to have some understanding of the common behavioural manifestations of risk-related and positive behaviour. These behaviours could be monitored to assist risk assessment and change measurement. Although risk may not be eliminated, it can be reduced through interventions including psychological treatment. The observations that participants provided in the current research can be used as examples of behaviours indicative of ongoing risk or prosocial change in important risk-related domains. In addition, custodial staff who may not be familiar with the relevant risk factors specific to sexual offenders, could benefit from access to

information about these manifestations to help their monitoring and management of offenders.

Study 3 within the current program of research included a trial of such a checklist of potential risk-related behaviours in a treatment program for sexual offenders, based on the results of this study. Weekly completion of this checklist by both offenders and custodial staff could provide information about changes in the frequency of various risk-related behaviours over the course of treatment. This form of monitoring could be used to assist with the completion of risk assessment and change measurement. A longer-term project could validate the checklist through a recidivism study that could evaluate whether changes in these behaviours (reductions in manifest dynamic risk factors and increases in the frequency of positive behaviour) over the course of treatment are related to changes in recidivism rates.

2.4.4 Conclusions

There are several practical and theoretical implications for the results of the current study, and two important conclusions. The first conclusion is that participants were able to identify behavioural manifestations that they considered relevant for each risk factor. Participants provided important information about the behaviours they expect to observe in incarcerated sexual offenders. The next step in furthering this field of research is to test the relationship between these behaviours and sexual offending following release into the community. One potential method may be to conduct a prospective evaluation to determine whether the most frequently observed behavioural manifestations are related to recidivism. That is, the behaviours identified in the current study have face validity, but further research is required to demonstrate their predictive validity. Such research could provide further insight into the behavioural manifestations of risk factors, which could facilitate more effective risk assessments, as Daffern and Ogloff (2017) asserted is important within this field. Not only could risk assessments benefit from this increased understanding of manifest risk, but so too could assessments of behavioural change over time.

The second conclusion relates to the differential ease with which behavioural manifestations were identified for different risk factors. This finding has implications for risk assessment. The finding that some identified behavioural manifestations were vague, or alternatively, more applicable to another risk factor, might suggest that at times, clinicians working with sexual offenders could improperly identify that a behavioural manifestation is an indicator of a particular risk factor when it might in fact

be an indicator for a different risk factor. This issue is relevant for the scoring of risk assessments; it might lead to an inflated estimation of risk if one behaviour is considered indicative of a range of risk factors. Another important implication of the current research is related to the finding that participants had more difficulty identifying prosocial behaviour than they did identifying problematic behaviour. Clinicians should be encouraged to monitor and discuss positive behaviour observed in offenders, to enhance the risk assessment and treatment process.

Given the complexity of the risk assessment process, it might be beneficial to present information regarding common manifestations of risk factors to less experienced psychologists and other staff (e.g., custodial staff) to assist in the identification of relevant behavioural manifestations for each offender. They would then be better equipped to understand the relevance of specific behaviours to dynamic risk items in structured risk assessment measures for each offender and use this information to assess change and risk. Additionally, it has been highlighted that clinicians need to accurately determine the risk areas that are active when particular behaviours are observed, through the process of functional analysis (e.g., whether aggressive communication is related to coping or interpersonal difficulties). This use of functional analysis may help determine whether the behaviour is relevant to a particular risk factor. With these potential modifications to the ways in which risk assessment tools are used, the use of structured risk assessment tools may become more reliable and item drift may decrease.

Chapter 3: Study 2

Findings from the first study revealed that offenders' behavioural manifestations of risk factors in custody can be identified, with some behaviours more easily monitored than others. In particular, positive behaviours appeared more challenging than negative behaviours, for clinicians to identify. Behavioural monitoring is crucial in order to gain greater insight into offenders' change processes throughout their incarceration. Determining the ways in which an offender has changed throughout a period of incarceration can assist with decisions such as classification progression, treatment progress, outstanding needs and granting parole. Therefore, it is crucial to further investigate and develop valid measurement tools for assessing changes in offenders' attitudes and behaviour.

This Introduction to Study 2 will further explore issues that were raised in Chapters 1 and 2. The focus will be the use of behaviour in custody, in particular, institutional behaviour such as misconducts, to evaluate offenders' overall behaviour. Prior research into the use of institutional behaviour in decision-making will be discussed, followed by additional exploration of the measurement of behavioural change. The background to the development of the Satisfactory Behaviour Rating Guide will also be provided.

3.1.1 Use of Institutional Behaviour

In order for institutional conduct to be used as a valid consideration within release decision-making, it must be predictive of future behaviour. That is, it must be established that the changes an offender has made in custody are likely to be transferred to the community. As discussed in section 1.1.10.1 (p. 32), empirical research has demonstrated cross-situational behavioural consistency between the custodial and community contexts (e.g., McDougall et al., 1995; Zamble & Porporino, 1990). It has been noted that the propensity for antisocial behaviour traverses varied social situations and institutional misconduct may be used as a representation of antisocial behaviour in the community (Bonham, Janeksela, & Bardo, 1986).

Prior research has focused on the factors that are considered in parole decision-making. Among the factors that parole boards consider, are post-sentencing variables (e.g., Gottfredson, 1979). Many empirical studies exploring parole board decision-making have found that an offender's institutional conduct was significantly associated with release decisions (Caplan, 2007). In the United States, a review of parole board decision-making literature revealed that decisions were primarily based on institutional

behaviour, crime severity, criminal history, incarceration length, mental illness, and victim input (French & Gendreau, 2006), instead of specific risk-relevant behaviours and associated changes while incarcerated. Gottfredson (1979) suggested that in theory, evidence of institutional adjustment, as indicated by compliance with regulations and lack of disciplinary actions, and participation in appropriate treatment programs, would permit a parole board to effectively gauge the offender's prognosis following release from custody. In order for this prediction to be valid, these factors must function as evidence for behavioural change subsequent to the offender's initial incarceration.

Gottfredson (1979) investigated the extent to which institutional behaviour provided information incrementally greater than that provided by the offender's sentence length, in parole decision-making, in a United States Federal prison. The following behavioural factors were considered: number of punishments for violation of rules; assault or threat to assault, resulting in disciplinary action; and, whether there was a record of escape or attempted escape. The results suggested that institutional behaviour may have influenced time served in custody, but the influence was not large. One suggested explanation for the smaller than expected influence was that overall, the behavioural factors included in this study were relatively rare events. The authors also noted this outcome might have been due to the context of a Federal, rather than a State prison, as offenders in State prisons tend to incur a greater frequency of institutional misconducts (Gottfredson, 1979). On the other hand, Bonham et al. (1986) similarly investigated parole release decision-making in Pennsylvania. Results showed that institutional behaviour was the most influential factor, which the parole board used to make judgements about future criminal behaviour. Questions remain regarding the validity of institutional behaviour as a measure of change and predictor of future behaviour. There have been mixed results in prior research investigating the association between institutional behaviour and future behaviour upon release.

Despite the finding that institutional behaviour influences parole decision-making, Carroll, Wiener, Coates, Galegher, and Alibrio (1982) found that after a one-year follow-up, parole board predictions in Pennsylvania had minimal relation to post-release outcomes. The decision-process used by the parole board in Carroll et al.'s research is best regarded as unstructured, which Gottfredson (1979) suggested may have been responsible for the disparity between decisions for offenders who were essentially in similar situations. This process highlights the need for a more structured process by which to determine offenders' change and therefore, suitability for release.

More recent studies have also demonstrated that official misconducts in custody had limited relationship with future reoffending. Trulson, DeLisi, and Marquart (2011) conducted a study in the United States within a “large southern juvenile correctional system” (p. 709). The authors used a variety of official measures for institutional misconduct: total misconducts; staff assaults; youth assaults; danger to others; possession of a weapon; and, gang related activity. Using these measures, there was only a small effect of total misconducts on frequency of arrests following release, with no effect on the binary outcome of whether or not an offender was arrested. They concluded that official misconducts had limited value as a predictor of recidivism. It should be noted that the sample within this study consisted of juvenile offenders rather than adults. It is possible that there is a difference between juveniles and adults in the factors that are predictive of future recidivism. This result was similar to that gained in subsequent research in which misconducts were found to predict recidivism in adult but not juvenile offenders (Cochran, Mears, Bales, & Stewart, 2012).

Contrary to these findings, a number of studies have demonstrated that prison misconducts predict future recidivism (Carroll et al., 1982). To date, the primary focus of research in this area has been the relation between institutional aggression and recidivism post-release (Trulson et al., 2011). Limited research has examined the relevance of aggression in custody to recidivism following release while controlling for violence risk, as measured by contemporary valid risk assessment instruments (Mooney & Daffern, 2011). Despite the limited research, aggressive behaviour in custody, or a pattern of drug use or serious non-compliance within the structured and controlled institutional environment are behaviours considered by parole boards to be indicative of an increased risk of recidivism upon release and compliance with parole conditions (Adult Parole Board of Victoria, 2015), which suggests insufficient behavioural change to warrant release from custody. In order to ensure the most valid assessment measures are used to determine change over time and subsequent risk level, it is important to gain further insight into whether incidents in custody add predictive power to this process.

3.1.2 Change Measurement

3.1.2.1 Importance of change measurement. Change measurement is key within the criminal justice system. It allows for modified predictions of future risk, which is the basis for many parole board decisions. The final decision related to incarcerated offenders is whether to grant release. Prior to release from custody, a determination

must be made about whether the offender in question has made enough changes such that the risk he/she presents to the community has lowered to a satisfactory level.

Parole decision-making has significant consequences for both the offender and the community into which the offender could be released. These decisions are complex and important. Political pressures and resource challenges add to the complexity (Serin, Gobeil, et al., 2016). Therefore, the methods by which change is assessed require scrutiny and ongoing improvement. Ultimately, this decision-making is based on an attempt to predict risk of breaching the conditions of parole and risk of criminal recidivism (e.g., violence to others), which includes evaluation of the changes undergone during offenders' incarceration periods.

3.1.2.2 Institutional misconducts. Whereas static risk measures might not capture changes in the likelihood of offending (Mooney & Daffern, 2011), prison misconducts may provide dynamic information relevant to recent attitudes and behaviour. Specifically, institutional behaviour may provide a useful source of information regarding an offender's response to incarceration and treatment, allowing for an assessment of change over time. Behaviour in custody is also readily available to assessors (Cochran et al., 2012).

In a meta-analysis including 68 studies published between 1952 and 2003, French and Gendreau (2006) assessed whether the results of prison misconduct studies had practical long-term consequences, such that there would be a positive relationship between the degree to which a treatment program reduces misconducts and subsequent recidivism. Treatment program categories were identified as: behavioural (radical behavioural, social learning, cognitive behavioural, or punishment); non-behavioural (nondirective therapy, psychodynamic, group milieu); educational/vocational; and, unspecified. The research revealed the strongest effects for behavioural programs in eliciting changes in frequency of offenders' misconducts. Specifically, there was a 26% reduction in institutional misconducts for behavioural program participants. Programs that were most effective in reducing prison misconducts also generated lower recidivism rates ($r = 0.13$, $CI = -0.04$ to 0.29). The results of French and Gendreau's meta-analysis demonstrated that on average, custody-based behavioural programs produced reductions in misconducts, which could lead to reductions in reoffending in the community. This study provides further evidence to suggest that institutional behaviour is predictive of post-release behaviour, and that changes in behaviour over time in custody are predictive of modified behaviour in the community.

3.1.2.3 Dynamic risk assessment. It has been asserted that the timing of release to parole is typically based on an analysis of risk and consideration of whether the offender has changed (Serin, Gobeil, et al., 2016). This consideration of change over time reduces the limitations inherent in reliance on static risk factors to assess future risk, upon which release decision-making have historically been based (Carroll & Burke, 1990). Reliance on past criminal behaviour as an indicator of future behaviour may be particularly prone to error for offenders incarcerated for lengthy periods (Mooney & Daffern, 2015). That is, offenders' attitudes and behaviours may have changed over the intervening period since the time of their index offences (e.g., as a result of maturation, development of impulse control).

Various risk assessment measures have been implemented for use by parole boards (see Pearson & McDougall, 2017), including those focused on dynamic risk factors. A sizeable amount of research has been conducted to identify the dynamic risk factors that predict sexual (Hanson et al., 2009; Hudson et al., 2002) and non-sexually violent reoffending (Mooney & Daffern, 2015). Change in dynamic risk factors is purportedly the main mechanism of change in offenders and comparison of dynamic risk factors over time may therefore indicate change in the propensity for criminal behaviour. Less research has been conducted on incarcerated populations in relation to the relevant behaviours in custody, which may be predictive of recidivism upon release, and which may indicate the presence of dynamic risk factors.

3.1.2.4 Behavioural manifestations of risk factors. The OPB (see Daffern et al., 2007) and OAB (see Gordon & Wong, 2009) frameworks were developed to assist with the risk assessment process and evaluation of change over time in custody, as described in section 1.1.10.2 (p. 38) and section 2.1.1.1 (p. 65). Identification and observation of OPB and OAB, and their prosocial alternative behaviours may be beneficial as they focus upon the identification of behavioural manifestations of dynamic risk factors and offending processes, which may provide important insights into incarcerated offenders' risk of recidivism and change in treatment.

In addition to the focus on misconducts specifically, research has also been conducted on the relationship between risk-related institutional behaviour and recidivism post-release (Mooney & Daffern, 2011). One risk assessment measure developed to assist with the monitoring of risk-related behaviours in custody is the VRS (Andrews & Bonta, 2010; Douglas & Skeem, 2005; Polaschek, 2017). The VRS contains factors that are theoretically and empirically related to violence and reflect an

offender's criminogenic needs. It captures the seven dynamic risk areas, which Douglas and Skeem (2005) identified as promising in the prediction of future violent offending: impulsivity; emotional regulation/control; mental disorder; criminal attitudes; substance abuse; interpersonal aggression, relationship with significant others; and, insight into violence (Cochran et al., 2012). One of the key requirements Douglas and Skeem (2005) identified in the conceptualisation of dynamic risk factors, was their ability to change.

The VRS has been the focus of several validation studies (Lewis, Olver, & Wong, 2013; Wong & Gordon, 2006). Lewis et al. (2013) concluded that the results of their study provided support for the VRS risk factors being causal and predictive of future offending, such that reductions in these risk factors lead to reductions in violent reoffending (i.e., these risk factors can be used to evaluate change over time). Gordon and Wong (2009) developed the offence analogue and offence reduction behaviour rating guide to be used in conjunction with the VRS, to monitor specific behavioural manifestations of the VRS' dynamic risk factors. Gordon and Wong (2010) argued that reductions in OAB and increases in ORB need to be related to the criminogenic needs relevant for the particular offender being assessed, in order to translate into a reduction in risk. This argument might be relevant for the use of institutional misconducts as markers for behavioural change over time. That is, if the misconducts are relevant to the offender's specific risk factors, this institutional behaviour could be a valid measure of the offender's change process.

Mooney and Daffern (2013) conducted a preliminary investigation into the application of the offence analogue and offence reduction behaviour rating guide. They found that behavioural manifestations of relevant dynamic risk factors (OAB) and their prosocial alternatives (ORB) are identifiable in the custodial context, with some behaviours associated with violent recidivism. This suggests that these behavioural manifestations change over time and these changes are predictive of future behaviour. The findings demonstrated that most of the behaviours that were associated with violent recidivism, were those reflective of prosocial skills (ORB). The authors noted that OABs related to interpersonal aggression and violence during institutionalisation were among the more frequently recorded behaviours. The results also revealed that several OABs and ORBs were not independently linked to violent recidivism. Mooney and Daffern offered explanations for their results. For instance, behavioural indicators of some dynamic risk factors might not be easily identifiable from file information.

Alternatively, the custodial environment might suppress or fail to provide triggers for some of these factors. Further, behavioural indicators of some dynamic risk factors may be less overt and thus less likely to be identified in this environment.

Research in the Victorian (Australian) correctional context has demonstrated that incarcerated violent offenders who had three or more aggressive incidents recorded against them reoffended violently more frequently and sooner after release, than those with no recorded incidents of aggression in custody (Mooney & Daffern, 2013). However, in this research there was no significant difference in recidivism rates between offenders with no recorded aggressive incidents or those with one or two recorded incidents. Therefore, Mooney and Daffern (2015) noted that the effect may be related to offenders who engage in repeated aggressive acts in custody. Repeated aggressive acts may be an indication that these offenders have not been able to change over time in custody. Further, in Colorado, Heil et al. (2009) found that non-sexual offenders' sexual offending in custody was predictive of sexual offending upon release. Not only are aggressive behaviours in custody predictive of future behaviour, but substance abuse-related behaviours in custody may be predictive of subsequent relapse upon release (VRS; Wong & Gordon, 2000).

As previously discussed, there has been an increase in the use of dynamic risk factors in the assessment of behavioural change over time and the associated assessment of risk for future recidivism. There is ongoing debate regarding the conceptualisation of dynamic risk factors and the ways in which they manifest. Behaviour in custody is relevant particularly in the form of manifestations of criminogenic needs, such as OABs. Using assessment tools to measure the presence of these behavioural manifestations and the changes in frequency over time, may promote more effective methods by which to evaluate offenders' behaviour throughout their incarceration.

3.1.2.5 Staff observations. Atkinson and Mann (2012) demonstrated that custodial officers' observations of offenders' behaviour in custody can assist with the assessment of recidivism risk. Furthering this idea, the trajectory of observations over time could assist with change measurement. However, these behavioural observations are restricted by the degree to which custodial officers monitor relevant behaviour and their interpretation of this observed behaviour. One potential method by which to minimise these limitations is the development of a checklist of relevant behaviours, to which staff can refer (McDougall et al., 1995), such as the predicted OABs and ORBs

identified in the VRS. This process can further assist with the measurement of offenders' change over time.

3.1.2.6 Limitations of using official records. Pearson and McDougall (2017) discussed the importance of assessing the likelihood that a serious offender could reoffend post-release, at an individual-level, prior to forming a decision regarding release to parole. In order to assess this individual-level risk, an evaluation of the offender's change throughout the incarceration period is required. Despite the importance of this process, the difficulty with accurate identification of future behaviour based on behaviour in custody was highlighted through Mooney and Daffern's (2015) findings. At the group-level, offenders with three or more aggressive acts were more likely to offend upon release; however, they also reported that 40.7% of offenders who were recorded as engaging in aggression on three or more occasions during their incarceration were not charged with a violent offence following release. Further, 26% of offenders who were not recorded as engaging in aggression in custody were charged with a violent offence following release.

These results suggest that while official records of aggressive misconduct in custody may provide some indication of an offender's propensity for future violence, its use has limitations. Mooney and Daffern (2015) suggested potential limitations, such as, official records: provide an underestimation of misbehaviour; lack detail regarding the nature, context, and relevance of institutional misconduct; focus exclusively on antisocial rather than prosocial behaviour; and, fail to consider how the custodial environment may suppress or alter an offender's behaviour.

Further, reliance on official records of misconduct or risk assessment results could lead to the loss of valuable information regarding individual offenders' behaviour and associated changes over time. A broader array of behaviours may be relevant to the determination of an offender's progress and prognosis. Mooney and Daffern (2015) noted that these records could serve to supplement judgements produced using formal assessment tools.

3.1.3 Parole System Review

More generally, some researchers have argued for a more structured approach to parole decision-making. Serin and colleagues developed a Structured Parole Decision-making Framework to assist with the decision-making process. One of the seven domains within this framework is "Institutional/community behaviour", which accounts for the offender's behaviour during the current sentence, both while incarcerated and

during previous periods of community supervision. Another of the domains is “Offender change”, which relates to gathering evidence for the offender having benefited from participation in programs, or other changes throughout the sentence. Both these domains highlight the importance of assessing offenders’ behaviour in custody and associated changes within their attitudes and behaviour.

These issues related to developing a structured measure to assist with individual-level decision-making are pertinent to the Victorian parole system. Highlighting the political nature of the parole system and associated decision-making processes, in 2013, Justice Callinan AC, a former High Court Judge, was appointed by the Victorian Minister for Corrections and Minister for Crime Prevention to conduct a review of the parole system in Victoria. It was noted that the effectiveness of the Adult Parole Board is fundamental to the integrity of the parole system. Therefore, the review was established to ensure the Adult Parole Board operates effectively and can respond to the ongoing reforms and changing demands within the justice system.

Within this review, Justice Callinan commented that his impression of the parole system was that the “balance in relation to the grant of parole, its cancellation and the revocation of cancellations may have been tilted too far in favour of offenders, and sometimes, even very serious offenders” (p. 11). He noted that if the public’s safety was the paramount consideration in parole decision-making, the Parole Board should be more risk averse than he perceived it to have become. He acknowledged the importance of maintaining a balance between humanity/reform and “the realities of the threat to society of recidivists at large, and the futility in some cases of the most careful and individually crafted programmes for the release and reform of some offenders” (p. 21); noting that since 2008, arrests for cancellation of parole had risen from 189 to an estimated 800 in 2013.

Justice Callinan indicated that good behaviour in custody is a crucial condition, which an offender must satisfy prior to being granted parole; however, he also noted that there are additional important conditions, such as “an unlikelihood of reoffending, insight into his or her failures, and their consequences, and the merits of a lawful life henceforth” (p. 24). Prior to the review of the parole system, Justice Callinan noted that there were minimal guidelines in the legislation with respect to how the Parole Board should gain information and act in considering whether to grant or deny parole, with no specified measures to use in the decision-making process. He indicated that the detail of

the Parole Board's functions and the tests it applied only existed in its internally compiled Members' Manual.

3.1.4 Relevance of Institutional Behaviour

Further, in 2015, in response to an offender committing murder whilst under a supervision order pursuant to the *Serious Sex Offenders (Detention and Supervision) Act 2009* (Victoria), the Victorian Minister for Corrections commissioned a Complex Adult Victorian Sex Offender Management Review Panel (the 'Harper Review') to provide advice on the Act's legislative and governance models. Within this review, there was minimal discussion related to the assessment of behaviour during an offender's incarceration, as a contributing factor in the ultimate decision regarding future supervision. However, there was reference to the consideration of an offender's behaviour during the initial period of imprisonment at the time of selection of the 'cohort' of eligible offenders under the Act. The review indicated that prison files could be examined to assist in assessing offenders. Further, an eligible offender not initially considered to present a high risk of serious harm to others, could be further reviewed if serious offences were committed in custody. The Public Protection Authority, which would oversee the cohort of eligible offenders, would monitor the offenders throughout their incarceration to facilitate ongoing assessment of progress and risk. This process underscores the relevance of institutional behaviour within parole decision-making.

Considering these two reviews in tandem, the use of behavioural monitoring in custody was highlighted. Specifically, Justice Callinan noted that "No person, ... should be granted parole who has not behaved satisfactorily for at least the person's second half of that person's time in prison. Failure to meet these requirements should be clear disqualification for parole." (p. 95). McDougall et al.'s (2013) exploratory study provided an indication that monitoring of behaviour in custody could be of value in assessing risk of serious harm by high-risk offenders being released into the community. A challenge for Corrections Victoria was the method by which to determine whether each parole applicant's behavior was satisfactory.

3.1.5 Satisfactory Behaviour Rating Guide

In response to Justice Callinan's recommendations, the Satisfactory Behaviour Rating Guide (SBRG) was developed (Daffern et al., 2014) to assist Corrections Victoria staff in operationalising serious sexual and non-sexually violent offenders' behaviour in the second half of their incarceration, to facilitate informed decisions in relation to granting parole. It is possible that the latter stages of an offender's

incarceration are more representative of his or her current functioning. The SBRG provides a method by which staff can identify and evaluate offenders' risk-relevant behaviours, and communicate these behaviours to release decision-makers. Within the SBRG, the risk-relevant behaviours relate to a group of eight central risk factors, based on four aspects of offenders' behaviour, which are predictive of violent offending (Wong & Gordon, 2000). Implementation of this monitoring system is drawn from research demonstrating the benefits of dynamic risk assessment through monitoring behavioural manifestations of offenders' risk-related behaviour in custody (McDougall et al., 2013).

This scale also encourages staff to refer to available sources of information to gain additional insight into the behaviours that offenders have engaged in throughout the latter stages of their incarceration, which allows for the integration of information about changes over time within the assessment. As a result, the final behaviour rating accounts for a variety of behaviours as identified through several sources. Completion of this measure represents the first stage of the parole suitability process. Caseworkers (custodial staff) determine whether the offender's behaviour is satisfactory; if so, they are referred to a parole officer for further assessment.

3.1.6 Current Study

The SBRG has not yet undergone empirical testing and research is required to ensure it is a valid and reliable tool. The aim of the current study was to conduct a preliminary test of the concurrent validity of the SBRG through a retrospective archival analysis of serious sexual and non-sexually violent male offenders who had applied for parole. The hypothesis was that the offenders' caseworkers' (custodial officers') final ratings on the SBRG would correspond with official records pertaining to the domains of: violent behaviour; substance abuse; participation in rehabilitation (as measured by attitude to employment, education and/or rehabilitation); and, response to direction and supervision. Subsequent to this validation process, the intention was for the SBRG to be included as one measure of behavioural change in Study 3.

3.2 Method

3.2.1 Participants

Corrections Victoria identified a consecutive sample of 150 incarcerated offenders, who were all the male violent and sexual offenders who applied for parole from 1 August 2015 from one of nine prisons in Victoria, Australia (see Table 3). The

ethics approval provided for this research did not allow for the identification of offence type within the sample. Participants were serving custodial sentences lasting between 20 and 5469 days ($M = 1280.49$ days, $SD = 1249.79$).

Table 3. *Number of Participants from Each Prison*

Location	Prison Security Level	Frequency	Percent of Total Sample
Barwon Prison	High	20	13.30
Beechworth Correctional Centre	Medium	5	3.30
Dhurringile Prison	Low	14	9.30
Hopkins Correctional Centre	Medium	21	14.00
Langi Kal Kal	Low	26	17.30
Loddon	Medium	18	12.00
Marngoneet Correctional Centre	Medium	19	12.70
Middleton	Low	10	6.70
Port Phillip Prison	High	17	11.30

3.2.2 Materials

The Satisfactory Behaviour Rating Guide was the measure of interest in the current study. It contains four categories, relating to violent behaviour, substance abuse, attitudes and commitment to employment, education and/or rehabilitation, and response to direction and supervision. In order to complete the SBRG, caseworkers are required to consider and review incidents recorded on the Prisoner Information Management System (PIMS), custodial officer reports, and any other information relevant to the four categories included in the SBRG.

Within the category violent behaviour, caseworkers are required to record incidents of physical assaults against others, verbal aggression (including threatening, intimidating or standover behaviours), sexual aggression, possession of weapons, or property damage. Information related to substance abuse is gained through involvement in alcohol or drug related activity, including positive urinalysis results, refusal to engage in random urinalyses, or possession of drug paraphernalia. Evidence of a positive attitude and commitment to employment, education and/or rehabilitation is gained from reports related to the offender's engagement in programs, education and industry. It can also be gained from other staff and may include information about the offender's attendance, participation and performance in these activities. Response to direction and supervision is gauged through evidence of the offender's ability to comply with instructions, abide by rules, and respond positively to direction from staff. Finally, the

caseworker considers the gravity and frequency of incidents within these four behavioural categories, which contributes to the overall rating.

Integration of this material results in a summary rating of the offender's overall behaviour. The caseworker provides a final assessment of the offender's behaviour as Satisfactory, Of Minor Concern, or Of Major Concern. Although there was no formal training for caseworkers in their application of the SBRG, written guidance was provided on the SBRG form and staff were instructed to complete the form for offenders applying for parole. Staff were also informed of the purpose of the SBRG within the parole suitability assessment process. The form included a list of information that caseworkers should seek in completion of the assessment and methods by which this information could be gained.

3.2.3 Procedure

Data gained through the method outlined in section 3.2.2 were de-identified and provided by Corrections Victoria. The central focus of this study was the prediction of caseworkers' (custodial officers') overall behaviour rating. Behaviours within the four variables were identified from the date of reception in custody to the SBRG review date. The data made available were quantified such that the four behaviour categories could be statistically analysed. The methods by which these data were quantified are as follows.

3.2.3.1 Violent behaviour. Recorded PIMS incidents were classified into five subcategories of violent behaviour: physical aggression; verbal aggression; sexual aggression; possession of a weapon; and, property damage. This process facilitated a more descriptive representation of incidents, which would allow further analysis to determine whether particular types of aggressive behaviour have a greater impact on the ultimate behaviour rating. The behaviour frequencies were calculated and are presented in Table 4.

Table 4. *Frequencies of Violence Subcategories*

Behaviour	Number of Incidents	Frequency	Percent of Total
Physical Aggression	0	102	68.00
	1	26	17.30
	2	7	4.70
	3	4	2.70
	4	4	2.70
	5	3	2.00
	6	1	0.70
	7	0	0.00
	8	2	1.30
Verbal Aggression	9	1	0.70
	0	103	68.7
	1	21	14.0
	2	13	8.70
	3	4	2.70
	4	2	1.30
	5	3	2.00
	6	0	0.00
	7	1	0.70
	8	0	0.00
	9	1	0.70
	10	1	0.70
	11	0	0.00
	12	0	0.00
	13	0	0.00
	14	0	0.00
	15	0	0.00
Property Damage	16	1	0.70
	0	145	96.70
	1	1	0.70
	2	1	0.70
	3	1	0.70
	4	1	0.70
Weapon Possession	5	1	0.70
	0	128	85.30
	1	15	10.00
Sexual Aggression	2	7	4.70
	0	149	99.30
	1	1	0.70

Although the five subcategories were each present within the data review, the low occurrence of some behaviours led to a decision to combine frequency scores across the subcategories to produce a Total Violent Incidents score. These total frequency scores are shown in Table 5.

Table 5. *Total Violent Incidents*

Number of Incidents	Frequency	Percent of Total
0	77	51.30
1	25	16.70
2	16	10.70
3	7	4.70
4	4	2.70
5	7	4.70
6	2	1.30
7	1	0.70
8	4	2.70
9	1	0.70
10	0	0.00
11	1	0.70
12	1	0.70
13	2	1.30
14	0	0.00
15	0	0.00
16	0	0.00
17	0	0.00
18	0	0.00
19	0	0.00
20	0	0.00
21	0	0.00
22	1	0.70
23	0	0.00
24	0	0.00
25	0	0.00
26	0	0.00
27	0	0.00
28	1	0.70

The victim of the aggressive behaviour was recorded as follows: staff; other offenders; non-prison workers/relative; and, prison property/objects. The frequencies for each victim type are displayed in Table 6.

Table 6. *Frequency of Total Violent Incidents Directed Towards Each Victim Type*

Victim Type	Number of Incidents	Frequency	Percent of Total
Staff	0	103	68.70
	1	28	18.70
	2	8	5.30
	3	1	0.70
	4	5	3.30
	5	0	0.00
	6	1	0.70
	7	2	1.30
	8	0	0.00
	9	0	0.00
	10	1	0.70
	11	0	0.00
	12	0	0.00
	13	0	0.00
	14	0	0.00
	15	0	0.00
	16	0	0.00
	17	1	0.70
Offenders	0	99	66.00
	1	13	15.30
	2	12	8.00
	3	7	4.70
	4	2	1.30
	5	2	1.30
	6	1	0.70
	7	1	0.70
	8	1	0.70
	9	1	0.70
Non-Prison Workers/Relatives	10	1	0.70
	0	148	98.70
Property/Objects	1	1	0.70
	2	1	0.70
	0	144	96.00
	1	2	1.30
	2	1	0.70
	3	1	0.70
	4	1	0.70
	5	1	0.70

3.2.3.2 Involvement in substance abuse or like behaviour. Information about offenders' involvement in substance use or related behaviour was gained from a review of PIMS incidents, and urinalysis test reports and test result summaries. Behaviours were coded according to eight subcategories: detection of a positive urinalysis; refusal

to undertake random urinalysis; possession of an illicit substance; possession of drug paraphernalia; suspected or observed medication diversion; suspected or observed dealing or receiving illicit substances; suspected or observed drug use; and, suspected or observed alcohol use. Table 7 shows the frequencies of these behaviours.

Table 7. *Frequency of Total Substance Abuse/Use Incidents in Each Subcategory*

Substance Abuse Behaviour	Number of Incidents	Frequency	Percent of Total
Positive Urinalysis	0	110	73.30
	1	15	10.00
	2	9	6.00
	3	1	0.70
	4	3	2.00
	5	5	3.30
	6	0	0.00
	7	2	1.30
	8	1	0.70
	9	1	0.70
	10	0	0.00
	11	0	0.00
	12	1	0.70
	13	0	0.00
	14	1	0.70
	15	0	0.00
	16	1	0.70
Urinalysis Refusal	0	132	88.00
	1	8	5.30
	2	8	5.30
	3	1	0.70
	4	1	0.70
Substance Possession	0	110	73.30
	1	24	16.00
	2	6	4.00
	3	5	3.30
	4	5	3.30
Paraphernalia	0	141	94.00
	1	7	4.70
	2	2	1.30
Diverting Medication	0	135	90.00
	1	8	5.30
	2	3	2.00
	3	2	1.30
	4	1	0.70
	5	1	0.70
Dealing or Receiving	0	137	91.30
	1	11	7.30
	2	1	0.70
	3	1	0.70
Drug Use	0	142	94.70
	1	6	4.00
	2	2	1.30
Alcohol Use	0	141	94.00
	1	7	4.70
	2	2	1.30

Due to the low occurrence of some of these behaviours, a decision was made to combine the subcategory scores to produce a Total Substance Use/Abuse score for each participant. The frequencies for these incidents are shown in Table 8.

Table 8. *Frequency of Total Substance Use/Abuse Incidents*

Number of Incidents	Frequency	Percent
0	82	54.70
1	22	14.70
2	10	6.70
3	5	3.30
4	2	1.30
5	7	4.70
6	5	3.30
7	1	0.70
8	3	2.00
9	2	1.30
10	3	2.00
11	2	1.30
12	3	2.00
13	0	0.00
14	0	0.00
15	0	0.00
16	0	0.00
17	1	0.70
18	0	0.00
19	0	0.00
20	1	0.70
21	1	0.70

3.2.3.3 Attitude to employment, education and/or rehabilitation. Using Corrections Victoria employment history reports, each offender's average pay from reception date to SBRG review date was calculated. Use of this calculation was based on information provided to the candidate that offenders who were positively involved in employment, education and/or rehabilitation received higher pay. Each participant's pay was averaged for the second half of their sentence, from mid-way through their sentence to the date of their application for parole. Although this calculation is not necessarily the most robust measure of attitudes to employment, education and rehabilitation, this proxy variable was considered the most valid method available (see Table 9).

Table 9. *Average Pay Frequency*

Average Pay Range (\$)	Frequency	Percent of Total
3.00-3.99	22	14.70
4.00-4.99	27	18.00
5.00-5.99	33	22.00
6.00-6.99	41	27.30
7.00-7.99	25	16.70
8.00-8.99	2	1.30

Further, the frequency of an offender being “banned” from employment was recorded, as this was considered another useful indicator of performance in employment (see Table 10).

Table 10. *Frequency of Times Banned from Employment*

Times Banned	Frequency	Percent of Total
0	124	82.70
1	13	8.70
2	6	4.00
3	5	3.30
4	0	0.00
5	1	0.70
6	1	0.70

3.2.3.4 Response to direction and supervision. This category was measured through a review of recorded incidents on PIMS. Any behaviour, which reflected that the offender refused to comply with instructions, disregarded institutional rules, or responded poorly to direction, was included within this category. A total score was created through summing the number of incidents (see Table 11).

Table 11. *Frequency of Response to Direction and Supervision Incidents*

Number of Incidents	Frequency	Percent of Total
0	75	50.00
1	31	20.70
2	13	8.70
3	11	7.30
4	4	2.70
5	5	3.30
6	4	2.70
7	1	0.70
8	1	0.70
9	1	0.70
10	1	0.70
11	0	0.00
12	0	0.00
13	0	0.00
14	1	0.70
15	0	0.00
16	0	0.00
17	1	0.70
18	1	0.70

3.2.4 Data Analysis

The relative importance of each of the four categories to the overall behaviour rating was analysed. The analyses involved Spearman's rank-order correlation and ordinal logistic regression, using IBM SPSS Statistics Version 22.0.

Spearman's rank-order correlation is a nonparametric test measuring the strength and direction of the monotonic relationship between variables, rather than the linear relationship. It was used to determine which variables were most strongly associated with the overall rating. Cohen's rule of thumb effect sizes for correlations was used to determine the strength of the associations, with the absolute value of 0.1 = Small, 0.3 = Moderate, and 0.5 = Large (Cohen, 1988). The variables with the strongest associations were used for the subsequent step within the analysis.

As the SBRG overall rating was a categorical variable (i.e., Satisfactory = 0, Of Minor Concern = 1, Of Major Concern = 2), ordinal logistic regression was used; the proportional odds assumption was not violated ($\chi^2 = 8.90$, $df = 4$, $p = 0.06$). The regression analysis was used to determine the relative importance of the variables to the prediction of the overall behaviour rating.

3.3 Results

3.3.1 Descriptive Statistics

The frequencies with which the three overall behaviour ratings were made are presented in Table 12. As seen in Table 12, the majority of the overall ratings were Satisfactory.

Table 12. *Overall SBRG Rating Frequencies*

Overall Rating Category	Frequency	Percent of Total
Satisfactory	102	68.00
Of Minor Concern	37	24.70
Of Major Concern	11	7.30

Of the 102 participants who were rated as Satisfactory, 56 (54.90%) had incidents recorded (including violent behaviour, substance use and noncompliance). Of the 37 participants who were rated as Of Minor Concern, 32 (86.49%) had incidents recorded. Of the 11 participants who were rated as Of Major Concern, 10 (90.91%) had incidents recorded; meaning there was one participant in this category who did not have any incidents recorded.

The prevalence of each behaviour category is presented in Table 13, outlining the number of participants who ever had an incident recorded and the overall frequency of these incidents across participants. Overall, it was less common for participants to have recorded incidents in relation to Times Banned from Work, and the frequency of these incidents was lower than for the other behaviour categories.

Table 13. *Prevalence of Behaviour Categories*

Behaviour	Number of Participants with Recorded Incidents (%)	Total Frequency of Recorded Incidents (%)
Violent	48 (32.00)	288 (31.93)
Substance Use/Abuse	68 (45.33)	325 (36.03)
Times Banned from Work	27 (18.00)	51 (5.65)
Direction/Supervision	75 (50.00)	238 (26.39)

3.3.2 Correlation Between Overall Rating and Incidents

The first step for the data analyses in this study included correlational analyses to determine the association between the overall rating on the SBRG and the quantified SBRG variable categories. These results are displayed in Table 14.

Table 14. *Association between Overall SBRG Rating and Variable Categories*

Variable	1	2	3	4	5	6
1. Overall Rating	-					
2. Violent Incidents	0.44*	-				
3. Substance Use/Abuse Incidents	0.44*	0.52*	-			
4. Average Pay	-0.28*	-0.41*	-0.28*	-		
5. Times Banned from Work	0.28*	0.42*	0.31*	-0.22*	-	
6. Direction/Supervision Incidents	0.48*	0.60*	0.66*	-0.42*	0.39*	-

Note. * = $p < 0.01$

All associations reached significance; however, the effect sizes varied. There was a moderate positive association between the Overall Rating and Violent Incidents, Substance Use/Abuse Incidents and Direction/Supervision Incidents. This result meant that an increase in the number of incidents was associated with a less satisfactory overall rating. Average Pay demonstrated a small negative association with the Overall Rating, reflecting that lower pay was associated with a less satisfactory overall rating. Times Banned from Work demonstrated a small positive association with the Overall Rating, reflecting that an increase in the number of times banned from work was associated with a less satisfactory overall rating. The strength of these two variables' associations with the Overall Rating were the same. Therefore, it was decided that only Times Banned from Work would be included in the regression model. If a future recommendation were that officers use data from one source, it would be easier to access this information than offenders' pay.

3.3.3 Predicting the Overall Rating

Based on the correlational analyses, Violent Incidents, Substance Use/Abuse Incidents, Times Banned from Work, and Direction/Supervision Incidents were included in the logistic regression analysis to predict the overall SBRG rating.

3.3.3.1 Group-level analysis. A test of the full regression model against an intercept only model was statistically significant. This result indicated that the predictors as a set reliably distinguished between whether the final rating was Satisfactory, Of Minor Concern, or Of Major Concern ($\chi^2 = 63.223$, $df = 4$, $p < 0.01$).

There has been debate about the use of various measures by which to model goodness of fit within this type of analysis. The pseudo R square statistics measure the usefulness of the model but do not measure the goodness of fit. Rather, they indicate how useful the explanatory variables are in predicting the response variable (Bewick, Cheek, & Ball, 2005). They can be referred to as measures of effect size. The various pseudo R square statistics have advantages and disadvantages. It has been suggested

that there is limited practical use for the interpretation of the pseudo R square statistics (Hosmer & Lemeshow, 2000). However, within the current research, based on the reported pseudo R square statistics, the model accounted for between 26% and 43% of the variance in the final SBRG rating.

The significance of the variables in the regression analysis can be interpreted in terms of the odds ratios. As there is more than one explanatory variable in the model, the interpretation of the odds ratio for one variable depends on the values of other variables being fixed. The results suggested that each unit increase in total Violent Incidents was significantly associated with an increase in the odds of receiving a higher overall SBRG rating. The odds ratio was 1.26 (95% CI, 1.06 to 1.51), Wald(1) = 6.52, $p = 0.01$. In addition, each unit increase in Direction/Supervision Incidents was significantly associated with an increase in the odds of receiving a higher overall SBRG rating. The odds ratio was 1.27 (95% CI, 1.04 to 1.55), Wald(1) = 5.60, $p = 0.02$.

On the other hand, neither increases in total Substance Use/Abuse Incidents, nor Times Banned from Work, were significantly associated with the odds of receiving a higher overall SBRG rating. The odds ratio for total Substance Use/Abuse Incidents was 1.08 (95% CI, 0.96 to 1.21), Wald(1) = 1.68, $p = 0.20$, while the odds ratio for Times Banned from Work was 1.14 (95% CI, 0.74 to 1.76), Wald(1) = 0.35, $p = 0.55$.

3.3.3.2 Individual-level analysis. At the individual-level, there were some participants whose overall SBRG rating was Satisfactory, even though they had several incidents recorded. For instance, one participant had five Violent Incidents, 17 Substance Use/Abuse Incidents, and three Direction/Supervision Incidents. Another participant had five Violent Incidents, 20 Substance Use/Abuse Incidents, and five Direction/Supervision Incidents. Conversely, there was one participant whose overall SBRG rating was Of Major Concern; however, he did not have any incidents recorded.

In relation to the total frequency of incidents within each behaviour category, although at the group-level, Times Banned from Work featured less prominently than did the other categories, individual-level analysis may also be informative. The maximum number of times a participant was banned from work within this sample was six times. This participant's overall SBRG rating was Of Minor Concern. Another participant was banned from work five times and his overall SBRG rating was Of Major Concern; he also had a greater number of Direction/Supervision Incidents (seven) in comparison with the previously mentioned participant (one).

3.4 Discussion

The primary aim of Study 2 was to conduct a preliminary validation of the Satisfactory Behaviour Rating Guide (SBRG). The SBRG is a recently developed measure designed to aid parole decision-making for serious sexual and non-sexually violent offenders. The ultimate rating of offenders' behaviour was designed to be assessed based on offenders' behaviour within specific categories: violent behaviour; substance use/abuse; attitude to employment, education and/or rehabilitation; and, response to direction and supervision. Within these categories, were behavioural manifestations of dynamic risk factors (OAB) and their prosocial alternatives (ORB), relevant for offenders' risk of violent offending. Each behavioural manifestation was measured using Likert-type scales. Offenders' overall behaviour was considered Satisfactory if it was generally viewed to be prosocial rather than antisocial. It was hypothesised that caseworkers' overall ratings on the SBRG would correspond with official records pertaining to these four behaviour categories.

An overview of the results will be provided. Subsequent to this overview, the results will be further explored in relation to custodial officers' overall ratings, the relation between the overall ratings and individual behaviour categories, discrepancies between the group- and individual-level results, use of behavioural manifestations of risk factors within assessment processes, and implications of the results for the use of custodial officers' ratings. Limitations of the current research will be discussed, followed by suggestions for future research.

3.4.1 Overview of Results

The results demonstrated that each of the four behaviour categories was significantly associated with the overall SBRG rating. Further, higher rates of Violent Incidents and Direction/Supervision Incidents were associated with greater odds of a more severe overall SBRG rating (i.e., Of Minor Concern or Of Major Concern). However, higher rates of Substance Use/Abuse Incidents and Times Banned from Work were not associated with more severe overall SBRG ratings. The group-level findings reflect that ratings of satisfactory behaviour in custody were reliably associated with certain behavioural markers. The results suggested that violence and noncompliance are of greater value in behavioural assessments when compared with other manifestations of risk. At the individual-level, there were some anomalies, which demonstrated a discrepancy between participants' number of reported incidents and the overall behaviour rating.

3.4.2 Custodial Officers' Ratings

Prior research has provided support for the use of institutional misconducts as a measure of adult offenders' behaviour change (e.g., Cochran et al., 2012). However, a potential limitation related to use of institutional misconducts as a measure of offenders' behavioural change is the possible unreliability of observed behaviours. Specifically, there may be variation across contexts within custody, and between staff members, which contribute to whether or not an offender receives a misconduct charge for a certain behaviour. For instance, some behaviours may be observed by some officers but not others and some behaviours may not be considered relevant by some officers whereas other officers may consider the behaviours important. Furthermore, some behaviours may be seen as relevant to reoffending for some offenders but not others. Therefore, there may be different outcomes, in terms of the behaviour being recorded and being regarded as a misconduct, for similar behaviours both between and within offenders. Atkinson and Mann (2012) reported on the differences between custodial officers' decisions to report offenders' problematic behaviour. These decisions were based on a variety of factors, including the perceived severity of the behaviour. Similarly, within the current study, there may have been inherent subjectivity between caseworkers' evaluation of offenders' behaviour, in addition to the overall rating provided based on this behaviour.

The majority of the overall ratings were Satisfactory or Of Minor Concern. However, most participants had institutional misconducts recorded during their incarceration. This result suggests that, at times, custodial officers did not view institutional misconducts as problematic enough to warrant an overall rating of Of Major Concern. The current research did not ascertain reasons behind custodial officers' tendency not to rate offenders' behaviour as concerning. Potential reasons for assessors' relatively lenient ratings include: reluctance to evaluate participants' behaviour as problematic; forgetting that participants had misconducts; and, perceiving behaviour as relatively insignificant compared with more severe infractions.

Further potential explanations for this pattern in the ratings may be gleaned from Atkinson and Mann's (2012) study, which categorised reasons custodial officers provided for deciding not to report offenders' problematic behaviour. It is possible that reasons for custodial officers' disinclination to rate behaviour as unsatisfactory are similar to the reasons for choosing not to report negative behaviour. For instance, custodial officers might require a high threshold for the quantity and frequency of poor

behaviour in order for the overall rating to warrant concern. In addition, they may view their primary role as maintaining order and security, rather than contributing to an assessment or interpretation of behaviour; therefore, they may be less inclined to provide a more severe rating. Relatedly, if they do not consider it to be their responsibility, they may choose not to rate the overall behaviour as concerning if they lack confidence in the evidence to support this rating, such that they do not think their decision is defensible and it could be challenged (e.g., by the offender). It may be beneficial to gain greater insight into potential reasons for custodial officers' ratings, in order to facilitate improvements to the SBRG.

Another possible reason for the pattern in overall ratings may be that the SBRG is insensitive and does not adequately differentiate between offenders' behaviour. However, the group-level analyses revealed that there was an association between the behaviour categories and the overall ratings, which reflects that on the whole, caseworkers' assessments were reflective of the offenders' behaviour, as recorded on incident forms and behaviour in the workplace.

3.4.3 Relation Between Behavioural Categories and Overall Ratings

The group-level results will be discussed in more detail. It is possible that greater frequencies of Violent Incidents and Direction/Supervision Incidents were related to a more severe overall SBRG rating because these offenders may be viewed in a more negative light due to their more overtly oppositional behaviour. The general perception in custody is that misconducts involving violence "most jeopardises the safety of the institution" (Gendreau et al., 1997, p. 416). The authors also noted that custodial officers often have a high level of discretion when charging offenders with nonviolent misconducts, whereas misconduct charges involving physical injury are less subject to interpretation. Therefore, violence towards other offenders may be reported more frequently if the incidents are particularly serious. Consequently, it would be expected that offenders who incur these misconducts are perceived as having greater behavioural problems, which contributes to the greater likelihood of being regarded as unsatisfactory (Of Minor Concern or Of Major Concern). This possible explanation is consistent with prior research in which an evaluation of an anger management course with adult long-term offenders unsuccessfully used disciplinary reports for interpersonal matters as a measure of change. It was noted that these reports were ineffective measures of change because the base rate of reports was too low; that is, custodial officers used official sanctions only as a last resort. Therefore, the official statistics

failed to adequately capture the data (McDougall et al., 1995). Further, violence directed towards staff members is considered to be a significant breach of rules and boundaries, such that these offenders may similarly be regarded as unsatisfactory in their behaviour.

Within this study, behaviours were classified as violent if they involved either physical or verbal aggression. These forms of aggression were combined to form the Total Violent Incidents. Prior research similarly combined forms of aggression to produce a single variable (Mooney & Daffern, 2015). However, since few participants were physically aggressive, and very few were repeatedly physically aggressive, it was not possible to determine whether the frequency of physically aggressive incidents was related to behaviour ratings. Furthermore, it was not possible to determine whether there was a difference between behaviour ratings for offenders who engaged in repeated physical versus verbal aggression. Future research on the SBRG could further investigate the changes in offenders' aggression over time and the association between a reduction in misconduct severity and overall behaviour ratings.

Polaschek (2017) discussed the positive influence of protective factors and responsiveness on offenders' relationships with staff and the ability of offenders to work effectively with staff to achieve change. The converse may be true for violent behaviour, which is likely to be more challenging for staff members working with offenders as compared with drug-taking behaviour culminating in a positive urine screen. Similarly, staff may have more negative perceptions of offenders, and regard their behaviour as problematic, when they respond poorly to direction/supervision. A perception of offenders as adversarial might lead to more severe overall ratings, as observed in the current results.

Turning to the behaviours that were not significant within the model, it is worth considering why neither an increase in Substance Use/Abuse Incidents nor Times Banned from Work was predictive of a more severe overall behaviour rating. In Atkinson and Mann's (2012) study in which custodial officers identified potential offence paralleling behaviours in sexual offenders, substance use/abuse-related behaviours did not feature. It is possible that in general, substance use/abuse behaviours are considered less severe rule breaches as compared with other behaviours. If they are perceived to be prevalent in custody and less harmful to others, they might not be emphasised in ratings regarding the satisfactoriness of an offender's behaviour. These

behaviours might also be considered a consequence of dependence, which assessors might excuse more than confrontational or aggressive behaviour.

Although statistically significant, Times Banned from Work only had a small association with the overall SBRG rating. This behaviour was recorded less frequently overall in comparison with the other behavioural categories, which might have contributed to its limited predictive value in the overall rating. That is, there might not have been enough variation in these incidents, for the regression to identify how it affected the overall rating. Problems at work might not have been considered concerning enough to warrant an incident report, and when incidents were recorded, caseworkers may have dismissed them as less relevant than the other behaviour categories in determining an overall rating. When considering problematic behaviour for the overall rating of satisfactoriness, assessors might focus more on offence-related behaviour such as violence, or behaviour that is relevant for the custodial routine.

3.4.4 Discrepancies Between Group- and Individual-Level

At the individual-level, it appeared that a higher number of Times Banned from Work did contribute to the overall behaviour rating, but seemingly only in combination with the other behaviour categories. For example, the primary difference between the two participants who were banned from work repeatedly but had different overall behaviour ratings (Of Minor Concern vs Of Major Concern), was their frequency of Direction/Supervision Incidents. That is, the participant with more repeated Direction/Supervision Incidents received the more severe overall rating.

The overall behaviour ratings should also be considered at the individual-level. For some participants, there were disparities between recorded incidents and overall behaviour ratings. Some participants who demonstrated poor behaviour reflected through recorded incidents, were assessed overall as Satisfactory. There were also participants who had no incidents recorded but gained an overall Of Minor Concern or Of Major Concern rating. It is possible that the participants who were rated as Satisfactory but had prior misconduct incidents had improved behaviourally as they progressed through their sentence; this would highlight the importance of exploring changes in behaviour over time to determine offenders' behaviour trajectories throughout their incarceration. As Dvoskin and Heilbrun (2001) noted, if only static factors are measured, the risk appraisal prior to release is unlikely to be changed by the course of offenders' incarceration. However, this change over time would not explain

the Of Minor Concern or Of Major Concern ratings for participants with no recorded incidents.

3.4.5 Behavioural Manifestations of Risk Factors

Dynamic risk factors have been the focus of research into the underlying causes of offending and methods by which to assess ongoing risk (Beech & Craig, 2012; Cording, Beggs Christofferson, & Grace, 2016; Ward & Fortune, 2016). The OAB framework was developed to assist with the assessment of dynamic risk factors in custodial settings. The use of this framework may be beneficial as monitoring OABs and ORBs throughout an offender's incarceration could clarify whether dynamic risk factors are decreasing, which is indicative of decreased risk (Gordon & Wong, 2010).

These issues are pertinent with respect to the current findings. The results suggested that caseworkers did not perceive offenders' substance use/abuse or difficulties at work to be important factors contributing to their overall behaviour ratings. It is possible that these behaviours are not relevant for offenders' risk; however, prior research has demonstrated that substance use and work ethic are risk factors predictive of future recidivism (Douglas & Skeem, 2005). In addition, Mooney and Daffern (2013) found that positive behaviours associated with work ethic were predictive of lower recidivism rates. It is also possible that these factors might not be reliable measures within the custodial context because assessors do not perceive them to be significant. This explanation may suggest the importance of providing training for assessors in relation to the risk factors relevant for offenders' future behaviour.

The behaviours assessed through the SBRG were not individualised, as is recommended within the implementation of the OPB and OAB frameworks. Caseworkers providing ratings for the offenders were likely unaware of the specific behavioural manifestations of the risk factors within the measure, which were relevant for the individual. Therefore, the relevance of each risk factor for the individual was not known. However, McDougall et al. (2013) identified the value in focusing on behaviour that might not meet the definition of OPB (e.g., it is not directly related to the offender's index offence), but is still generally problematic and offence-related. Within their research, they did not attempt to identify behaviour as OPB. They found consistency between the behaviours exhibited in custody and those subsequently exhibited in the community. Their findings suggested that both low- and high-level negative behaviours in custody are relevant in predicting future behaviour in the community.

It is pertinent that as the first stage of the parole decision-making process, the SBRG was not used as a risk assessment tool, with caseworkers providing an evaluation of offenders' general behaviour within custody without focusing on risk of recidivism. The overall SBRG rating provided by caseworkers should not be interpreted as a risk rating. However, caseworkers' overall ratings were based on offenders' behaviour relevant to empirically derived risk factors for violent recidivism. If an offender's behaviour was rated as Satisfactory, it would generally mean that overall, he engaged in more prosocial rather than antisocial behavioural manifestations of these risk factors. As a result, assuming the behavioural manifestations were relevant for the offender's individual case, based on the theories underlying the OAB framework, he would likely be at lower risk of recidivism. While many risk assessments primarily measure static risk, the use of institutional misconducts as a measure of risk-related attitudes and behaviours can be considered dynamic (Mooney & Daffern, 2011), as these attitudes and behaviours can fluctuate throughout an offender's period of incarceration. Therefore, an offender's overall rating may indicate that the risk factors relevant to his offending are less prominent and may lead to a reduction in future offending. On the other hand, it is also possible that the overall rating is an indication of static risk, rather than change over time. In this scenario, an offender who gains an overall rating of Satisfactory may remain at the same risk level for reoffending as he did upon his incarceration. As previously indicated, further investigation of the relation between these ratings and subsequent recidivism is warranted.

3.4.6 Use of Staff Ratings

Prior research has suggested the potential use of custodial officers' observations of behavioural manifestations of risk factors, to gather valuable information about offenders' behaviour (Atkinson & Mann, 2012). In the current research, custodial officers' collation and interpretation of offenders' misconducts and other available information was used to ascertain an overall behaviour rating. The discrepancies between some ratings and the institutional misconduct incidents may suggest that there are limitations to the use of custodial officers' evaluations of offenders' behaviour. Such limitations might be due to custodial officers' biases (e.g., some custodial officers may perceive that an offender who dismisses direction and instruction from those in authority has more problematic behaviour than an offender who uses substances in custody, while other custodial officers may have different perceptions of the differential severity of behaviour in custody) and consequently the behaviours they focus on to

provide final ratings (e.g., using the previous example, offenders who were charged for non-compliance with directions may have been rated as more problematic overall as compared with offenders who were charged for substance abuse-related behaviours). Alternatively, these limitations may be due to caseworkers' lack of training in how to complete the SBRG. As noted previously, there was no formal training in the SBRG so it may be that the aforementioned discrepancies were attributable to unreliable completion of the SBRG.

Regardless of the assessor's knowledge base, clinicians and decision-makers must remain mindful that all assessment tools have the potential for errors. For instance, McDougall et al. (2013) found that for one offender out of the eight for whom recidivism or recall to custody was predicted, a high frequency of negative behaviour in custody did not lead to this outcome. Further, Mooney and Daffern (2015) found that 26% of offenders in their research who had no recorded aggression in custody were charged with a violent offence post-release. Due to the potential for errors and assessors' subjectivity, it is unsurprising that there were discrepancies between ratings of satisfactory behaviour and the frequency of official misconducts in the current study.

3.4.7 Limitations

There were some limitations in the current research. Ethical considerations prevented identification of sexual and non-sexually violent offenders. Therefore, there could be no comparison between offence type in relation to assessing the validity of the measure for violent as opposed to sexual offenders. Further, this lack of differentiation meant that it is unknown which results were specifically relevant for sexual offenders.

More generally, there was minimal demographic information on participants available within the current research. Consequently, the differential impact of sample characteristics could not be assessed. Therefore, there are limits to the information that can be gained in relation to specific factors contributing to the results. For instance, the relation between risk level and behaviour ratings could not be compared. There is minimal research into the relevance of behaviour in custody while controlling for risk level and it remains an important area to examine (Mooney & Daffern, 2011). Since many characteristics of the current sample remain unknown, there are also limitations to the generalisability of these results to other populations.

It is common for offenders' misconducts to decrease in frequency over their incarceration, in particular for offenders with long sentences, and if incarcerated at a young age. For instance, a longitudinal study demonstrated that offenders who were

incarcerated for periods of approximately seven to 10 years became more adaptive within the custodial environment over time; this effect was greater for offenders who were initially incarcerated at a younger age (Zamble, 1992). However, due to the constraints of the current research, recorded incidents throughout participants' period of incarceration were used, rather than limiting the time period to the second half of their incarceration as Justice Callinan AC recommended in his review. Therefore, there was no consideration of changes that participants may have made since their initial incarceration. This is important since participants may have engaged in numerous problematic behaviours early in their sentence but changed and then desisted from these behaviours. An additional limitation was that participants' sentences were varying lengths. Therefore, they may have been incarcerated for different lengths of time, which may have had an effect on the behaviours observed and the misconducts accrued.

3.4.8 Future Directions

Atkinson and Mann (2012) reported that there were various reasons why custodial officers did not report offenders' poor behaviour. These reasons included the behaviour being perceived as 'normal' for the individual or within the prison culture. Prior research has demonstrated that the frequency of negative behaviour in custody was a statistically significant predictor of recall to custody (McDougall et al., 2013). Frequency of negative behaviour has been shown to more effectively predict cross-situational behavioural frequency than the seriousness of behaviour in custody (Hill, 1985; McDougall et al., 2013). Therefore, it is important to ascertain the frequency of all misconducts, including low-level incidents. As such, recording of all negative behaviour is important. Nevertheless, future research might seek to determine how to assist custodial officers to distinguish between 'normal' behaviour that is not risk-related, either generally or for a particular individual, and behaviour that is particularly meaningful because it relates to criminal offending.

In the current research, there was no indication of the level of severity of the recorded incidents. Therefore, lower-level behaviours within these categories might have remained unreported, which could have influenced the pattern of results. Additionally, the threshold for custodial officers' decisions to report misconducts might have varied, either by custodial officer, offender or prison. If information about the severity of recorded misconduct incidents is available, future research could be conducted to investigate whether it is a relevant factor in predicting the overall rating.

The disparity between some participants' incidents and overall behaviour ratings, in addition to prior research findings including recidivism data, suggests that additional factors may be relevant (e.g., criminogenic needs, length of incarceration, motivation to change), which contribute to both short- and long-term outcomes (i.e., institutional misconduct and recidivism). The nature of these contributing factors could be investigated in the future. It may also be beneficial to conduct further research into the factors staff consider in their decision-making for overall behaviour ratings.

Discrepancies between custodial officers' ratings and reported incidents suggests there are some limitations to the use of custodial officers' observations. It may be beneficial to conduct further research into the use of custodial officers' behavioural assessments of offenders to gain greater understanding of the value in this process. If custodial officers' observations are to be used more widely within the decision-making process, it is vital that potential limitations are identified and modified, and the benefits are further developed. Further, it may be useful to provide additional training to caseworkers in the completion of the SBRG if it is to be implemented more widely.

Evidence remains inconclusive regarding the predictive accuracy of offenders' behaviour in custody for recidivism. The current research provided preliminary support for the validity of the SBRG as a measure of offenders' behaviour in custody that is completed by custodial officers. It could be extended in the future to include recidivism outcomes. Future research could use the current sample to determine whether those offenders who were rated as Satisfactory are less likely to breach parole or reoffend than those who were rated as Of Minor Concern or Of Major Concern. Added benefit might be gained for this measure's efficacy by comparing its use with different groups of serious offenders, such as sexual and non-sexually violent offenders. It may also be beneficial to compare the use of this measure of change with other measures of change, to gain further evidence of its utility.

3.4.9 Conclusions

This research provided a preliminary validation of the SBRG as a structured tool by which to measure individual offenders' behaviour in custody to assist with the assessment of change and parole decision-making. SBRG ratings were related to offenders' behaviour, providing additional support for the presumption that behavioural manifestations of risk factors assist decision-making in terms of an offender's readiness for release to parole. Consequently, there is further support for monitoring OPB and OAB, and their prosocial equivalents, within a custodial environment.

Chapter 4: Study 3

This final study incorporates information from the first two studies. The first study was a survey for psychologists, who provided examples of expected behavioural manifestations of empirically derived risk factors for sexual offending, in addition to prosocial alternative behavioural manifestations. The psychologists' survey culminated in the development of a behavioural checklist, for implementation in a custody-based treatment program for sexual offenders. Subsequently, a recently developed behavioural monitoring measure, the Satisfactory Behaviour Rating Guide (SBRG), was validated in Study 2. Although it was not designed to be a measure of change, it was designed to assess offenders' behaviour and consider this behaviour when making judgements that contribute to an ultimate decision about an offender's suitability for release on parole. In Study 2, results revealed that overall, custodial officers' SBRG ratings were consistent with information gained through officially recorded information (e.g., institutional misconducts). The results from Study 2 raised some potential limitations of using custodial officers' ratings of offenders' behaviour, through their reluctance to conclude offenders' behaviour was unsatisfactory.

This Introduction to Study 3 will provide an overview of prior research specific to the aims of the current study. The information contained in this section was outlined in Chapter 1 and will now be further discussed in relation to the rationale for conducting this study. It will cover information about the behavioural manifestations of risk factors in custody, providing structure within behavioural observation, and change measurement and related issues. These ideas are central to the current study.

4.1.1 Behavioural Manifestations of Risk Factors

In controlled environments, such as custody, blatant antisocial or offending behaviours that are easily detectable in the community could be altered or inhibited due to the absence of potential victims, triggers, or circumstances (Gordon & Wong, 2015). For instance, an offender incarcerated for sexual offences against children may engage in alternative behaviours in custody such as viewing, and masturbating to, images of children in magazines. These proxy behaviours (offence analogue behaviours; OABs) within custodial settings are often indications that the causes underlying the problematic behaviours have been maintained. Often, OABs are more socially acceptable proxies of the individual's criminogenic needs when the open expression of the deviant behaviours are heavily sanctioned, such as watching children on television programs rather than in person (Olver & Wong, 2016).

Similar to these negative behaviours, positive behaviours might also be less easily observable in custody, or if they are observed, they might not be documented due to the punitive culture within the custodial context (McDougall et al., 2013; Mooney & Daffern, 2013). It is relevant that in the psychologists' survey (Study 1), some positive behavioural manifestations of risk factors were reported as the absence of the associated negative behaviour. However, within the OAB framework, offence reduction behaviours (ORBs) are the socially appropriate skills and behaviours that the individual has developed to replace or manage past problematic behaviours. Therefore, these positive behaviours are not the absence of OABs, but rather, they are skills added to the offender's behaviour repertoire.

Olver and Wong (2016) noted that observations of ORBs are particularly critical in controlled settings such as custody, due to the many artificial situations an offender may be placed in, which could inhibit and reduce problem behaviours. There are situations in which it might be assumed that the absence of the problem behaviours reflects positive changes, while it is merely an indication that the problem behaviours have been inhibited as a function of the context. Alternatively, as Daffern et al. (2007) described, the offender might have developed detection evasion skills, which allow maintenance of the problematic behaviour without detection. Therefore, in the development of a checklist to assist staff monitoring behaviours, it is important to ensure the positive behaviours are not merely the absence of negative behaviours, but are positive skills that offenders have learned.

4.1.2 Guidance in Behavioural Observation

Structured measures that decrease potential bias in assessments and decision-making are useful in a custodial context (Gordon & Wong, 2015). Such systematic measures allow staff to view and rate offenders' behaviour with greater objectivity. As a result, the observation and monitoring of risk-related behaviour could be improved through the use of objective monitoring forms. Such measures could also assist in providing increased balance to the types of behaviours monitored, such that the focus is not only on negative behaviours but also includes positive, prosocial behaviour. As Mooney and Daffern (2013) noted, it is important to increase the focus from solely institutional misconduct, to the consideration of both reductions in problematic behaviour and the development and maintenance of positive behaviours, in the assessment of behavioural change and progress over time. Further research into the use of staff observations to inform decision-making may be beneficial.

As demonstrated in prior research, although some behaviours are regularly recorded in custody (e.g., substance use and violence), other risk-relevant behaviour is rarely recorded in official records; in particular when it is related to prosocial behaviour that demonstrates an improvement in risk (Mooney & Daffern, 2013). In order to assess risk and behavioural change, it is crucial that relevant behaviours are identified, observed and monitored. An assessment of treatment change has been developed for violent offenders, through the VRS and its associated OAB and ORB rating guide (Gordon & Wong, 2009). This rating guide was designed to assist clinicians to assess the presence of relevant behavioural manifestations of risk or improvement for each VRS dynamic risk factor (Mooney & Daffern, 2013). Similarly, this process in which relevant behavioural manifestations of risk factors are identified, has been implemented in the HCR-20v3 (Douglas, Hart, Webster, & Belfrage, 2013), a structured professional judgement violence risk assessment measure. For each item within the HCR-20v3, a list of “indicators” (illustrative examples) is provided to assist with rating the presence of each risk factor (Strub, Douglas, & Nicholls, 2014). Of note is that whereas the VRS encourages observation of prosocial behaviours, the HCR-20v3 focuses only on negative behaviour.

The use of the VRS OAB and ORB rating guide, to identify risk-related behaviour relevant to individual offenders, and to increase staff awareness of more subtle behaviours that might otherwise remain undetected, emphasises the importance of this process with various offender groups. Guidance of this nature within sexual offender research is in its earlier stages, with an OAB and ORB rating guide recently developed for use with the VRS:SO (Olver, Gordon, & Wong, 2017). Mann, Thornton, et al. (2010) emphasised that it is crucial to have a clear understanding of what constitutes risk factors for sexual offending in order to effectively implement frameworks used to monitor and assess behavioural manifestations of dynamic risk factors. This understanding is also important in identifying behavioural manifestations of risk factors for the purpose of measuring treatment change. The current research serves to contribute to this field through the development of a behavioural checklist outlining suggested behavioural manifestations of risk factors relevant to incarcerated sexual offenders. Further contribution is related to implementing a measure that includes both negative and positive behaviours rather than focusing on negative behaviour, which remains the focus of most common assessment instruments.

4.1.3 Change Measurement

Gordon and Wong (2015) noted that one primary challenge for forensic service providers is to determine whether clients' risk of violence, antisocial and offending behaviours have been moderated while under their care; this task is particularly difficult in relation to offenders who have been incarcerated for substantial periods. Current functioning can assist with this process, to determine whether changes have already been made (i.e., the person is not more prosocial). However, changes in risk-related behaviour must be relevant to criminogenic needs in order to reduce risk (Gordon & Wong, 2010). Therefore, the benefits of the methodology of the current research are highlighted, such that the risk factors forming the basis of the behavioural checklists were gained from theoretically driven risk assessment protocols, namely the RSVP and VRS:SO.

As described in section 1.1.12 (p. 48), there are various methods by which change assessment can be conducted. There are advantages and limitations in all forms of measurement. Within the current study, both self-report and observer ratings were implemented with the behavioural checklist. These two forms of measurement will now be elaborated.

4.1.3.1 Self-report. A common approach to change measurement is using offender self-report. For instance, pre- to post-treatment psychometric measures are often used to determine whether offenders' attitudes and beliefs have changed through treatment. However, conclusions about the use of these measures have been mixed (Barnett et al., 2012; Beggs & Grace, 2010; Stevens et al., 2016; Wakeling et al., 2011; Walters, 2006).

Pearson and McDougall (2017) indicated that information communicated in risk management meetings is often characterised as offender self-report and major behavioural misconducts resulting in disciplinary sanctions. They labelled this information as "tip of the iceberg" risk behaviour. They noted that this form of behaviour identification and communication does not allow for awareness of lower-level or hidden problem behaviours, which might be offence-related behaviour. They discussed case study examples involving serious offenders, who had completed treatment programs and convinced decision-makers that they had made clinical changes, but subsequently continued offending upon release from custody. These individual cases provide evidence for the challenges associated with reliance on offender self-report. It highlights the value of implementing multiple measures of change to assist in

decision-making. Another form of measuring change is to engage staff in the process of behavioural monitoring. The use of observer ratings will now be further discussed.

4.1.3.2 Observer ratings. Prior research has encouraged the use of observer ratings, in order to supplement self-report and gain a more accurate representation of offenders' change (Serin et al., 2013). There are various forms of observer ratings that can be used, such as clinicians, custodial staff, education staff, and employment overseers. Due to the diverse roles of the observers, and different contexts in which observers interact with offenders, varied information can be gained regarding offenders' behaviour when reports from different staff are gathered.

Gordon and Wong (2015) emphasised that offenders' behaviours can be observed through different lenses, which may be "coloured by one's professional training, experience, situational demands, political and organisational pressures, not to mention subjective opinions and personal biases" (p. 97). For instance, one issue that might arise in assessing treatment change, is that it is open to subjectivity and bias when the assessment is conducted by the treatment provider, due to potential vested interests (Olver & Wong, 2016).

To counter this potential bias and gain greater depth of information, Olver and Wong (2016) noted that collateral information sources are especially important, such as observations from other staff who have the opportunity to observe the offender in contexts outside of treatment. In particular, custodial officers can enrich the depth of information available in the change measurement process, through providing observations of offenders' behaviour within the context of their daily living (see Atkinson & Mann, 2012). Behavioural monitoring by observers within the custodial context can be a complex task. Some of the complexities will now be further explored.

4.1.3.2.1 Behavioural monitoring. It has been suggested that when using structured professional judgement tools, the reliability and validity of the outcome can be affected by the sources of information that are examined and as such, it is important for a variety of different sources to be examined and used in the decision-making process (Beech et al., 2016). These sources of evidence should include information from others who have insight into the offender's daily functioning, such as wing staff. However, it may not be sufficient to ask observers for their opinion, as they may not understand the relevance of behaviours they observe. In addition, like other staff, their opinion might be biased due to other factors such as their personal view of the offender.

Further, Mann et al. (2010) indicated that most behavioural observations recorded in secure settings come from staff who may not have been trained in differentiating risk-relevant from irrelevant behaviours. As a result, they may tend to focus on behaviour indicative of an offender's risk to himself or others, or to the security of the environment, rather than behaviour indicative of recidivism risk. Therefore, it is important to formulate potential indicators of risk and progress for the offender and ask others to rate these items in a more objective manner. Pearson and McDougall (2017) also discussed the use of structured behavioural monitoring. They suggested that for serious offenders who are deemed high risk of serious harm, risk management should be informed by the consistent application of a behaviour monitoring protocol to examine the continuity of offence-related behaviour across community and custody, including post-release.

Behavioural monitoring to inform assessment has a long history of use in non-forensic fields. For instance, a highly-structured behavioural monitoring scale was developed for nursing staff in aged care: the Cohen-Mansfield Agitation Inventory (Cohen-Mansfield, Marx, & Rosenthal, 1989). It is less common in forensic fields. However, Atkinson and Mann (2012) focused on the value that might be gained through use of custodial officers' observations of offenders' behaviour, based on predicted risk-related behaviours. One potential limitation for this methodology is the risk of confirmation bias, such that observers are primed towards noticing similar behaviours to those presented on checklists. Confirmation bias might lead to an overestimation of the presence of these behaviours.

Despite the potential limitations, behavioural monitoring might assist the risk and change assessment process, and clinical practice more generally. Mann, Thornton, et al. (2010) indicated that risk assessment frameworks for sexual offenders are stronger at identifying past risk factors (e.g., in the lead-up to offending), but weaker at identifying currently active risk factors. These currently active risk factors can be measured through their behavioural manifestations; that is, OAB and ORB.

The measurement of current functioning can be achieved through behavioural monitoring of change over time, in particular through prospective research. Prior research involving violent offenders has engaged a retrospective design (Mooney & Daffern, 2013). However, a prospective study design may facilitate greater accuracy in direct behavioural monitoring. It is important to observe and record this information, in

order to monitor changes in offenders' risk level while in custody. It is expected that the current research will serve to further promote this goal.

Skeem et al. (2002) emphasised the use of multiple time points in behavioural monitoring, through conducting 26 weekly interviews with patients in a psychiatric facility, in parallel with gaining collateral information from others who were familiar with the patients' activities. Multiple time points can assist in determining whether behaviours are truly dynamic and to identify individuals who are involved in repeated problematic behaviour.

Gordon and Wong (2015) noted that if intervention is successful, at its completion, offenders should have replaced the majority of their OABs with ORBs, and this change should have promoted a reduction in recidivism risk. With further practice and support, ORBs should become a central component of the offender's behavioural repertoire. Prior to an offender's release, an offender should have reduced his OABs and increased his ORBs such that in the community, the majority of his behaviour is related to the ORBs.

This change from antisocial to prosocial behaviours may also be evidenced through the SBRG. Within the SBRG, if the majority of an offender's ratings on the scales within the domains reflects prosocial behaviours, he/she should be more likely to receive a satisfactory behaviour rating. In Study 2, the majority of participants' ratings were Satisfactory; however, due to the absence of raw data, it was not known whether the behaviours were mostly rated as prosocial. Use of a behavioural checklist outlining positive and negative behaviours, such as those identified in Study 1, might provide greater depth of information regarding offenders' change over time. McDougall et al. (1995) reported that using a checklist to monitor offenders' behaviour provided preliminary evidence for behavioural change on some scales.

4.1.4 Current Study

The setting for Study 3 was the Custody-Based Intensive Treatment (CUBIT) program at Long Bay Correctional Complex in Sydney, Australia. It is a treatment program for adult males who have been convicted of a sexual offence against a child or adult. It is conducted within a therapeutic community. Further details about CUBIT are provided in section 4.2.1.1. As with many therapeutic models, which take an approach goals perspective to guide clients on how they can best achieve their identified goals rather than avoiding potential threats (Fortune, Ward, & Willis, 2012), offenders in CUBIT are encouraged to focus on working towards goals that promote a more

supportive and helpful lifestyle. A combination of internal conditions (skills and capabilities) and external conditions (opportunities and supports) are central to the possibility of individuals replacing unhealthy ways of achieving good lives with appropriate and healthy ways of achieving their core values (Ward & Brown, 2004). Treatment in a therapeutic community such as CUBIT provides the opportunity and support for offenders to engage in this process. As a result, changes in risk-related behaviour would be expected to take place throughout treatment, such that these behaviours are replaced with more adaptive, prosocial behaviours.

4.1.4.1 Behavioural checklists. Within treatment, and the associated measurement of change, it is important to ensure that there is a balance between observation of negative behaviour with recognition of positive behaviour. As previously mentioned, the behavioural checklist, developed from the results of Study 1, was a central source of data for Study 3. The checklist contains both negative and positive behavioural manifestations of risk factors relevant for sexual offending, as identified by professionals with experience in the field.

The checklist was developed such that numerical responses were sought, indicating the frequency with which each behaviour was engaged in over the past week. The format of the checklist will be further described in section 4.2.2. This format, in which frequencies were required, is in contrast to that which McDougall et al.'s (1995) checklist described. These checklists provided options for 'Often', 'Sometimes', or 'Always' engaged in the behaviour, rather than quantifying the responses. This method allows for greater observer subjectivity in relation to what is considered 'Often', 'Sometimes', and 'Always'. Further, from a research perspective, this format provides greater difficulty for statistical analyses as it is not quantified. From a clinical perspective, it also increases the difficulty for clinicians to evaluate changes made over time, unless scores are assigned to the ratings. Whereas the behavioural checklist in the current study was completed on a weekly basis, McDougall et al.'s (1995) checklist was completed once prior to the commencement of intervention, once upon completion of the intervention, and "at various stages of follow-up" (p. 89).

The current research extended McDougall et al.'s (1995 and 2013) work, in relation to the development of a checklist that could be used to monitor offenders' behaviour. That is, the behavioural checklist was completed by both custodial officers and offenders; the latter as a measure of self-report. Not only can a combination of self- and observer-report gain valuable information about offenders' change, but use of self-

report might also be advantageous for participants. Ward (2017) promoted the idea of increasing offenders' sense of agency, which, he argued, would facilitate desistance. Through completion of weekly checklists in which participants take responsibility for the behaviours they are engaging in, an increase in their sense of agency could be expected. The process might focus participants' attention towards their daily behaviours and the ways in which their decisions are impacting on their lives, in the short-term if not on an ongoing basis. Additionally, despite the potential challenges associated with using offender self-report to measure change, such as positive impression management, Morrison-Beedy, Carey and Tu (2006) described contemporaneous self-report as the 'gold-standard' of data collection. In the current research, the completion of weekly behavioural checklists served as contemporaneous self-report.

4.1.4.2 Multiple change measures. An important part of the current research was to investigate the processes by which change-related information can be integrated in a clinical context, such that the benefits of different approaches could be consolidated. This method can inform forensic mental health or judicial decision-making about matters such as conditional release and supervision (see Lewis, Olver, & Wong, 2013). Identifying changes in behaviour over the course of treatment might assist in the development of a more reliable method of gaining information about future risk.

Providing support for the use of multiple measures of change is the idea that some offenders who are cooperative and compliant in custody subsequently reoffend after release to the community (see Mooney & Daffern, 2015). Therefore, there might be subtle indicators for ongoing problematic beliefs or behaviour, to which some change measures might be more sensitive. Further, the potential bias through both self-report and observers with different roles provides support for the use of multiple assessment modalities to limit the impact of biases.

Therefore, in combination with the behavioural checklist, three additional sources of information about offenders' change were used. These included measures focusing on negative attitudes and behaviour, and those focusing on both negative and positive attitudes and behaviour: pre- and post-treatment psychometric measures; SBRG; and, Treatment Gain: Short Scale. Both participant self-report and staff observations were used to gain information about behavioural change over time. Multiple time points were used through the implementation of these four different measures. Further, within the behavioural checklist alone, information was gained at multiple time points. As Olver and Wong (2016) highlighted, it is beneficial to gain information about an offender's

base rate of engaging in behavioural manifestations of risk factors. In this study, the first time point at which data were collected for the behavioural checklists was prior to participants' treatment commencement. This procedure facilitated the identification of the behavioural frequency prior to intervention. Therefore, a more accurate assessment of change over time could be gained.

4.1.4.3 Aims and hypotheses. The three aims and hypotheses for this study are reviewed below, as previously outlined in section 1.1.1.5 (p. 61).

Aim 1:

- a) To determine whether dynamic risk factors and prosocial equivalent behaviours manifest in a custodial therapeutic community in a sample of sexual offenders;
- b) To determine whether the behavioural manifestations were evident to both offenders and custodial officers; and,
- c) To determine whether some behavioural manifestations were more evident than others.

Hypothesis 1:

- a) Sexual offenders' dynamic risk factors for sexual offending and prosocial equivalent behaviours would manifest in a custodial therapeutic community;
- b) The behavioural manifestations would be evident to custodial staff and offenders; and,
- c) Some behavioural manifestations would be observed with greater frequency than others.

Aim 2:

- a) To determine whether the frequency of dynamic risk factors and prosocial equivalent behaviours change over the course of treatment; and,
- b) To determine whether custodial staff and offenders' perceptions of the presence and change in manifest dynamic risk factors and prosocial equivalent behaviours correspond, and whether there is a change in their correspondence over the course of treatment.

Hypothesis 2:

- a) Over time, there would be a reduction in dynamic risk factors and an increase in prosocial equivalent behaviours; and,
- b) Prior research has demonstrated the tendency for the group process (particularly within therapeutic communities) to increase offenders' self-

awareness and motivation for increased openness (e.g., Corey, 2017; and Ware et al., 2010). There has also been evidence for the impact of treatment on offenders' development of alternative strategies to manage risk-related behaviours (Andrews et al., 2011), and the importance of emphasising for observers the behaviours that are relevant (see McDougall et al., 1995). Therefore, it was hypothesised that initially, custodial officer observations of participants' behaviour would differ from participant self-report due to participants' tendency to minimise reporting of dynamic risk factors and emphasise prosocial equivalent behaviour. However, over time in treatment, the two would converge as participants' insight increased through treatment and custodial staff observed more prosocial and less risk-related behaviour.

Aim 3:

- a) To determine whether changes in these behaviours correspond with other markers of change, such as reliable and clinically significant change in the psychometric measures used in pre- and post-treatment testing, and the Treatment Gain scale and Satisfactory Behaviour Rating Guide.

Hypothesis 3:

- a) These changes would correspond with the reliable and clinically significant change outcomes measured through pre- and post-treatment psychometric tests. It was also hypothesised they would correspond with change as assessed using the Treatment Gain scale. Additionally, those participants who demonstrated positive change (reduced dynamic risk factors and increasing prosocial equivalent behaviour) would be more likely to be regarded as Satisfactory on the Satisfactory Behaviour Rating Guide.

4.2 Method

4.2.1 Participants

4.2.1.1 Custody-Based Intensive Treatment (CUBIT) program. This study was conducted within the CUBIT program at Long Bay Correctional Complex in Sydney, Australia. CUBIT is a 40-bed residential treatment program for high risk sexual offender prisoners with a current and/or prior sexual offence against children and/or adults. Treatment is group-based, with 10 offenders in each group. Offenders engage in seven hours of group therapy each week, for approximately eight to 10 months. Some additional interventions are provided individually to complement the group work if

required (e.g., to increase understanding, or if problems arise that would be more effectively processed in a one-to-one forum).

The eligibility criteria to enter the program are as follows:

- Offenders must meet the criteria of a sexual offender as per the Corrective Services New South Wales (CSNSW) definition: any convicted offender whose current offences include one of sexual violence; any convicted offender whose history of offences includes a conviction for sexual violence; any convicted offender who informs Corrective Services NSW that he/she has committed acts of sexual aggression (whether they be officially known or not, e.g., includes 'no billed' charges, in which prosecution does not proceed due to insufficient evidence); or, any convicted offender whose offence(s) are determined to have entailed an underlying sexual motivation, e.g., a violent offence with a sexual motivation;
- Offenders must at the time of referral be serving a custodial sentence; and,
- Offenders must be adult males.

If offenders are eligible for a treatment program, a Senior Psychologist within CSNSW Sex Offender Programs will assess them for program suitability. Suitability for CUBIT is based on offenders' STATIC-99R risk category in combination with a STABLE-2007 dynamic risk/needs assessment. This combined static and dynamic risk assessment must place them in at least a Moderate-High risk level. Alternatively, offenders assessed as having responsivity issues that might benefit from a residential program may be considered for entry into this program (e.g., mental illness, low motivation).

In general, many offenders commence CUBIT after their Earliest Release Date (ERD). That is, they are already in their parole period but have not yet been granted parole and, therefore, have not been released from custody.

4.2.1.2 Participant selection. The candidate is a psychologist in CUBIT and facilitated a treatment group throughout the data collection period. Therefore, for the current study, offenders who were in the candidate's treatment group were excluded. All other offenders admitted to CUBIT were eligible for participation in the research. A description of the ethical considerations within this study is provided in section 4.2.4.

A total of 36 offenders commenced a CUBIT treatment group between 1 January 2014 and 15 July 2015. The data collection commencement date was selected as it was the first possible date after ethics approval was granted for the research. The date for the

cessation of recruitment was selected to allow for the completion of data collection in a reasonable time frame within the candidate's enrolment. Fourteen offenders were excluded because they were admitted to the candidate's treatment group, leaving 22 eligible participants. After they were approached in relation to the research (see section 4.2.3), three offenders did not consent to participate; two reported their preference to focus solely on treatment, and one cited poor memory and anxiety. Therefore, 19 participants were included in the sample.

Demographic characteristics and other information for these 19 participants are described below. Following presentation of this information, an outline is provided regarding the way in which the available data was restricted due to some participants' withdrawing from some parts of the study prematurely.

4.2.1.3 Participant information. Participants were aged from 21 to 64 years ($M = 39.47$, $SD = 11.35$) at the commencement of treatment. Four participants identified as Aboriginal, one participant identified as Algerian and the remaining 14 were Caucasian Australian. Participants' sentence lengths varied between 36 months and life. The time between participants' Earliest Release Dates and treatment commencement varied. One participant commenced treatment 226 days prior to his Earliest Release Date. The remaining participants were in their parole periods (beyond their Earliest Release Date and eligible for parole but prior to their sentence expiration) upon treatment commencement, by between 105 and 536 days ($M = 204.44$, $SD = 117.39$).

The 19 participants were in treatment for between 26 and 69 weeks ($M = 45$, $SD = 10.80$). The participant who was in treatment for the shortest time had a Court Based Release date due to his sentence length (three years); this Court Based Release date meant that he was court ordered to be released from custody on his Earliest Release Date. Two participants had previously completed CUBIT; one was incarcerated with a life sentence and had not been released in between his two periods of treatment, while the other had completed treatment on two prior occasions, but subsequently reoffended on both occasions, which led to additional periods of incarceration. Three participants received suspensions from treatment during the course of the research for behavioural misconduct (e.g., physical aggression); one of these participants received multiple suspensions at various stages of treatment.

Participants' STATIC-99R scores ranged from one to nine ($M = 4.95$, $SD = 1.99$). As seen in Table 15, these STATIC-99R scores placed participants in the Low to High risk categories, with the majority (52.63%) in the Moderate-High risk category. As

previously indicated, participants whose STATIC-99R scores placed them in a risk category lower than Moderate-High were found suitable for CUBIT based on a combined risk assessment that included consideration of dynamic risk factors.

Table 15. *Participant Risk Level and Offence Type*

Participant Information	Number of Participants (%)
STATIC-99R Risk Category	
Low (-3 - 1)	1 (5.26)
Low-Moderate (2 - 3)	2 (10.53)
Moderate-High (4 - 5)	10 (52.63)
High (6+)	6 (31.58)
Current Offence Victim Type	
Child	12 (63.16)
Adult	7 (36.84)
Prior Offence Type	
None	4 (21.05)
Non-Sexual only	6 (31.58)
Sexual only (child victim)	2 (10.53)
Sexual only (adult victim)	0 (0)
Sexual (child victim) and Non-Sexual	5 (26.32)
Sexual (adult victim) and Non-Sexual	2 (10.53)

Most participants' current sexual offence victim was a child victim (63.16%) and most participants had prior criminal histories (78.95%); approximately half had a history of prior sexual offending (47.37%).

4.2.1.4 Data availability. Although 19 participants commenced, there was variability in their completion of all measures and the length of their participation in the study over the course of their treatment also varied (see Table 16).

Table 16. *Number of Participants with Available Data for Each Change Measure*

Change Measure	Number of participants (%)
Behavioural checklists	
All self-report and observer data	10 (52.63)
Some self-report and all observer data	9 (47.37)
Some self-report and observer data (premature discontinuation of both checklists)	2 (10.53)
Pre- and post-treatment psychometric measures	18 (94.74)
Treatment Gain scale	18 (94.74)
Satisfactory Behaviour Rating Guide	18 (94.74)

Note. $N = 19$.

Table 17 outlines information about each participant's engagement in the research, including (where relevant) their withdrawal of consent, descriptions of missing data, the frequency ranges for reported behavioural engagement, and similar information for custodial officer observations. Eight of the 19 participants withdrew consent for the self-report behavioural checklists during the course of the study. Three of these eight participants agreed to complete these again, at a later date, as they progressed through treatment; one only completed one additional checklist before he again withdrew from the study. One completed checklists in the final month of treatment, and one completed checklists in the final two months of treatment. The remaining five participants' length of participation in the research varied between one and 21 weeks. The reasons participants provided for their withdrawal from the research included: no perceived personal gain from participation; inconvenient being requested to complete a weekly checklist; and, it was "too stressful". One participant was suspended from treatment for a period of time subsequent to the fifth time point in data collection; this marked the end of his behavioural checklist completion through both self-report and custodial officer observations. One participant was repeatedly suspended from the program. He withdrew consent for self-report data after the first week of participation; however, custodial staff continued to complete observer rated behavioural checklists. Due to time constraints within the research, and the unknown length of time this participant would remain in treatment, the decision was made to conclude data collection for this participant prior to his treatment completion.

Table 17. *Behavioural Checklist Descriptive Information*

Participant	Participant remained in research throughout their treatment in CUBIT	Number of time points participants completed self-report (starts at Time 0)	Number of time points missing from self-report	Self-reported behaviour frequency range	Number of possible time points completed by custodial officer	Number of time points missing data from custodial officers	Custodial officer reported behaviour frequency range
1	Yes	31 (suspended from treatment for 4 weeks and removed from the wing, no data collected during this time)	4	0-50+	31	8	0-25
2	No	12 (up to week 11, withdrew, attempted once more in week 52)	43	0-100	62	15	0-40
3	Yes	53	0	0-50	53	25	0-30
4	Yes	47 (started just after Time 1 rather than between Time 0 and 1)	0	0-100+	47	13	0-30
5	No	13	37	0-12	50	25	0-30
6	No	3	37	0-100	40	15	0-40

7	No	4	No data collected after Time 5	0-50	No data collected after Time 5	No data collected after Time 5	0-45
8	Yes	37 (removed from the wing before he could complete his final checklist)	1	0-50	38	9	0-35
9	Yes	26	0	0-100+	26	8	0-25
10	No	9 (4 at the beginning, 1 in the middle, 4 at the end)	28	0-50	36	21	0-30
11	Yes	32	0	0-100	32	8	0-25
12	No	14 (did not complete at Time 2, withdrew after Time 5, back in at Time 29)	24	0-31	38	12	0-30
13	Yes	39	0	0-60+	39	7	0-15
14	Yes	60	0	0-35	60	13	0-30
15	Yes	46 (no checklist completed after treatment but had one from custodial officer)	1	0-50	46	15	0-15
16	No	21	28	0-17	49	13	0-20

17	Yes	45	0	0-100+	45	9	0-15
18	No	1 (Time 0, completed at Time 1 but invalid)	48	0-45	49	13	0-15
19	No	0 (attempted 1 checklist but withdrew without completing all items)			Collected custodial checklists up to Time 19 – time taken in treatment due to difficulties and treatment suspensions led to researchers deciding to cease data collection		0-45

Note. Self-reported behaviour frequency range = the overall range of frequencies each participant reported across behaviours on the checklist within the data collection period (behavioural checklists were completed weekly, such that participants indicated the frequency with which they engaged in each behaviour on the checklist in the preceding week); Custodial officer reported behaviour frequency range = the overall range of frequencies custodial officers reported observing across behaviours on the checklist within the data collection period (behavioural checklists were completed weekly, such that custodial officers indicated the frequency with which they observed participants engaging in each behaviour on the checklist in the preceding week).

4.2.2 Materials

4.2.2.1 Behavioural checklists. The behavioural checklist was developed using the results obtained in Study 1, the psychologists' survey. In this survey, psychologists identified behaviours they would expect to observe if each of 24 risk factors was relevant for a sexual offender and if each of these risk factors was no longer relevant for a sexual offender. These 24 risk factors were drawn from two risk assessment tools: Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003) and Violence Risk Scale: Sexual Offender version (VRS:SO; Wong et al., 2003). The RSVP is a widely used risk assessment tool for sexual offenders, both in Australia (e.g., Corrections Victoria) and internationally (e.g., Scotland; see Darjee, Russell, Forrest, Milton, Savoie, Baron, Kirkland, & Stobie, 2016). There is content overlap between the RSVP and several other professional guidelines for assessing sexual violence risk (see RSVP Protocol; Hart et al., 2003). Each of the 17 dynamic risk factors from the RSVP was used within the behavioural checklists. The five static risk factors were excluded, as behaviours related to these risk factors would not be expected to change over time due to their basis on unchanging historical factors (Hanson & Harris, 2001). Two risk factors from the VRS:SO were included as additional factors: Interpersonal Aggression and Offence Planning. The candidate and supervisors considered these constructs to be absent from the RSVP.

Each behavioural checklist (see Appendix C) consisted of 38 behaviours indicative of manifest dynamic risk factors (two behavioural manifestations for each risk factor) and 38 behaviours identified as prosocial equivalents to the dynamic risk factors (two behavioural manifestations of each prosocial equivalent that were considered inconsistent with the risk factor, such that the presence of these prosocial behaviours were considered indicative of positive behaviour within the domain). Written instructions were provided on the checklist. For each checklist item, examples of more specific behaviours were provided. These examples did not represent an exhaustive list; they were included as cues to assist participants and custodial officers in their identification of behaviours. A column was provided on the checklist, in which a frequency could be written. Additionally, as an alternative, a scale from zero to 50 was provided, with markings in intervals of five (measured in millimetres on the page). A mark could be placed at any point on the scale as a representation of an approximate frequency. These numbers referred to the number of times the participant engaged in the specified behaviours over the week prior to completion of the checklist. For participants

who marked a point on the scale between the intervals, a ruler was used to determine the point at which the marking was made, to gain a numerical frequency value (e.g., if a mark was placed in between two points on the scale, and measured seven millimetres from point zero, it was assigned a value of seven).

4.2.2.2 Psychometric measures. Consent to engage in CUBIT includes completion of pre- and post-treatment psychometric measures. The measures included in this pre- and post-treatment testing were determined by Corrective Services NSW. The purpose of these measures is to evaluate changes made during treatment, in participants' attitudes and beliefs surrounding areas such as sexual offending, relationships and sexual interests. As a primary aim of this study was to compare different measures of change, this section will also include some discussion of change-related psychometric information, and the use of these measures to determine changes over time, from prior research.

There are two social desirability measures included in this testing battery; the Paulhus Deception Scales (PDS; Paulhus, 1998) was not included in the current research as it was unavailable for some participants. All other psychometric measures from this pre- and post-treatment testing battery were included in the current research.

4.2.2.2.1 Bumby Rape Scale. This measure (BRS; Bumby, 1996) contains 36 items and assesses cognitive distortions related to sexual violence against women. Items are scored on a Likert-type scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). Responses are summed to gain a total score, with higher scores indicating beliefs that are more consistent with deviant sexual behaviour. Bumby (1996) reported excellent internal consistency with Cronbach's $\alpha = 0.96$, while the test-retest reliability after a two-week interval was $r = 0.86$. Previous research has shown that sexual offenders with offences against adults scored higher on the BRS than did non-sexual offenders (Bumby, 1996). However, Blumenthal, Gudjonsson and Burns (1999) reported there was no difference between sexual offenders who offended against adults and those who offended against children on this scale. That is, sexual offenders who offended against adults endorsed no more positive attitudes associated with rape compared with child offenders.

A previous study provided meta-analytically determined normative sample mean scores, which included three samples of university students (Nunes, Petterson, Hermann, Looman, & Spape, 2016). Using this normative sample, Nunes et al. (2016) demonstrated that sexual offenders tended to already be in the functional range (i.e.,

within the range of the functional population) on this measure prior to treatment. This result suggests that the measure might not be sensitive to sexual offenders' problematic beliefs; potential reasons for this insensitivity are that the beliefs it is measuring are not relevant for the sample, or that the measure is vulnerable to offenders providing socially desirable responses.

Nunes et al. (2016) reported that the majority of those offenders who had dysfunctional pre-treatment scores on the BRS demonstrated significant improvement through treatment (62.3% 'recovered' to the functional range and 10.5% 'improved' while remaining in the dysfunctional range). Despite these changes through treatment, Nunes et al. (2016) indicated that there was no significant relation between positive change on the BRS and subsequent sexual recidivism, although scores tended to be higher for those who reoffended than those who did not. Nevertheless, the results of the change analyses suggested that sexual recidivists' scores generally improved slightly more than offenders who did not reoffend (Nunes et al., 2016).

4.2.2.2.2 Bumby Molest Scale. This measure (BMS; Bumby, 1996), which consists of 38 items, assesses cognitive distortions related to sexual offending against children. Items are scored on a Likert-type scale ranging from 1 (*strongly disagree*) to 4 (*strongly agree*). Responses are summed to gain a total score, with higher scores indicating beliefs that are more consistent with deviant sexual behaviour. Bumby (1996) reported excellent internal consistency with Cronbach's $\alpha = 0.97$ and test-retest reliability after a two-week interval, $r = 0.84$. Researchers have also validated its use with a sexual offender population (e.g. Blumenthal et al., 1999; Marshall et al., 2003). Blumenthal et al. (1999) reported that sexual offenders who offended against children were more likely to endorse cognitive distortions in this measure than did those who offended against adults ($p < 0.001$). A previous study provided meta-analytically determined normative sample mean scores, which included one sample of university students and one community sample (Nunes et al., 2016).

Consistent with the BRS results, Nunes et al. (2016) reported in their meta-analysis that the majority of sexual offenders already had functional pre-treatment BMS scores, but the majority of those with dysfunctional pre-treatment scores made significant gains through treatment. As noted for the BRS, functional pre-treatment scores on the BMS may suggest insensitivity to sexual offenders' beliefs. Similar to their findings with the BRS, sexual offenders who reoffended generally had higher scores on the BMS. However, positive changes on the BMS did not significantly relate

to recidivism. Nevertheless, the same trend was observed as for the BRS, such that there was a trend for positive changes to predict recidivism (Nunes et al., 2016).

4.2.2.2.3 Coping Inventory for Stressful Situations. This measure (CISS; Endler & Parker, 1990) assesses general coping styles. It contains three scales, which assess Task-Oriented, Emotion-Oriented, and Avoidance-Oriented coping strategies. The Avoidance scale is further divided into two additional subscales, Distraction and Social Diversion. There are 48 items overall, with 16 items per scale. All responses are provided on a five-point Likert-type scale. Endler and Parker (1990) did not report the internal consistency of the measure, but reported the test-retest reliability for each of the subscales: Task ($r = 0.73$); Emotion ($r = 0.68$); Avoidance ($r = 0.55$); Distraction ($r = 0.51$); and, Social Diversion ($r = 0.54$). This measure has previously been used with sexual offenders whose offences were against children (Serran, Moulden, Firestone, & Marshall, 2007). These authors reported that in this group of sexual offenders, compared with a waiting list control group, there was an increase in task-oriented coping through treatment; however, there was no significant change over time in emotion-oriented coping or the Distraction subscale of Avoidance-Oriented coping. There was an increase in the use of Social Diversion.

4.2.2.2.4 Coping Using Sex Inventory. This measure (CUSI; Cortoni & Marshall, 2001) has 16 items and was developed for use with sexual offenders.. The items relate to both consenting and non-consenting sexual behaviour with children and adults in four areas: fantasies; masturbation; pornography use; and, actual sexual behaviour with a partner. Responses are provided on a five-point Likert-type scale, indicating the frequency with which the behaviour is engaged in when a stressful or difficult situation is encountered. The internal consistency of the overall scale has been reported as Cronbach's alpha = 0.88 (Cortoni & Marshall, 2001). No test-retest reliability data is available. Cortoni and Marshall (2001) demonstrated that sexual offenders made significantly more use of sexual coping strategies than did non-sexual violent offenders ($p < 0.01$). The test developer was contacted for the current research, to ascertain whether data exists for non-offenders; the response was that at this stage no such data are available (F. Cortoni, personal communication, April 29 2016). No prior research was available regarding the use of the CUSI as a change measure.

4.2.2.2.5 Wilson Sexual Fantasy Questionnaire. This measure (WSFQ; Wilson, 1978) is a 40-item questionnaire, which assesses four types of sexual fantasies: Exploratory, Intimate, Impersonal and Sadomasochistic. Responses are provided on a

five-point Likert-type scale (*never to regularly*), for questions assessing sexual desires, preferences and activities. Scores are summed on each subscale such that the higher the score on each subscale, the more fantasies a respondent endorsed. Skovran, Huss, and Scalora (2010) reported the internal consistency for each subscale: Exploratory (Cronbach's $\alpha = 0.84$); Intimate (Cronbach's $\alpha = 0.92$); Impersonal (Cronbach's $\alpha = 0.77$); and, Sadomasochistic (Cronbach's $\alpha = 0.81$). This measure was validated on child sexual offenders (e.g., Baumgartner, Scalora, & Huss, 2002; and, Allan, Grace, Rutherford, & Hudson, 2007). Both these studies provided some evidence of sexual offenders scoring lower than the non-offender population on two subscales, Intimate and Exploratory. However, Allan et al. (2007) demonstrated that the WSFQ loaded on a single factor and was positively correlated with sexual recidivism. The test developer was contacted to determine the availability of test-retest reliability. To date, no such data are available (G. Wilson, personal communication, 1 October 2016 and 20 March 2018).

Some research has been conducted in relation to the WSFQ as a measure of change. Treatment has been found to have an effect on a group of reconvicted sexual offenders' use of impersonal sexual fantasies and a slight increase in the use of sadomasochistic fantasies in comparison with the non-reconvicted group's slight decline (Bakker, Hudson, Wales, & Riley, 1998). On the contrary, prosocial change on the Impersonal and Sadomasochistic subscales have been associated with increased recidivism (Hudson et al., 2002).

4.2.2.2.6 Revised University of California (UCLA) Loneliness Scale. This measure (LS-UCLA; Russell, Peplau, & Cutrona, 1980) assesses emotional loneliness. It contains 20 items, which target respondents' beliefs about the extent to which they have meaningful relationships, have people close to them, or are lonely. Responses are provided on a four-point Likert-type scale, with higher scores reflecting greater loneliness and fewer close and meaningful relationships. Internal consistency has been reported as excellent, Cronbach's $\alpha = 0.95$, and the test-retest reliability has been measured at $r = 0.73$ (Olver et al., 2014). The scale has been used in research with sexual offenders (e.g., Beech, 1998; Fisher, Beech, & Browne, 1999; Hudson et al., 2002). Fischer et al. (1999) found that men who sexually offended against children differed significantly from non-offenders on this scale. Beech (1998) found that high-deviance sexual offenders demonstrated significantly higher emotional loneliness on this measure than did low-deviance sexual offenders.

Prior research has demonstrated that the LS-UCLA was not susceptible to change over time in treatment (Hudson et al., 2002). Consistent with these findings, Barnett, Wakeling, Manderville-Norden, and Rakestrow (2013) reported that the majority of participants in their study were unchanged on the LS-UCLA pre- to post-treatment. Further, improvements on this measure were not related to reduced recidivism. Nunes et al. (2016) reported their findings at both the group- and individual-level. They reported that at the group-level, there was a significant improvement pre- to post-treatment. At the individual-level, 53.7% of participants were in the dysfunctional range pre-treatment and 37.7% of these participants were classified as having recovered (i.e., were in the functional range) post-treatment.

4.2.2.2.7 Social Intimacy Scale. This measure (SIS; Miller & Lefcourt, 1982) is a 17-item scale of intimacy and loneliness in close relationships. Responses are provided on a 10-point Likert-type scale. Higher scores indicate higher levels of perceived intimacy. Miller and Lefcourt (1982) reported internal consistency between Cronbach's $\alpha = 0.86$ to 0.91 , and test-retest reliability between $r = 0.84$ (over a two month interval) and 0.96 (over a one month interval). Prior research has used the SIS to compare levels of intimacy between offender-types (i.e., different groups of sexual offenders such as adult vs child, and wider groups of offenders such as sexual and non-sexually violent) and between offenders and non-offenders (e.g., Looman, Abracen, DiFazio, & Maillet, 2004; Marshall, Champagne, Brown, & Miller, 1998; Seidman, Marshall, Hudson, & Robertson, 1994). These studies demonstrated that the SIS can be used to distinguish between offender types and between offenders and non-offenders.

In relation to using the SIS as a change measure, Nunes et al. (2016) reported that at the group-level, participants did not significantly improve pre- to post-treatment. At the individual-level, 37.8% of participants were in the dysfunctional range pre-treatment and 27.3% of these participants were classified as having recovered (i.e., were in the functional range) post-treatment.

4.2.2.2.8 Social Self Esteem Inventory. This measure (SSEI; Lawson, Marshall, & McGrath, 1979) was designed to assess self-worth in social situations. It contains 30 items, reported on a six-point Likert-type scale (*unlike me* to *exactly like me*). Lower scores reflect lower levels of social self esteem. Lawson et al. (1979) reported the test-retest reliability at $r = 0.88$ and internal consistency was good, Cronbach's $\alpha = 0.60$. Previous research has validated the use of the SSEI with sexual offenders, such that it successfully distinguished between sexual offenders and comparison participants (see

Marshall, Anderson, & Champagne, 1997). Research on the SSEI as a measure of change has demonstrated that self esteem was significantly increased over treatment as assessed through this measure (Marshall, Champagne, Sturgeon, & Bryce, 1997).

4.2.2.2.9 Marlowe-Crowne Social Desirability Scale. This measure (MC-SDS; Crowne & Marlowe, 1960) assesses the extent to which respondents tend towards positive self-presentation, which is of particular concern in forensic populations. For instance, prior research has evidenced the tendency for sexual offenders to provide socially desirable responses within self-report measures (e.g., Haywood, Grossman, Kravitz, & Wasyliv, 1994; Langevin, 1991; Lanyon, Dannenbaum, & Brown, 1991). The MC-SDS includes 33 true-false statements. According to the authors, for inclusion in the scale, the items had to meet the criteria of cultural approval but improbable occurrence, and have “minimal pathological or abnormal implications” if responded to in either direction (Crowne & Marlowe, 1960, p. 350). Internal consistency has been reported as ranging from acceptable to excellent, Cronbach’s $\alpha = 0.72$ to 0.96 (Ballard, 1992; Crowne & Marlowe, 1960; Fischer & Fick, 1993; Loo & Thorpe, 2000; Reynolds, 1982). One-month test–retest reliability was $r = 0.89$ (Crowne & Marlowe, 1960).

The MC-SDS has been widely used with non-forensic populations, for example as a measure of defensiveness in medical research (Deshields, Tait, Gfeller, & Chibnall, 1995; Helmers et al., 1995; Mann & James, 1998), to detect positive impression management in a sample of university students, some of whom endorsed psychopathic personality characteristics (Edens, Buffington, Tomicic, & Riley, 2001), and within general student and community samples (e.g., Ballard, 1992; Fraboni & Cooper, 1989; Loo & Thorpe, 2000; Reynolds, 1982; Strahan & Gerbasi, 1972; Vella-Brodrick & White, 1997).

Andrews and Meyer (2003) further established its use within the forensic context, with offenders including those who engaged in: physical abuse of children; child neglect; domestic violence; child sexual abuse; pre-trial competency defendants; disability examinees; and, various individuals within civil court proceedings. The MC-SDS has been used in studies of sexual offenders, with results suggesting that social desirability as measured by this scale was associated with reduced deviance and risk of recidivism (e.g., Allan, Grace, Rutherford, & Hudson, 2007).

The MC-SDS was used in this study to determine the veracity of participant’s self-report. A score of 26 or above (one standard deviation above the mean for a

forensic sample) was used as a basis for exclusion, consistent with prior research in a forensic population (Andrews & Meyer, 2003).

4.2.2.3 Treatment Gain: Short Scale. The Treatment Readiness Responsivity Gain Scale: Short Version (TRRG:SV; Serin, Kennedy, & Mailloux, 2005) is a clinical rating scale designed for use with offenders referred to correctional programs. It has three domains: Treatment Readiness, Treatment Responsivity, and Treatment Gain. The Treatment Gain: Short Scale was of primary interest in the current research. Therefore, this scale was the only one used from the TRRG:SV.

The Treatment Gain scale is a post-treatment rating scale, completed by the participant's treating therapist. Its purpose is to capture overall pre- to post-treatment gain. It provides an overall estimate of an offender's performance at the completion of treatment, rather than a measure of change in specific treatment-targets. It consists of eight items: evidence of increased skills from program; disclosure in program; application of knowledge; application of skills; depth of emotional understanding of program content; appropriateness of behaviour in group; participation; and, therapeutic alliance. These items are rated on a scale of 0 (*poor*) to 3 (*very good*), with descriptions within each item to assist reliable scoring. The maximum score is 24, with higher scores considered indicative of greater treatment gains. Sowden (2013) demonstrated its validity as an overall measure of change, through its relationship with the VRS:SO. Of note is that change on all VRS:SO scales have been shown to be associated with change, such that they significantly predicted recidivism in a sample of sexual offenders (Olver, Wong, Nicholaichuk, & Gordon, 2007). Therefore, Sowden's (2013) findings that the Treatment Gain scale was positively correlated with the VRS:SO Dynamic ($r = 0.67, p < 0.01$), VRS:SO Sexual Deviance ($r = 0.51, p < 0.01$), VRS:SO Criminality ($r = 0.49, p < 0.01$), and VRS:SO Treatment Readiness ($r = 0.57, p < 0.01$) provide support for the validity of the Treatment Gain scale as a measure of change. Further research has reported that the inter-rater reliability for the Treatment Gain scale is 0.95 (Sowden & Olver, 2017). The Treatment Gain scale had a strong correlation with the VRS:SO Change score, $r = 0.66$.

4.2.2.4 Satisfactory Behaviour Rating Guide (SBRG). The development of the SBRG (Daffern et al., 2014) was described in Chapter 3. The SBRG comprises four domains: No violence during institutionalisation; Not involved in substance abuse or like behaviour; Has a positive work ethic in relation to employment, education and/or rehabilitation; and, Responds well to supervision and direction. There are 10 sub-scales

related to these domains, ranging from prosocial to negative behaviour: Behaviour is prosocial and cooperative; Has a good work ethic; Peer/associate preference; Emotional control; Violence during institutionalisation; Substance abuse or like behaviour persists; Relationships with significant others; Interested in and maintaining community supports; Working towards increased freedom; and, Sexual aggression during institutionalisation.

In the current study, each participant was rated on the SBRG at their completion of treatment. The treating therapist and a custodial officer each completed the SBRG in relation to the participant's behaviour whilst housed in CUBIT. An ultimate rating was also provided, as to whether the participant's behaviour was considered Satisfactory or Unsatisfactory.

4.2.3 Procedure

The candidate approached each eligible participant after the primary therapist had sought initial contact with the offender and oriented him to the CUBIT community. The candidate introduced herself to the offender and a brief description of the research was verbally provided, after which a written information sheet was provided for the offender to read. All offenders who were approached for the research reported confidence in their reading ability. Offenders were given the option of reading the information sheet at the time of the initial discussion, or they could read in their own time. They were informed that participation in the research would have no impact on their treatment and that no data collected for the purpose of this research would be provided to others within CUBIT (either staff or offenders). Offenders who verbally agreed to participate after reading the information sheet were provided with an opportunity to ask the candidate questions and discuss any concerns about the research. Subsequent to this discussion, they were provided with a written consent form to read and sign. They were informed that they could withdraw from the research at any time, including prior to commencement of their participation, with no detriment to their treatment.

The time frame between participants consenting to engage in the research and completing their first checklist ranged between one and 12 days ($M = 4$, $SD = 3.24$). The time frame between participants completing their first checklist (Time 0) and commencing treatment ranged between zero and eight days ($M = 2$, $SD = 2$). The time frame between participants consenting to engage in the research and commencing treatment ranged between one and 20 days ($M = 6$, $SD = 4.38$). This variation was due to factors such as primary therapist availability, lock downs (offenders were

occasionally locked in their cells for operational reasons), and offender movements between gaols.

Participants were informed about the day they would be required to complete their first behavioural checklist. They completed the checklist on the same day each week for the duration of their treatment, unless they withdrew consent prior to the completion of treatment. Participants were divided between weekdays, such that on each weekday there were between two and three participants scheduled to complete their checklists. It was intended that this allocation of participants to days of the week would decrease the workload for custodial officers, who also completed the behavioural checklists on the same days as did the participants. There were times when the researcher was unable to provide participants with the checklist on the allocated day. Reasons included: the researcher was unavailable; all offenders were locked in their cells for operational reasons; and, a participant was on escort out of the wing (e.g., at a medical or legal appointment). If the situation was identified in advance, participants were provided with the checklists up to six days early and were asked to complete the checklists on the relevant day. If the situation was unforeseen, participants were provided with the checklist at the next opportunity; generally, the following day.

Prior to completion of the first checklist, the researcher discussed the checklist in detail with each participant. Although instructions were written on the checklist, these instructions were also presented verbally, such that additional details could be provided and each participant's level of understanding could be gauged. Examples were provided to participants for items that were more complex or abstract (e.g., "Difficulty engaging in treatment"). The process for estimating frequencies of engaging in behaviours was explained. Strategies for accurately completing the checklists were discussed (e.g., creating a tally to remember the number of times a behaviour was engaged in throughout the week). In cases when participants had difficulty understanding the requirements of the checklist, there was opportunity for ongoing discussion with the candidate; additionally, treatment progression facilitated the process of gaining further awareness and understanding of the relevant behaviours, since these are common areas of discussion within group treatment. Participants were encouraged to mark items on the checklist if they were uncertain how to respond or which behavioural examples might be relevant. These items were further discussed with the participant when the checklist was collected from him.

After participants completed each checklist, the candidate considered their responses. If the reported frequencies appeared unrealistic (e.g., were much higher or lower than expected or did not match patterns observed throughout previous checklists), the participant's understanding of the instructions was gleaned and the candidate clarified the responses with the participant. If this process reflected a participant's poor understanding of the requirements, further direction and clarification was provided. It was important for the candidate to be flexible in the provision of instructions and to ensure individual responsivity factors were considered, in the same way they would be within treatment (e.g., in relation to participants' cognitive capacity and ability to generalise behavioural examples to each item). Despite discussion and clarification, some participants continued to provide seemingly unrealistic frequencies (e.g., engaging in behaviours 100 times over the week). Through discussion with participants, it was ascertained that these responses were provided as a result of their continued interpretation of the checklist as a Likert-type scale (e.g., never, sometimes or always engaging in the behaviours).

If participants expressed their declining motivation to engage in the research, or if the candidate perceived their motivation to be waning through comments made when the checklists were collected, time was spent discussing this with them and attempts were made to assist them as required. Through ongoing discussion with the participants, it appeared that many were influenced by other CUBIT treatment participants' negative attitudes towards treatment and comments about the research. Once a participant stated that he was considering his withdrawal from the research, he was asked if he would be willing to provide the candidate permission to follow up with him the following week to determine his level of motivation. If this permission was provided, further discussion was held the following week. Participants who indicated they did not wish to continue were asked if they would consider recommencing in the last month of treatment. This procedure was implemented to facilitate a comparison between the first and last months of treatment, in order to assess change throughout treatment.

Attention was consistently focused on ethical considerations within the research context (see section 4.2.4). A balance was required between providing participants with support and encouragement to continue their involvement in the research, such as discussing and alleviating their concerns and providing them with an understanding of the potential benefits they may gain (e.g., greater insight into their ongoing behaviour such that they might gain additional benefit from the treatment process), and ensuring

they did not experience perceived pressure to engage in the research if they wished to withdraw their consent.

Custodial officers who were regularly rostered on within CUBIT were asked to complete the behavioural checklists on the same days as the participants completed the checklists. Within the area of the correctional centre in which CUBIT operates, custodial officers are assigned a “regular” role, for instance in a particular wing. There were generally four custodial officers whose daily shifts were regularly in CUBIT. At the time of data collection for this study, there were two custodial officers rostered in CUBIT in the morning and two in the afternoon. These officers attended the weekly staff meetings and community meetings, and had a thorough understanding of the program and knowledge of the treatment participants. During the data collection period, there was a rotation of CUBIT’s regular custodial officers, such that there were four new regular custodial officers.

Prior to commencing data collection, the candidate introduced the justification for the research and the nature of expectations to custodial staff within CUBIT. Additionally, the potential benefits of the research for the therapeutic community context were outlined. Custodial officers raised some concerns regarding the process, such as the burden on their time. These concerns were discussed and it was agreed that there would be flexibility such that if operational demands were particularly high, they could complete the checklist at another time (either the same day or the following day). They were aware that completion of the checklists was voluntary; as the data collection period progressed, some custodial officers expressed their desire to discontinue involvement in the process. For those officers who were motivated to assist in the process, ongoing discussions were held over time, to ensure the checklists were not creating unnecessary burdens on them. At times, clarification was provided regarding specific items on the checklists. Throughout the data collection process, consistent contact was maintained with those custodial officers who assisted with the research. The process described above in relation to engaging custodial officers, was repeated as necessary, with the rotation of regular custodial officers in and out of CUBIT.

If there was no regular officer available on the relevant day, the data was not collected. If a regular officer was available within a reasonable time period after this day, for example, within the next day or two, the checklist was completed at this time. For all participants, there were weeks in which no regular custodial officers were rostered on in the unit. In addition, as with participant checklists, there were

approximately five times when the researcher was unable to disseminate the checklists to custodial officers (e.g., due to time away from work). Where possible, alternative arrangements were made, for instance another therapist was asked to provide officers with checklists. Overall, there were between seven and 25 time points per participant, for which there was no custodial officer data collected (see Table 17).

Consent was sought from participants who withdrew from the research prior to their treatment completion, for the continuation of data collection from custodial officers. Consent was also sought for the use of these participants' pre- to post-treatment psychometric measures, in addition to the staff completion of the Treatment Gain scale and SBRG. All participants consented to these parts of the procedure.

At the conclusion of treatment, participants and custodial officers completed one final behavioural checklist. The participant's treating therapist completed the Treatment Gain: Short Scale questionnaire and the SBRG, and a custodial officer completed the SBRG. The pre- and post-treatment psychometric measures were also collected.

4.2.4 Ethical Considerations

4.2.4.1 Conflict of interest. An ethical consideration of this research was the candidate's dual roles. Data was collected from participants engaged in the treatment program in which the candidate was also a clinician. Consequently, there was a power imbalance between the candidate and potential research participants because the candidate was in a position of authority as a clinician. Therefore, there was a possibility that potential research participants could feel coerced into providing consent to engage in the research, due to perceived potential detrimental effects on their treatment progress should they decide not to consent or the perception that they would be privileged in some way by participating. In addition, it raised the possibility that the candidate could be biased in conducting the research due to potential vested interests. These possible difficulties were overcome in the following ways: the research did not include participants from the candidate's treatment group; in the consent form provided to potential participants, it clearly stated that there would be no detrimental effects to the individual should he decide not to provide consent to participate in the research (i.e., the decision to participate or not would have no impact on the individual's relationship with CSNSW, or with their engagement in treatment or determination of release from custody); and, there were no treatment-related incentives provided to potential participants.

4.2.4.2 Consent. All participants were provided with a Participant Information Sheet regarding the study and on the basis of this information, each individual could make an informed decision about participation. All participants were required to provide written informed consent prior to commencing participation in the research. Participants were informed that they could withdraw consent at any stage of the research, with no detrimental effects (e.g., on their treatment, sentence, placement, classification, or parole considerations). No incentives were offered to participants.

4.2.4.3 Confidentiality. Information gathered through the research process was used exclusively for research purposes, such that only the candidate had access to potentially identifiable information gained through the research. Participants were informed that the information gained through the study would not be distributed to the treating psychologist unless there was an over-riding ethical obligation to do so (i.e., related to risk of harm to self or others, escape risk, or contravention of specific rules and regulations). All material was de-identified prior to data analysis.

4.2.5 Data Analysis

Both group- and individual-level change was measured. Group-level change was measured to ascertain the areas in which the sample made changes. Individual-level change was measured to gain a richer and more detailed representation of the processes by which offenders change and the ways in the disparate measures of change can be used to assist clinicians in their overall assessment of offenders' change.

Two data analysis processes were used for the behavioural checklists. Both these processes required the data to be recoded prior to analysis. The first recoding process allowed for determining whether there were differences between risk factors and between participants and custodial officers in the consistency with which behaviours were reported. It also allowed for comparisons between the two behavioural manifestations of each risk factor, which had been conflated for the comparisons between participants and custodial officers. Univariate comparisons were used for these categorical variables. The data were placed in contingency tables, which will be further described below. An overall binomial test was used, after which, the individual contingency tables were tested for significance using Fisher's exact tests. These analyses provided information about the overall patterns in the ways in which the risk factors manifested for offenders in treatment, both through self-report and custodial officer observations. Therefore, initial differences between self-report and observer data could be identified.

Subsequent to these analyses, the behavioural checklist data were again recoded to allow for analyses of behavioural change over time. Due to the quantity and complexity of the data collected, the risk factors were categorised into psychological domains from the RSVP. This process will be described in more detail below. Tests for skewness were conducted and transformations followed to reduce the skewness of the data. Group-level regression analyses were conducted to determine change over time within each of the psychological domains. The structure of the data was longitudinal. In order to account for the nested study design (i.e., multiple domains nested within participants and observers) and to most effectively capture participants' behavioural changes over time and detect individual differences in change, a multilevel regression model using Hierarchical Linear Modeling (Version 7) software (HLM; Raudenbush, Bryk, & Congdon, 2011) was used for the group-level analyses. Regression analyses were conducted across participant self-report, custodial officer observations, and between the two to compare the changes in responses over time. Individual-level analyses were conducted through use of a split file regression, for both self-report and custodial officer observations. SPSS 24.0 was used to analyse behavioural change on an individual-level.

Data for the pre- and post-treatment psychometric measures were entered into SPSS. Data screening was conducted to detect outliers. Missing data analyses were conducted for this dataset. Multiple imputation was used to replace the missing values. After testing for assumptions (such as normality), group-level analyses were conducted to gain information about the sample as a whole in relation to this measure of change. Repeated measures *t*-tests were used for those psychometric measures that met the assumptions at both pre- and post-testing; this was substituted with Wilcoxon signed rank tests where data violated the assumptions. Individual-level analyses were then conducted, to gain a more detailed understanding of the change process. That is, tests of reliable and clinically significant change were used to determine whether individual participants made changes from pre- to post-treatment in the relevant attitudes and beliefs measured within the psychometric measures.

Overall, the four measures of change (psychometric testing, behavioural checklists, Treatment Gain scale and SBRG) facilitated a comparison between the different change measures at the group-level, in addition to the individual-level. The individual-level analyses served to mirror a portion of the process clinicians are confronted with through facilitation of intervention. That is, throughout the provision of treatment and at its conclusion, clinicians must integrate various measures of change for

each of their clients in order to assess each client's level of change, which also provides important information for the risk assessment and release decision-making process.

The remainder of this section provides a more detailed description of the data preparation and analysis process.

4.2.5.1 Behavioural checklists. As previously indicated, the 76 items that formed the behavioural checklists were initially developed from 19 risk factors derived from the RSVP and VRS:SO. Each risk factor consisted of two negative behavioural manifestations and two positive behavioural manifestations. The first step in data analysis for the behavioural checklists was to assess whether there were associations between participants' tendency to self-report the presence or absence of behaviours compared with custodial officers' reports. To enable this process, the data was recoded. Prior research has outlined various methods by which behavioural data can be coded (e.g., MacLaren-Chorney, McMurtry, Chambers, & Bakeman, 2014). Data can be continuous or categorical. Both these coding methods were used in the current research; however, within the first step for data analysis, categorical coding was implemented. A common method is dichotomous coding in which categories are imposed on a continuum, such as behaviour coded as present or absent.

Despite a limitation of this method being the loss of specific information (Hammond, 2013), it was used in the current research as follows. If a behaviour elicited a response of zero (i.e., the behaviour was not engaged in/observed over the preceding week) every week throughout the time in which the participant was involved with the research, this behaviour received a code of 0. Conversely, if a behaviour was ever engaged in/observed across the time in which the participant was involved with the research, this behaviour received a code of 1.

To facilitate the comparisons, the items were collapsed into the 19 risk factors. Each pair of behavioural manifestations (i.e., the two negative behaviours associated with one risk factor and the two positive behaviours associated with one risk factor) was combined to gain a rating of 0 or 1. That is, if neither of the behaviours in the pair was engaged in/observed throughout the testing period, a code of 0 was given ('never' observed). If either one or both of the behaviours in the pair was engaged in/observed throughout the testing period, a code of 1 was given ('ever' observed). This coding facilitated the cross-tabulation of ratings, such that the associations between participant self-report and custodial officer ratings could be determined.

As the two behavioural manifestations of each risk factor were combined to facilitate these analyses (i.e., the two negative behaviours were combined and the two positive behaviours were combined when determining presence or absence), any differences in the likelihood of one of the behavioural manifestations being reported compared with the other manifestation were potentially concealed. Therefore, a final set of analyses was conducted to determine whether there was an association between the reported presence or absence of the two behavioural manifestations of each risk factor (both negative and positive behaviours) according to participants and custodial officers. This analysis could reflect whether one of the behaviours provided a greater amount of information than the other, within each risk factor; that is, it could assist in determining whether one of the two behaviours within the risk factor was more discriminative than the other by being reported more frequently.

These analyses were related to the findings in Study 1, in which some risk factors appeared more easily identifiable than others. That is, psychologists demonstrated greater knowledge of potential behavioural manifestations of some risk factors (e.g., Problems with Stress or Coping) compared with others (e.g., Sexual Deviance). This inherent difference between the risk factors may lead to differences in the patterns with which certain behaviours may be observed compared with other behaviours.

The binomial test of significance is a non-parametric form of probability test, which is used to examine the distribution of a single dichotomous variable in small samples (Wagner-Menghin, 2005). The current data met the binomial test assumption of independent observations, such that participants' responses were independent of each other. This test was run once for the comparison between participant and custodial officer reports of 'never' versus 'ever' reported, to determine whether there was an association between participants' and custodial officers' reports of the frequency with which participants engaged in the behavioural manifestations. That is, this test was used to investigate whether the frequency with which participants 'never' reported the behaviour was associated with the frequency with which custodial officers 'never' reported the behaviour and vice versa with 'ever', as opposed to an association between the frequency with which participants 'never' reported the behaviour and the frequency with which custodial officers 'ever' reported the behaviour, and vice versa.

Specifically, the consistency between participants and custodial officers in reporting each behaviour was determined by placing reported behaviours in 2 x 2 contingency tables with the absence or presence of a behaviour ('never' or 'ever'

reported) constituting the two levels for each factor. Table 18 provides an example, with the risk factor Problems with Substance Abuse (negative behavioural manifestations). For a particular behaviour, both the custodial officers and the participants ‘never’ classified it as being present (Quadrant Row 1, Column 1) or ‘ever’ classified it as being present (Quadrant Row 2, Column 2) where ‘never’ is indicated by “0” and ‘ever’ is indicated by “1”. The remaining two quadrants indicated disagreement between the participant and custodial officer concerning the absence or presence of the behaviour.

Table 18. *Problems with Substance Abuse Contingency Table*

		Problems with Substance Abuse (Negative) Custodial Officer	
		0	1
Problems with Substance Abuse (Negative) Participant	0	12	3
	1	2	0

For each contingency table, the number of participants for whom there was agreement between the custodial officers and participants could be added (add Quadrant Row 1, Column 1 to Quadrant Row 2, Column 2) and compared with the number of disagreements (add Quadrant Row 1, Column 2 to Quadrant Row 2, Column 1). Each of the 76 contingency tables then could be placed into two categories: (a) Where agreements exceeded disagreements or (b) where disagreements exceeded or were equal to agreements. Using a null hypothesis of 50% in the two categories, a binomial test was used to test whether over all of the behaviours, agreement significantly exceeded disagreement. In order to further investigate these overall results, each individual contingency table was tested for significance to determine whether there were any individual effects.

Subsequently, the binomial test was also run once for the comparison between each of the 19 risk factors’ two behavioural manifestations, to determine whether there was an association between the frequencies of the two behavioural manifestations, based on both participants’ and custodial officers’ reports. Each of the 38 individual contingency tables was also tested for significance to determine whether there were any individual effects.

For all the contingency table analyses, the expected count in some cells was less than five. Therefore, Fisher’s exact test was used rather than the Pearson chi square test

of association. Fisher's exact test does not provide coefficients with the p values; therefore, within section 4.3, only the p values are cited.

These analyses were conducted in order to provide an overview of the data. The following data preparation and analyses are related to determining whether there were changes in behaviours over time in treatment.

4.2.5.1.1 Coding weekly scores and creating domain scores. A key aim in relation to the behavioural checklists was to compare frequencies of behaviours over time in treatment. However, participants and custodial officers provided disparate observations for the behaviours in the checklists (i.e., the methods by which data was provided varied between individuals), which meant the reported frequencies were not easily compared. For instance, some participants appeared to count the frequency with which they engaged in behaviours, whereas others provided a number that represented a proportion of their time over the preceding week (e.g., either not at all or very often, translated to 0 or 100). As a result, some participants' data provided a pattern in which all negative behaviours were reported as 0 and all positive behaviours were reported as 100. Based on discussions with the participants, it was determined that this did not necessarily represent engagement in the behaviour 100 times during the week, but was reflective of their perception that they consistently engaged in that behaviour.

Consequently, several steps were taken to enable data analysis. First, the numerical range for each participant each week was determined, both for participant self-report and custodial officer observations (e.g., 0-100 or 0-25). These ranges were then divided into quartiles such that a number from one to four was assigned to each observation point based on the quartile in which it was contained (see Tavakoli, 2012). Subsequent to this process, each risk factor was assigned to one of the four dynamic risk factor domains from the RSVP, based on the domain from which the risk factor was initially determined (i.e., five risk factors in Psychological Adjustment, five risk factors in Mental Disorder, four risk factors in Social Adjustment and three risk factors in Manageability). The two risk factors, which were adapted from the VRS:SO were also categorised into the RSVP domains. In order to determine which domains these two risk factors should be classified into, the examples provided in the RSVP manual were referred to and matched to the risk factors. Table 19 displays the risk factors within the four domains. The quartile assigned to each of the risk factors within a domain was averaged such that there was one score (based on the quartiles) for each domain every week during which the participant was in treatment.

Table 19. *Risk Factors Categorised within Psychological Domains*

Psychological Adjustment	Mental Disorder	Social Adjustment	Manageability
Attitudes That Support or Condone Sexual Violence	Sexual Deviance	Problem with Intimate Relationships	Problems with Planning
Problems with Self-Awareness	Psychopathic Personality Disorder	Problems with Non-Intimate Relationships	Problems with Treatment
Problems with Stress or Coping	Major Mental Illness	Problems with Employment	Problems with Supervision
Problems Resulting from Child Abuse	Problems with Substance Abuse	Non-Sexual Criminality	Offence Planning
Interpersonal Aggression	Violent or Suicidal Ideation		

The positive and negative behaviours remained separated such that there were eight domain scores each week for each participant, both for participant self-report and custodial officer observations. The eight domain scores related to: Psychological Adjustment Negative; Psychological Adjustment Positive; Mental Disorder Negative; Mental Disorder Positive; Social Adjustment Negative; Social Adjustment Positive; Manageability Negative; and, Manageability Positive. The overarching psychological constructs subsumed by these domains are as follows: coping and problem solving (Psychological Adjustment); mental disorder; relationships and social influences (Social Adjustment); and, cooperation with supervision and self-management (Manageability).

4.2.5.1.2 Data distribution. Prior to analyses, the distribution of the data related to domain scores was scrutinised to consider the symmetry in its distribution. Broadly, there are two ways in which to determine symmetry in the distribution. One is a ‘rule of thumb’ method, in which the absolute value of the skewness can be interpreted. Bulmer (1979) suggested that values from 0 to 0.5 are relatively symmetrical, while 0.5 to 1 is moderately skewed and 1 or more is highly skewed. An alternative idea within this ‘rule of thumb’ has been proposed, which is that values between -2 and 2 are sufficient to prove normal univariate distribution (George & Mallery, 2010). The other interpretation method is a statistical calculation to determine whether the skewness is significantly different from zero. For the current data, to determine whether the skew was significant at the $p < 0.05$ level, the z score was calculated as Statistic/Standard Error and if the

absolute value was greater than 1.96, it was determined to be significantly skewed (Field, 2009).

Through these analyses, it was determined that the data were not normally distributed. The ranges for the skewness and kurtosis statistics, and their associated z scores (SE = 0.13 and 0.25 respectively), are presented in Table 20. The original and transformed data in relation to both participant self-report and custodial officer observations are outlined. Two domains consistently reflected the largest skew and kurtosis indices: Mental Disorder Negative and Manageability Negative. Although the data transformations increased the normality of these distributions, they remained non-normal. Further details for the skewness and kurtosis are provided below.

Table 20. *Data Distribution for Psychological Domains*

	Skewness statistic range		Skewness z score range		Kurtosis statistic range		Kurtosis z score range	
	Min	Max	Min	Max	Min	Max	Min	Max
Participant O	-0.22	3.77	-1.70	29.70	-1.29	18.47	-5.10	73.00
Participant R	-1.74	1.27	-13.68	10.01	-0.96	2.63	-3.79	10.41
Custodial O	0.59	7.03	4.61	55.35	0.15	67.40	0.59	266.41
Custodial R	-3.85	0.69	-30.33	5.40	-1.29	17.31	-5.08	68.41
P-C O	-0.14	2.35	-1.11	18.46	-0.82	14.41	-3.26	56.94
P-C NL	-2.85	0.43	-22.40	3.38	-0.61	34.67	-2.40	137.02

Note. O = Original; R = Reciprocal; NL = Natural Log; P-C = Difference between participant and custodial officer values.

The following domains were positively skewed for participants: Psychological Adjustment Negative; Mental Disorder Negative; Social Adjustment Negative; and, Manageability Negative. The following domains were positively skewed for custodial officer observations: Psychological Adjustment Negative and Positive; Mental Disorder Negative; Social Adjustment Negative; and, Manageability Negative and Positive. The positive skew indicated that for these domains, the majority of the data was at the lower frequency of behavioural observations.

Although not every domain was skewed, each domain was transformed in order to allow for comparison within the analyses. This transformation served to reduce the extent of the skew (Manikandan, 2010). Using both the natural log and reciprocal methods of transforming the data, the domains' skewness generally decreased, with the reciprocal method more successfully decreasing the skew. Despite the decrease in the

skewness values, the data remained skewed except for custodial officer observations for Psychological Adjustment Positive and Manageability Positive.

All but one of the HLM analyses was conducted using the reciprocal data. This method transformed the data such that the new values represented the relative time between each occasion of behaviour engagement, rather than the values representing the frequency of behaviour. This was determined to be the most appropriate transformation method as it was more effective in normalising the distribution than was the natural log transformation.

HLM was also used to analyse the differential change over time between participant self-report and custodial officer observations. As previously described, calculations were conducted to determine the quartiles in which each observation point was contained. For each participant, the quartile value at each time point was averaged across the risk factors in each of the eight domains; therefore, there were eight values at each time point throughout the participant's engagement in the research (either for the duration of their treatment participation or until their point of withdrawal from the research). The same calculations were conducted for the custodial officer observations. Subsequently, the custodial officer observations were subtracted from the participant self-report values to gain a difference score at each time point. The custodial officer observations were generally lower frequencies than the participant self-report values, thereby resulting in positive values rather than negative values.

Skewness was analysed separately. Data in the following domains were skewed: Psychological Adjustment Positive; Mental Disorder Positive; Social Adjustment Negative; and, Manageability Negative. The data were transformed using the natural log, rather than the reciprocal method, based on the following rationale. As this dataset consisted of the differences between quartiles, as opposed to the quartiles themselves, there were values between zero and one. Using the reciprocal method, low values were transformed into large values, thereby creating significantly skewed data, whereas the natural log process did not create this challenge. The natural log transformation cannot be conducted on negative values. Therefore, a constant of three was added to each value to ensure all values were greater than zero (McDonald, 2014). Note that in order to interpret the coefficients subsequent to analysis, reverse transformations were conducted such that the constant was subtracted and the log transformation was reversed.

Although data remained skewed following the transformations, the majority of domains had skewness values between -1 and +1 (see Bulmer, 1979). Prior to the transformation, Mental Disorder Negative did not meet significance for skewness; however, subsequent to the transformation, it had a skewness value greater than one. For consistency and greater ease of interpretation, the transformed data for this domain was still used in the analysis.

4.2.5.1.3 Group-level regression analysis. Two datasets were developed for the group-level analyses. The first dataset contained the weekly domain scores for each participant's self-report data and the custodial officer observations. Ten custodial officers completed the checklists, with the identity of the custodial officer unknown for some checklists. The largest proportion of checklists completed by one custodial officer was 62.89%. The other custodial officers completed the following proportion of checklists each: 13.16%; 10.24%; 4.02%; 2.93%; 2.56%; 1.46%; 0.37%; and, two officers each completed 0.18%. The unidentified checklist observers completed 2.01%. In order to determine whether there was a difference between individual custodial officers' observations, dummy variables were assigned to each custodial officer; another variable was created for the unidentified observers' checklists. These dummy variables were also included in the first dataset.

The second dataset contained the following descriptive information for each participant: age at treatment commencement; STATIC-99R score; sentence length in months (the sentence length for the participant who was serving a life sentence was calculated based on his age at sentencing until the average male life expectancy); victim type (adult or child); prior offences (none or sexual); and, number of days from Earliest Release Date to treatment commencement. Within HLM, covariates must be either numerical or binary categorical values. This requirement limited the possible included detail within the descriptive information (e.g., further information about prior offences). These covariates were all included in exploratory HLM analyses to determine whether they had a significant impact on behavioural change over time. Through the exploratory analyses, estimated level-2 coefficients and their standard errors were obtained by regressing the empirical Bayes residuals on the level-2 predictors (i.e., covariates) selected for possible inclusion in subsequent HLM analyses. Using the t-value gained through this exploratory analysis, a determination was made regarding whether each covariate should be included in the model. Any covariates with a t-value greater than the absolute value of 1.96 were included in the subsequent model. These covariates

were not included in analyses related to the remainder of the change measures, which will be further considered in section 4.4.

The sample size was reduced by two for the analyses related to change over time within participant self-report, as two participants withdrew consent after completing only one behavioural checklist. All participants were included in the analyses related to custodial officer observations over time, as those participants who withdrew from the research consented to custodial officers' continued completion of behavioural checklists. In addition, a comparison was conducted between the individual custodial officers involved in data collection, to determine if there was a difference between their observations. In order to avoid multicollinearity, all but one of the dummy variables assigned to each observer was included in the model. The custodial officer who provided the most data was excluded from the model as the base category against which all others could be compared.

4.2.5.1.4 Individual-level regression analysis. A split file regression analysis was conducted for individual-level change. This process allowed for an analysis of the measurement of change within individual participants in each domain over their time in treatment. Analyses included change in behaviours through self-report and custodial officer observations.

4.2.5.2 Psychometric measures. The pre- and post-testing psychometric measures were analysed using SPSS 22.0. Little MCAR's test was used to analyse the missing data. Data were missing completely at random ($p = 1.00$), which meant the missing data was unlikely to bias estimates (Little, 1988). Therefore, it was appropriate to use multiple imputation to replace the missing values. Multiple imputation involves replacing missing data with a number of plausible estimates (Graham, Olchowski, & Gilreath, 2007). Analyses are subsequently conducted on each separate dataset. For many analyses, the datasets were pooled or averaged to create a single set of values. Excluding one participant who did not have post-treatment psychometric testing, the participant with the highest percentage of missing data had 24.9% missing data across the pre- and post-treatment psychometric measures. Graham et al. (2007) recommend that for datasets that have between 10% and 30% missing values, 20 imputations should be conducted.

The multiple imputation process therefore resulted in 20 separate datasets. Within the subsequent analyses conducted, some provided a set of data that represented "pooled" data from the 20 datasets. This pooled dataset is automatically calculated

within SPSS by taking into account the variation among the 20 imputations and combining the imputations to form one overall dataset. However, some analyses provided only the individual results for each of the 20 datasets, with no pooled data. Therefore, as an alternative to pooled data, another dataset was developed in which each value was averaged across the 20 imputations. This averaged dataset was used for the analyses that did not provide pooled results.

4.2.5.2.1 Data distribution. The distributions of the pre- and post-treatment data were analysed for normality prior to conducting the analyses. The data distribution for the original dataset, prior to multiple imputation being conducted, can be seen in Table 21. The majority of measures demonstrated a normal distribution both pre- and post-treatment. However, using the *z* score method for both skewness and kurtosis indices (*SE* = 0.54 and 1.54 respectively), data were not normally distributed in the CISS Task subscale pre-treatment, the SSEI post-treatment, and the WSFQ Sadomasochism subscale and SIS both pre- and post-treatment.

Table 21. *Data Distribution for Psychometric Measures using Original Dataset*

Measure	Skewness statistic		Skewness z score		Kurtosis statistic		Kurtosis z score	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Bumby Rape Scale	0.25	0.27	0.46	0.48	-1.20	-1.50	-1.16	-1.41
BMS	0.86	1.04	1.61	1.94	1.58	1.05	1.52	1.01
CISS								
Task	-1.54	-0.09	-2.88	-0.16	2.37	-1.02	2.28	-0.91
Emotion	-0.28	-0.87	-0.51	-1.51	-0.96	0.27	-0.90	0.24
Avoidance	0.65	0.36	1.18	0.61	0.48	0.79	0.45	0.70
Avoidance Dist	0.27	0.32	0.49	0.55	0.84	0.50	0.79	0.45
Avoidance SD	-0.11	-0.15	-0.21	-0.26	-0.59	-0.42	-0.56	-0.37
CUSI	0.10	0.59	0.19	1.07	-1.54	0.43	-1.48	0.40
WSFQ								
Intimate	0.58	0.04	1.09	0.08	0.02	-0.81	0.02	-0.78
Impersonal	0.89	0.38	1.66	0.72	2.89	-1.19	2.78	-1.15
Exploratory	0.84	-0.19	1.56	-0.36	0.05	-0.25	0.04	-0.24
Sadomasochism	0.01	3.01	0.01	5.61	5.09	11.16	4.90	10.75
LS-UCLA	-0.07	0.13	-0.14	-0.24	0.01	0.31	0.01	0.30
SIS	-1.65	-2.01	-3.07	-3.74	2.37	3.42	2.28	3.30
SSEI	-0.59	-1.24	-1.10	-2.32	0.61	2.92	0.59	2.81

Note. BMS = Bumby Molest Scale; CISS = Coping Inventory for Stressful Situations; CISS Avoidance Dist = Avoidance Distraction subscale; CISS Avoidance SD = Avoidance Social Diversion subscale; CUSI = Coping Using Sex Inventory; WSFQ = Wilson Sexual Fantasy Questionnaire; SIS = Social Intimacy Scale; SSEI = Social Self Esteem Inventory.

Tables 22 and 23 outline the skewness and kurtosis indices respectively, for the datasets developed through the multiple imputation process. The values in these tables represent the ranges for the 20 multiple imputation datasets. The following measures demonstrated data that were not normally distributed at either one of pre- or post-treatment: BMS; CISS Task subscale; WSFQ Impersonal and Sadomasochism subscales; and, SIS.

Table 22. *Skewness Index for Psychometric Measures using Multiple Imputation Datasets*

Measure	Pre-Treatment				Post-Treatment			
	Skewness range		Skewness z score range		Skewness range		Skewness z score range	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Bumby Rape Scale	0.34	0.73	0.63	1.35	0.12	0.27	0.23	0.51
BMS	0.71	0.98	1.32	1.82	1.20	1.37	2.23	2.55
CISS								
Task	-1.34	-1.18	-2.51	-2.20	-0.38	0.12	-0.71	0.22
Emotion	-0.69	-0.42	-1.29	-0.78	-1.04	-0.55	-1.94	-1.02
Avoidance	0.53	0.79	0.99	1.48	0.16	0.55	0.29	1.02
Avoidance Dist	0.16	0.40	0.29	0.75	0.10	0.55	0.18	1.02
Avoidance SD	-0.21	0.01	-0.40	0.02	-0.43	0.09	-0.80	0.17
CUSI	0.10	-1.54	0.19	0.19	0.51	0.79	0.96	1.47
WSFQ								
Intimate	0.37	0.51	0.68	0.95	-0.13	-0.02	-0.24	-0.04
Impersonal	1.38	1.73	2.58	3.23	0.16	0.41	0.29	0.77
Exploratory	0.52	1.00	0.97	1.87	0.14	0.27	0.25	0.50
Sadomasochism	1.98	2.71	3.69	5.06	3.09	3.46	5.76	6.46
LS-UCLA	0.00	0.20	0.00	0.38	0.04	0.40	0.07	0.22
SIS	-1.11	-0.66	-2.07	-1.23	-2.36	-2.19	-4.41	-4.09
SSEI	-0.03	0.08	-0.05	0.16	-0.02	0.11	-0.04	0.21

Note. BMS = Bumby Molest Scale; CISS = Coping Inventory for Stressful Situations; CISS Avoidance Dist = Avoidance Distraction subscale; CISS Avoidance SD = Avoidance Social Diversion subscale; CUSI = Coping Using Sex Inventory; WSFQ = Wilson Sexual Fantasy Questionnaire; SIS = Social Intimacy Scale; SSEI = Social Self Esteem Inventory.

Table 23. *Kurtosis Index for Psychometric Measures using Multiple Imputation Datasets*

Measure	Pre-Treatment				Post-Treatment			
	Kurtosis range		Kurtosis z score range		Kurtosis range		Kurtosis z score range	
	Min.	Max.	Min.	Max.	Min.	Max.	Min.	Max.
Bumby Rape Scale	-0.52	0.09	-0.50	0.08	-1.68	-1.43	-1.61	-1.37
BMS	-0.23	0.54	-0.22	0.52	0.89	1.48	0.86	1.43
CISS								
Task	1.77	2.25	1.70	2.16	-0.82	-0.53	-0.79	-0.51
Emotion	-0.61	0.02	-0.58	0.01	-0.06	0.76	-0.06	0.73
Avoidance	0.35	0.80	0.34	0.77	0.69	1.28	0.67	1.23
Avoidance Dist	0.72	1.29	0.70	1.24	0.68	1.19	0.66	1.14
Avoidance SD	-0.71	-0.43	-0.68	-0.41	-0.90	-0.22	-0.87	-0.22
CUSI	-1.54	-1.54	-1.48	-1.48	0.30	1.08	0.29	1.04
WSFQ								
Intimate	-0.10	0.22	-0.09	0.21	-1.07	-0.90	-1.03	-0.86
Impersonal	2.70	4.13	2.60	3.98	-1.39	-0.92	-1.34	-0.88
Exploratory	-0.70	0.92	-0.67	0.89	-1.06	-0.69	-1.02	-0.66
Sadomasochism	3.54	8.40	3.41	8.10	10.39	12.89	10.00	12.41
LS-UCLA	-0.39	-0.01	-0.37	-0.01	0.40	0.54	0.38	0.52
SIS	-0.08	1.61	-0.08	1.55	5.49	6.20	5.29	5.97
SSEI	-0.50	-0.41	-0.48	-0.40	-0.19	-0.13	-0.18	-0.12

Note. BMS = Bumby Molest Scale; CISS = Coping Inventory for Stressful Situations; CISS Avoidance Dist = Avoidance Distraction subscale; CISS Avoidance SD = Avoidance Social Diversion subscale; CUSI = Coping Using Sex Inventory; WSFQ = Wilson Sexual Fantasy Questionnaire; SIS = Social Intimacy Scale; SSEI = Social Self Esteem Inventory.

4.2.5.2.1 Group-level analyses. The assumptions underlying paired samples *t*-tests were analysed using the averaged dataset described above. The Shapiro-Wilk test demonstrated that normality was violated for the following measures: post-treatment BMS; pre-treatment CUSI; post-treatment SIS; pre-treatment WSFQ Impersonal subscale; and, both pre- and post-treatment WSFQ Sadomasochistic subscales. Therefore, to calculate group-level change on the psychometric measures, repeated-measures *t*-tests were used for the: BRS; LS-UCLA; SSEI; CISS; and, WSFQ Intimate and Exploratory subscales. The Wilcoxon signed rank analysis was used to calculate group-level change for the remaining measures: BMS; CUSI; SIS; and, WSFQ Impersonal and Sadomasochistic subscales. Due to the small sample size, the Wilcoxon signed rank analysis was also used to calculate group-level change for those psychometric measures for which repeated-measures *t*-tests had been used, to determine if similar results were found using this more conservative analysis.

Each psychometric measure could be classified within the four psychological domains on which the RSVP was developed (Table 24). This categorisation of psychometric measures into domains facilitated subsequent comparisons between results from the behavioural checklists and the pre- and post-treatment psychometric measures. To determine the psychological domain to which each psychometric measure should be categorised, the RSVP manual was consulted, in which the risk factors and their potential behavioural manifestations were described. The domain Psychological Adjustment includes characteristics such as: cognitive distortions regarding sexual violence; attitudes of sexual entitlement; problems with anger and impulsiveness; and, maladaptive coping. The psychometric measures that similarly target these characteristics are: BRS (assessing cognitive distortions related to sexual violence against women); BMS (assessing cognitive distortions related to sexual offending against children); CISS (assessing coping styles); and, CUSI (assessing the use of sex as a coping strategy). The domain Mental Disorder includes attitudes and behaviours consistent with sexual deviance, which is consistent with the WSFQ. The domain Social Adjustment includes characteristics associated with relationships, such as: loneliness; attachment problems; and, poor social skills. The psychometric measures that assess these characteristics are: LS-UCLA (assessing feelings of loneliness and social isolation); SIS (assessing intimacy and loneliness in close relationships); and, SSEI (measuring social rejection and loneliness). The domain Manageability relates to issues surrounding planning, compliance with supervision, motivation and attitudes towards authority. It was determined that none of the psychometric measures targeted these areas and as such, this domain was not included in the pre- to post-treatment psychometric measures' analyses. The MC-SDS was used purely as a measure of social desirability in order to determine whether any participants engaged in positive impression management to an extent that would deem their responses invalid. Therefore, it was not categorised into a psychological domain.

Table 24. *Classification of Psychometric Measures into Psychological Domains*

Psychological Adjustment	Mental Disorder	Social Adjustment	Manageability
Bumby Rape Scale	Wilson Sexual Fantasy Questionnaire	Loneliness Scale-UCLA	Nil
Bumby Molest Scale		Social Intimacy Scale	
Coping Inventory for Stressful Situations		Social Self Esteem Inventory	
Coping Using Sex Inventory			

4.2.5.2.2 Individual-level analyses. Subsequent to the group-level analyses, the data were analysed at the individual participant-level. The methods by which these analyses were conducted are described below.

4.2.5.2.2.1 *Calculating the reliable change index.* Reliable change guides the interpretation of change for an individual (Viljoen et al., 2017). It is a z-value, which represents the probability that for a given individual, change would be observed due to chance alone. To assess the level of change each participant achieved pre- to post-treatment based on the results of the psychometric measures, and to determine whether this change was reliable, the reliable change index (RCI) was calculated. The formula used was that suggested by Christensen and Mendoza (1986, p. 305):

$$RCI = X_2 - X_1 / S_{diff}$$

This formula provides a more stringent rule than that proposed by Jacobson and Truax (1991), which considers inherent measurement error. X_2 represents a participant's post-test score, X_1 represents the same participant's pre-test score, and S_{diff} represents the standard error of the difference between X_2 and X_1 :

$$S_{diff} = \sqrt{2(SE)^2}$$

The standard error of measurement (SE) is computed with the formula:

$$SE = SD\sqrt{(1 - r_{xx})}$$

SD refers to the pre-treatment SD for the offender group for each measure and r_{xx} refers to the test-retest reliability of that measure. Test-retest reliability data was not available for all psychometric measures (CUSI and WSFQ). Although some research has suggested internal consistency should not be used (e.g., McCrae, Kurtz, Yamagata, & Terracciano, 2011), other research has suggested that internal consistency can be

used as an alternative when test-retest reliability is not available (e.g., Tingey, Lambert, Burlingame, & Hansen, 1996). The CUSI and WSFQ test developers were contacted to ascertain the availability of test-retest reliability data; however, the authors reported that no such data was available at that time. Therefore, Cronbach's alpha was used to calculate the RCI for the CUSI and WSFQ.

Jacobson and Truax (1991) suggest that an RCI of 1.96 or above reflects real (reliable) change for a two-tailed test with a 95% confidence level; that is, an RCI of this value or greater would mean that obtaining this change score by chance would be less than 5%. Therefore, it could be concluded with 95% confidence that the individual had demonstrated reliable change. To delineate reliable change status categories, participants who achieved an RCI of 1.96 or above in the direction of improvement were classified as 'demonstrating reliable change'. Participants who did not reach this threshold were classified as 'not having demonstrated reliable change'.

4.2.5.2.2.2 *Calculating clinically significant change.* According to Jacobson, Follette, and Revenstorf (1986), clinically significant change (CSC) is related to an individual's return to normal functioning. Jacobson and Truax (1991) suggest that there are several potential definitions of CSC. They propose the least arbitrary definition as one in which, subsequent to intervention (i.e., post-treatment), the level of an individual's functioning on the variable of interest is closer to the mean of the functional population than to the mean of the dysfunctional population. This process requires a determination of the point the individual must cross post-treatment in order to be assessed as changed to a clinically significant degree.

The cut-off points for each psychometric measure were calculated using the formula provided by Jacobson and Truax (1991, p. 13):

$$c = s_0M_1 + s_1M_0 / s_0 + s_1$$

s_0 and M_0 represent the *SD* and mean for the normative population, respectively; s_1 and M_1 represent the *SD* and mean for the offender population, respectively.

According to Jacobson and Truax (1991), if normative data are not available for a given measure, the cut-off point can be estimated using the two SD solution; that is, outside two SD from the mean of the offender population. This estimate is not as accurate as it does not account for the functional population. The CUSI did not have normative data available. The CUSI's authors were contacted in relation to the availability of normative data for this test, and they indicated that there was no normative data available. Therefore, the two SD method was used to calculate CSC for the CUSI.

The values used in the CSC cut-off, and the cut-off scores used in the calculations are presented in Table 25.

Table 25. *CSC Calculations Including Means, Standard Deviations, and Cut-Off Scores Used for Each Measure*

Psychometric measure	Normative sample M and (SD)	Offender sample M and (SD)	Cut-off score
Bumby Rape Scale	63.83 (13.43) Nunes et al. (2016)	58.64 (15.78) Olver et al. (2014)	61.44
Bumby Molest Scale	48.86 (10.01) Nunes et al. (2016)	61.47 (17.77) Olver et al. (2014)	53.40
CISS			
Task	58.56 (9.95)	58.68 (7.62)	58.63
Emotion	39.21 (11.54)	50.85 (8.69)	45.85
Avoidance	38.10 (9.59)	46.69 (8.18)	42.74
Avoidance Distraction	17.53 (5.51)	23.25 (5.47)	20.40
Avoidance SD	13.31 (4.13) Endler & Parker (1990)	17.52 (2.94) Serran et al. (2007)	15.77
Coping Using Sex Inventory	Unavailable	31.5 (9.22) Cortoni & Marshall (2001)	13.06
WSFQ			
Intimate	31.70 (9.31)	28.10 (13.3)	30.22
Impersonal	11.70 (6.77)	13.00 (9.1)	12.26
Exploratory	14.30 (7.87)	13.60 (10.3)	14.00
Sadomasochism	4.88 (5.71) Plaud & Bigwood (1997)	4.80 (6.9) Baumgartner et al. (2002)	4.84
Loneliness Scale-UCLA	37.06 (10.91) Russell et al. (1980)	41.91 (10.46) Olver et al. (2014)	39.54
Social Intimacy Scale	134.90 (21.9) Miller & Lefcourt (1982)	134.41 (25.02) Olver et al. (2014)	134.67
Social Self Esteem Inventory	132.00 (21) Lawson et al. (1979)	110.34 (13.01) Marshall, Anderson, et al. (1997)	118.63

Note. CISS = Coping Inventory for Stressful Situations; CISS Avoidance SD = Avoidance Social Diversion; WSFQ = Wilson Sexual Fantasy Questionnaire; Authors reported are the sources for the normative and offender samples.

Participants were categorised into one of five treatment-change categories as presented in Table 26, adapted from Wakeling et al. (2013). These categories facilitated data analysis related to clinically significant change.

Table 26. *Post-Treatment Change Categories*

Outcome Category	Description
Recovered	An individual who demonstrated reliable change and whose score moved from the dysfunctional to the functional range pre- to post-treatment
Improved	An individual who demonstrated reliable change but remained in the dysfunctional range post-treatment
Already okay	An individual in the functional range pre- and post-treatment, irrespective of reliable change
Unchanged	An individual who did not demonstrate reliable change pre- to post-treatment
Deteriorated	An individual who demonstrated reliable change such that the post-treatment score was significantly worse than the pre-treatment score and remained in the dysfunctional range

Note. These categories were described in Wakeling et al. (2013) based on Jacobson, Roberts, Berns, and McGlinchey (1999); the category names used in the current research are the same.

4.2.5.3 Clinically significant change and Treatment Gain scale. Correlational analyses were used to determine the relation between participants who achieved various levels of CSC and their Treatment Gain scale scores. Due to the distribution of the data, Spearman's rank-order non-parametric analysis was used.

4.2.5.4 Overall change. The results from the four change measures were subsequently interpreted in combination, for each participant. These results included various methods by which to monitor change over time: participant self-report; custodial officer observations; and, therapist ratings.

4.3 Results

The results for each measure of change are described below in relation to the three aims: (1) within the behavioural checklists, determining the overall differences between participant self-report and custodial officer observations, and between the presence and absence of different risk factors; (2) determining change in the behavioural checklists and the psychometric measures at both the group-and individual-level; and, (3) comparison between the four measures.

4.3.1 Behavioural Checklists

Results from the behavioural checklist self-report and observer data are presented below. Analyses served to answer questions regarding the manifestations of risk factors and more specifically, the processes of change over time in the risk-related behaviours that participants engaged in as they progressed through treatment. These questions included overall differences between self-report and observer reports within each of the risk factors, in addition to differences between each of the four domains addressed within the RSVP. The relative frequencies of each risk factor's behavioural manifestations are explored, with comparisons drawn between participant self-report and custodial officer observations of these behaviours on average throughout treatment (i.e., without taking into consideration the changes over time). The group-level changes in relation to the psychological domains are presented next, providing an overall perspective of the ways in which the group changed over the course of treatment. Participant self-report and custodial officer observations are each explored. After these group-level changes are presented, the data are further described at the individual-level, with analyses providing greater depth of information about the ways in which individuals changed or did not change in different domains.

4.3.1.1 Comparison between reports of behaviours. Overall, the binomial tests revealed that there was a significant association between participant and custodial officer reports of the combined behavioural manifestations for each risk factor ($p < 0.01$) and between the two behavioural manifestations of each risk factor ($p < 0.01$). That is, there was a greater likelihood of both participants and custodial officers similarly reporting the presence or otherwise of the behaviours, than of conflicting reports. Likewise, within the comparison between the two behavioural manifestations of each risk factor, there was a greater likelihood that the two behaviours were similarly reported as 'never' or 'ever' being present, than the presence of one behaviour but not the other being reported.

The contingency table results revealed no significant associations between any individual participant and custodial officer behavioural observations of the risk factors. Despite the failure to find any individual associations, some of the data from some of the contingency tables deserve additional discussion.

One example of a non-significant association is in relation to Sexual Deviance. In Study 1, relative to other risk factors, clinicians appeared to have more difficulty identifying its behavioural manifestations. A similar difficulty would be expected for

custodial officers' observations, compared with participant self-report. However, for the negative behavioural manifestations, 0/6 who self-reported ever engaging in the collection of inappropriate images or inappropriate sexual arousal were ever observed engaging in these behaviours, while 1/11 who never self-reported engaging in these behaviours were ever observed by custodial officers to do so ($p = 1.00$).

As seen in Table 27, there were several comparisons that could not be computed. Within these comparisons, at least one of the variables was a constant, such that every data point either had a code of 0 or 1. This situation occurred for one of the negative behaviour comparisons and all but three of the positive behaviour comparisons between participants and custodial officers. For seven of the 17 comparisons in which there was a constant, both the participant and custodial officer observations represented a constant; on eight occasions, it was only the participants' observations that were reported as a constant; and, on one occasion, only the custodial officers' observations represented a constant. For all these comparisons, the constant was the 'ever' observed variable; that is, the behaviour was reported at one time point at least.

Table 27. *Comparison between Participants' and Custodial Officers' Reports of Each Risk Factor*

Risk Factor	Negative Behaviour <i>p</i>	Positive Behaviour <i>p</i>
Extreme Minimisation or Denial	0.60	N/A
Attitudes that Support or Condone Sexual Violence	0.21	N/A
Problems with Self Awareness	0.12	N/A
Problems with Stress or Coping	N/A	N/A
Problems Resulting from Child Abuse	1.00	N/A
Sexual Deviance	1.00	0.47
Psychopathic Personality Disorder	1.00	N/A
Major Mental Illness	1.00	N/A
Problems with Substance Abuse	1.00	0.15
Violent or Suicidal Ideation	0.59	0.60
Problems with Intimate Relationships	1.00	N/A
Problems with Non-Intimate Relationships	1.00	N/A
Problems with Employment	0.34	N/A
Non-Sexual Criminality	0.52	N/A
Problems with Planning	1.00	N/A
Problems with Treatment	0.60	N/A
Problems with Supervision	0.10	N/A
Interpersonal Aggression	0.54	N/A
Offence Planning	1.00	N/A

Note. * = $p < 0.05$; ** = $p < 0.01$; N/A = comparison could not be computed as one of the variables was a constant.

The results in Table 27 combined the two negative behaviours for each risk factor manifestation with each other, and the two positive behaviours for each risk factor manifestation with each other. Hence, any differences between the reporting of these two behaviours were potentially concealed. Therefore, another set of analyses was conducted to determine whether there was an association between participants' and custodial officers' tendency to report one of the two behavioural manifestations in each pair (i.e., the two negative behavioural manifestations of each risk factor and the two positive behavioural manifestations of each risk factor). These analyses could indicate whether for each risk factor, one of the two behaviours was a more sensitive measure of the risk factor than was the other behaviour.

As seen in Table 28, similar to the previously reported analyses, all but four of the participant self-report positive behavioural manifestation comparisons could not be computed because at least one of the two variables was a constant. These behaviours

were constant because they were classified as having ‘ever’ been engaged in (i.e., they were all coded as 1 because there was at least one time point during which they were reported to have occurred) for all participants. There was one risk factor for which both the two negative behavioural manifestations were classified as constants due to custodial officers ‘never’ observing the behaviours in any participants. This risk factor was Sexual Deviance.

Within the negative behavioural manifestations, there were five significant associations through participant self-report and four significant associations through custodial officer observations. Within the positive behavioural manifestations, there were no significant associations through participant self-report as the majority of analyses could not be computed, while there were four significant associations through custodial officer observations.

The majority of the significant associations between the two behavioural manifestations were for either the risk factor’s negative behaviours or positive behaviours. However, there was a significant association between both the two negative behavioural manifestations and between the two positive behavioural manifestations, for two risk factors, as observed by custodial officers: Problems with Treatment and Interpersonal Aggression. An example of these significant associations is provided below for Interpersonal Aggression.

Within the negative behavioural manifestations, there was a strong association between intimidation of others and verbal aggression, with 5/5 of those who ever displayed intimidation of others also displaying verbal aggression, as compared with just 1/14 of those who never displayed intimidation displaying verbal aggression ($p < 0.01$). Within the positive behavioural manifestations, there was a strong association between appropriate interactions and assertive communication, with 12/13 of those who ever displayed appropriate interactions also displaying assertive communication, as compared with 0/6 of those who never displayed appropriate interactions displaying assertive communication ($p < 0.01$).

An example of a non-significant association is from custodial officer observations of the two negative behavioural manifestations of Problems with Stress or Coping. Within Study 1, behavioural manifestations of this risk factor appeared more easily identifiable than were manifestations of some of the other risk factors. It may be expected that custodial officers would observe the two behavioural manifestations with similar ease. Within this risk factor, 9/10 of those who ever displayed fluctuating

emotions also displayed withdrawal/isolation, as compared with 8/9 of those who never displayed fluctuating emotions ever displayed withdrawal/isolation ($p = 1.00$).

In order to adjust for multiple comparisons and the potential for Type I errors, the Bonferroni correction was applied, such that the 0.05 significance level was divided by the number of comparisons conducted, i.e., 76. This adjustment required p values to be less than 0.00066 in order to attain statistical significance. With this adjustment, those results that had demonstrated significance were no longer significant.

Table 28. *Comparison between Two Behaviours of Each Risk Factor*

Risk Factor	Negative Behaviour <i>p</i>		Positive Behaviour <i>p</i>	
	Participant	Custodial Officer	Participant	Custodial Officer
Extreme Minimisation or Denial	0.34	0.39	N/A	0.27
Attitudes that Support or Condone Sexual Violence	1.00	1.00	N/A	0.04*
Problems with Self Awareness	0.05*	0.07	N/A	N/A
Problems with Stress or Coping	1.00	1.00	N/A	0.11
Problems Resulting from Child Abuse	0.19	0.30	N/A	1.00
Sexual Deviance	0.35	N/A	0.06	<0.01**
Psychopathic Personality Disorder	0.23	0.02*	N/A	0.32
Major Mental Illness	1.00	0.37	0.18	N/A
Problems with Substance Abuse	N/A	0.05*	0.54	N/A
Violent or Suicidal Ideation	0.23	0.32	0.33	1.00
Problems with Intimate Relationships	0.25	1.00	N/A	0.16
Problems with Non-Intimate Relationships	0.01**	1.00	N/A	N/A
Problems with Employment	<0.01**	0.30	N/A	N/A
Non-Sexual Criminality	0.16	1.00	N/A	N/A
Problems with Planning	0.05*	1.00	N/A	1.00
Problems with Treatment	0.64	0.05*	N/A	0.05*
Problems with Supervision	0.05*	0.57	N/A	N/A
Interpersonal Aggression	0.08	<0.01**	N/A	<0.01**
Offence Planning	0.06	1.00	N/A	1.00

Note. * = $p < 0.05$; ** = $p < 0.01$; N/A = comparison could not be computed as one of the variables was a constant.

In summary, these results provided comparisons between behavioural observations of each risk factor for participants and custodial officers. Two primary groups of analyses were conducted, to determine whether there were associations between: participants' and custodial officers' observations; and, the two behavioural manifestations of each risk factor. The overall results revealed significant effects, such that across all the comparisons, there was a greater likelihood of both participants and custodial officers similarly reporting the behaviours as 'never' or 'ever' occurring, than of conflicting reports between participants and custodial officers. Further, there was a

greater likelihood of the two behavioural manifestations of each risk factor being reported as ‘never’ or ‘ever’ occurring, than one behaviour but not the other behaviour occurring.

The individual contingency table results for the first group of analyses revealed no significant associations. Within the second group of analyses, there were five significant associations within the negative behavioural manifestations; however, after adjusting for multiple comparisons, there were no significant associations. One finding was that many analyses could not be computed due to at least one variable representing a constant. The majority of these analyses involved positive behaviours, in which participants reported ‘ever’ engaging; that is, there were no positive behaviours within these analyses, for which participants reported no occurrences.

4.3.1.2 Group-level participant self-reported change. Change over time will first be described according to participant self-report. In order to determine whether missing data influenced the results of the self-report data at the group-level, three sets of analyses were conducted: (1) the first analysis included data from all participants, regardless of the presence of missing data; (2) the second analysis excluded two participants, both of whom withdrew from participation in the research after a week of data collection. The absence of these participants did not alter the results, with the same domains demonstrating significant change over time as with all participants included; and, (3) the third analysis excluded an additional two participants who only contributed to three and four weeks of data collection respectively. The results of these analyses reflected that while the p value decreased for those domains that did not demonstrate significant change over time, it did not alter whether the result was statistically significant or not. Therefore, the results including all participants are presented.

As displayed in Table 29, the results revealed (from participants’ self-report) a change over time in the Psychological Adjustment domain, such that positive behaviours increased and negative behaviours decreased. Within the remaining three domains, there was an increase in positive behaviours over time, with no significant change in negative behaviours.

Table 29. *Participant Self-Report of Group-Level Change in Each Domain throughout Treatment*

Domain	<i>b</i>	<i>p</i>
Psychological Adjustment Negative	< 0.01	0.01*
Mental Disorder Negative	< 0.01	0.67
Social Adjustment Negative	< 0.01	0.35
Manageability Negative	< 0.01	0.41
Psychological Adjustment Positive	< -0.01	< 0.01**
Mental Disorder Positive	< -0.01	< 0.01**
Social Adjustment Positive	< -0.01	< 0.01**
Manageability Positive	< -0.01	< 0.01**

Note. These coefficients are based on the reciprocal value of the quartile as described in the Method section; * = $p < 0.05$; ** = $p < 0.01$.

These same findings are also presented in graphical form through Figures 1 and 2, in order to provide greater clarity regarding the patterns in behavioural frequency changes over time. That is, the frequencies of positive and negative behaviours over time in treatment through participant self-report, are presented in Figures 1 and 2 respectively. The time frame used in these figures was the mean number of weeks participants remained in treatment ($M = 45$). The frequencies on the Y-axis refer to the quartiles used in the analyses, after the coefficients' reciprocal values were reversed. Despite not achieving statistical significance, the negative behaviours within Mental Disorder, Social Adjustment and Manageability appeared to show a trend towards a decrease over time (see Figure 1). As indicated, all domains achieved statistically significant increases in relation to positive behaviour; as seen in Figure 2, Manageability appeared to demonstrate the most prominent increase over time. None of the covariates described in section 4.2.5.1.3 (p. 168) had a significant impact on behaviours exhibited in treatment as reflected through participant self-report.

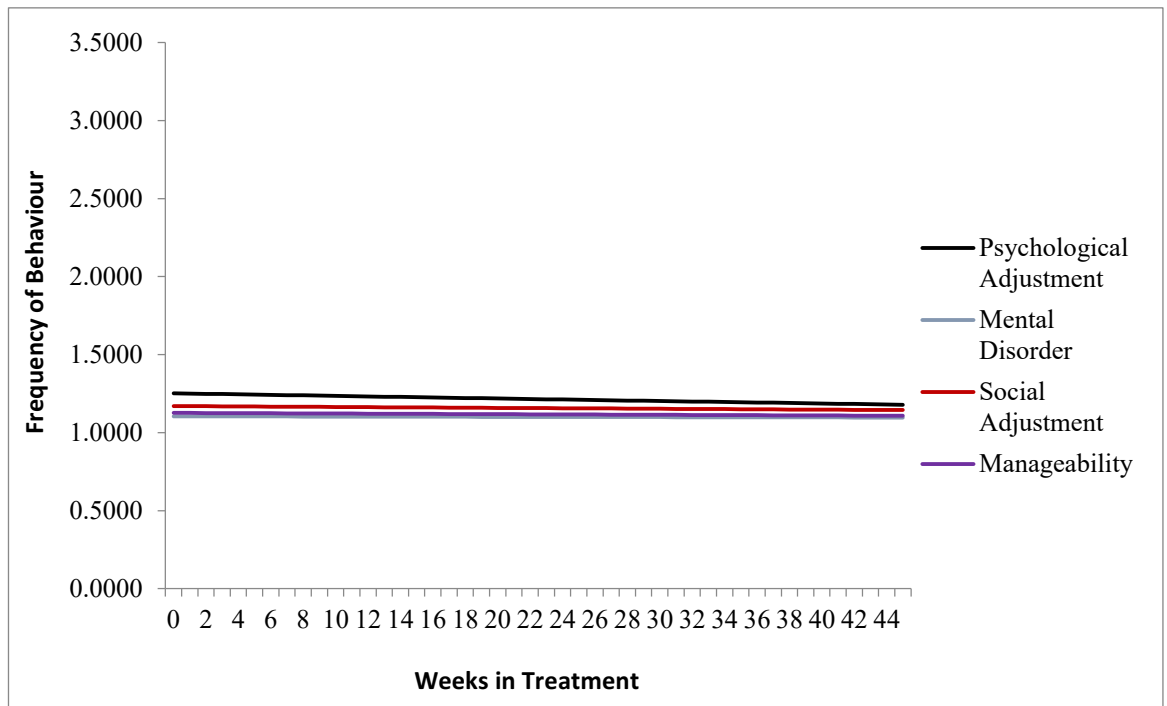


Figure 1. Participant self-report of group-level change in each negative domain throughout treatment.

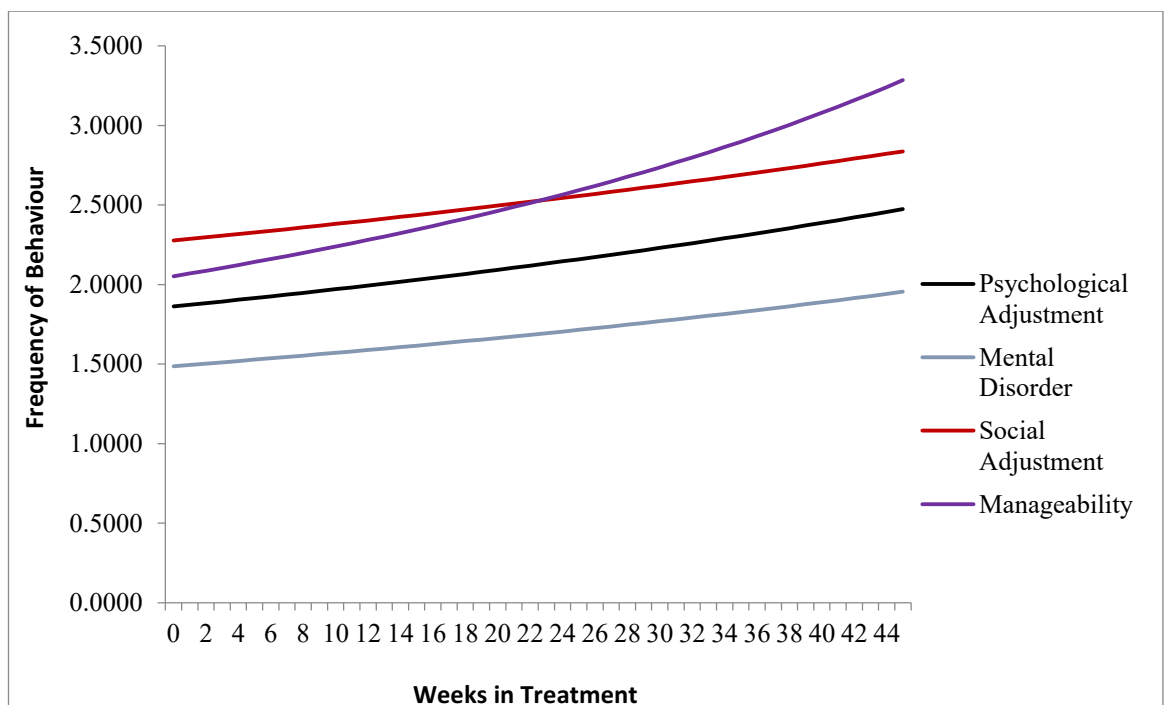


Figure 2. Participant self-report of group-level change in each positive domain throughout treatment.

4.3.1.3 Group-level custodial officer observed change. Due to the number of different observers, analyses were conducted to determine whether there was a difference between individual custodial officers' observations of participants' change over time. This comparison was conducted by using a base category for the custodial officer with the majority of observations, against which to compare each custodial officer's observations.

The results presented in Table 30 reflect the difference between each custodial officer's observations and the observations of the excluded custodial officer (i.e., the base category). There was a significant difference between custodial officers' observations of participants' behaviour over time. Specifically, of the 10 observers, compared with the custodial officer excluded as the base category, three reported significantly different observations in the Psychological Adjustment and Mental Disorder negative domains, one reported significantly different observations in the Social Adjustment negative domain, and six of the observers had significant differences in the Manageability negative domain.

Five observers displayed significant differences within all four positive domains. Two observers were significantly different in three of the four positive domains, while the remaining three observers had either one or two significant differences within the domains. Observers 6, 8, and 9, who demonstrated significant differences in the fewest domains, completed the fewest checklists (0.37%, 0.18% and 0.18% respectively).

Table 30. Comparison between Custodial Officer Observations throughout Treatment

Domain	Observer (% total checklists)									
	2 (4.02)	3 (13.16)	4 (1.46)	5 (10.24)	6 (0.37)	7 (2.56)	8 (0.18)	9 (0.18)	10 (2.93)	Unknown (2.01)
PAN <i>b</i>	-0.01	<0.01	-0.15	-0.13	0.05	-0.01	0.04	-0.43	<-0.01	-0.04
<i>p</i> value	0.83	0.94	<0.01**	<0.01**	0.58	0.76	0.80	<0.01**	0.98	0.42
MDN <i>b</i>	<-0.01	-0.02	-0.13	-0.06	-0.03	<0.01	-0.01	-0.02	-0.04	-0.07
<i>p</i> value	0.73	0.06	<0.01**	<0.01**	0.61	0.85	0.86	0.76	0.11	<0.01**
SAN <i>b</i>	-0.06	0.06	0.01	-0.01	0.07	0.09	0.08	-0.26	0.02	0.02
<i>p</i> value	0.06	<0.01**	0.78	0.58	0.46	0.03*	0.56	0.07	0.56	0.63
MNN <i>b</i>	-0.07	-0.03	-0.14	-0.05	-0.03	-0.01	<-0.01	<-0.01	-0.10	-0.08
<i>p</i> value	<0.01**	0.02*	<0.01**	<0.01**	0.68	0.83	0.99	0.96	<0.01**	<0.01**
PAP <i>b</i>	-0.26	-0.19	-0.38	-0.12	<-0.01	-0.29	0.10	-0.29	0.05	-0.32
<i>p</i> value	<0.01**	<0.01**	<0.01**	<0.01**	0.97	<0.01**	0.55	0.07	0.25	<0.01**
MDP <i>b</i>	-0.11	-0.16	-0.08	-0.09	0.13	-0.14	0.14	-0.02	0.11	-0.12
<i>p</i> value	<0.01**	<0.01**	0.03*	<0.01**	0.05*	<0.01**	0.14	0.88	<0.01**	<0.01**
SAP <i>b</i>	-0.20	-0.18	-0.23	-0.07	0.26	-0.03	0.27	0.08	0.27	-0.14
<i>p</i> value	<0.01**	<0.01**	<0.01**	<0.01**	<0.01**	0.44	0.03*	0.54	<0.01**	<0.01**
MNP <i>b</i>	-0.25	-0.22	-0.39	-0.08	0.08	-0.15	0.13	-0.25	0.15	-0.31
<i>p</i> value	<0.01**	<0.01**	<0.01**	<0.01**	0.42	<0.01**	0.36	0.07	<0.01**	<0.01**

Note. Observer = dummy variable assigned to custodial officer who completed behavioural checklists, with Observer 1 excluded as the base variable against which all others were compared; PAN = Psychological Adjustment Negative; MDN = Mental Disorder Negative; SAN = Social Adjustment Negative; MNN = Manageability Negative; PAP = Psychological Adjustment Positive; MDP = Mental Disorder Positive; SAP = Social Adjustment Positive; MNP = Manageability Positive; These coefficients are based on the reciprocal value of the quartile as described in the Method section; * = $p < 0.05$; ** = $p < 0.01$.

As presented in Table 31, at a univariate level, participants' positive behaviour was observed to increase over time within all domains. However, there was no observed change in negative behaviours within Psychological Adjustment, Social Adjustment or Manageability. There was an observed increase in negative behaviours in Mental Disorder.

Consistent with the way in which participant self-report findings were presented (i.e., in both tabular and graphical forms), the custodial officers' observations are also presented in graphical form through Figures 3 and 4, in order to provide greater clarity. That is, the frequencies of custodial officers' observations of participants' positive and negative behaviours over time in treatment are presented in Figures 3 and 4 respectively. As with Figures 1 and 2, the time frame used in these figures represented the mean number of weeks participants remained in treatment ($M = 45$). Similarly, the frequencies on the Y-axis refer to the quartiles used in the analyses, after the coefficients' reciprocal values were reversed. Although not gaining significance, the negative behaviours within Psychological Adjustment and Social Adjustment tended towards decreasing frequencies, while Manageability demonstrated an upward trend such that there was a slight increase in negative behaviour over time (see Figure 3). Positive behaviours within Social Adjustment appeared to have a higher baseline overall as compared with the other three domains (see Figure 4).

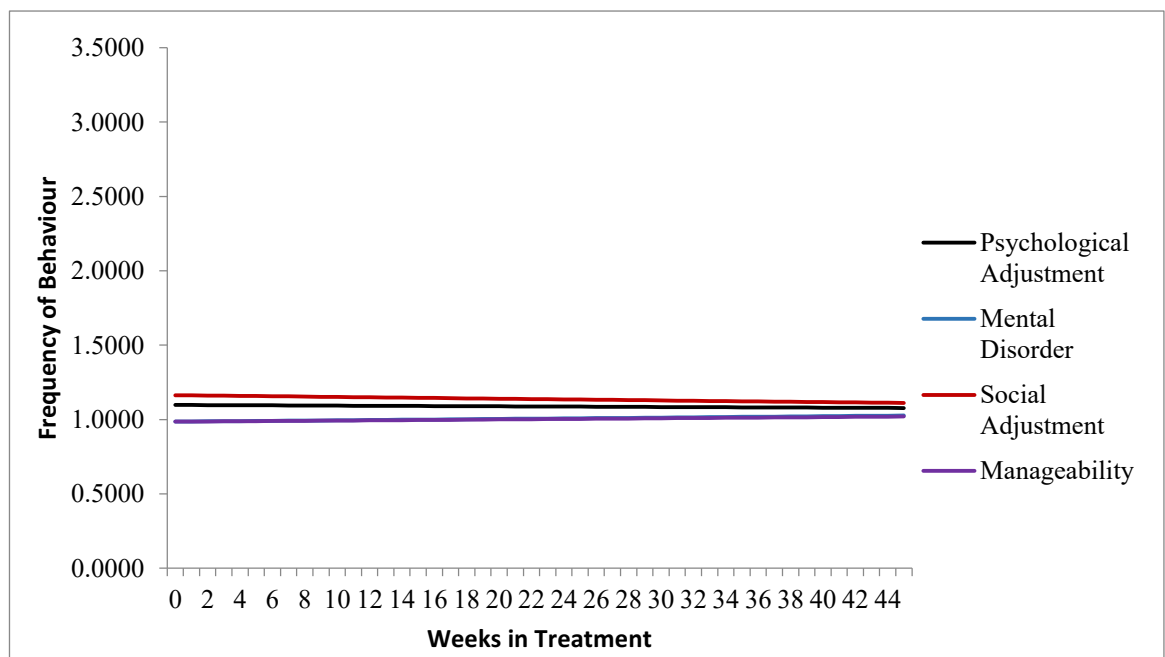


Figure 3. Custodial officer observations of group-level change in each negative domain throughout treatment.

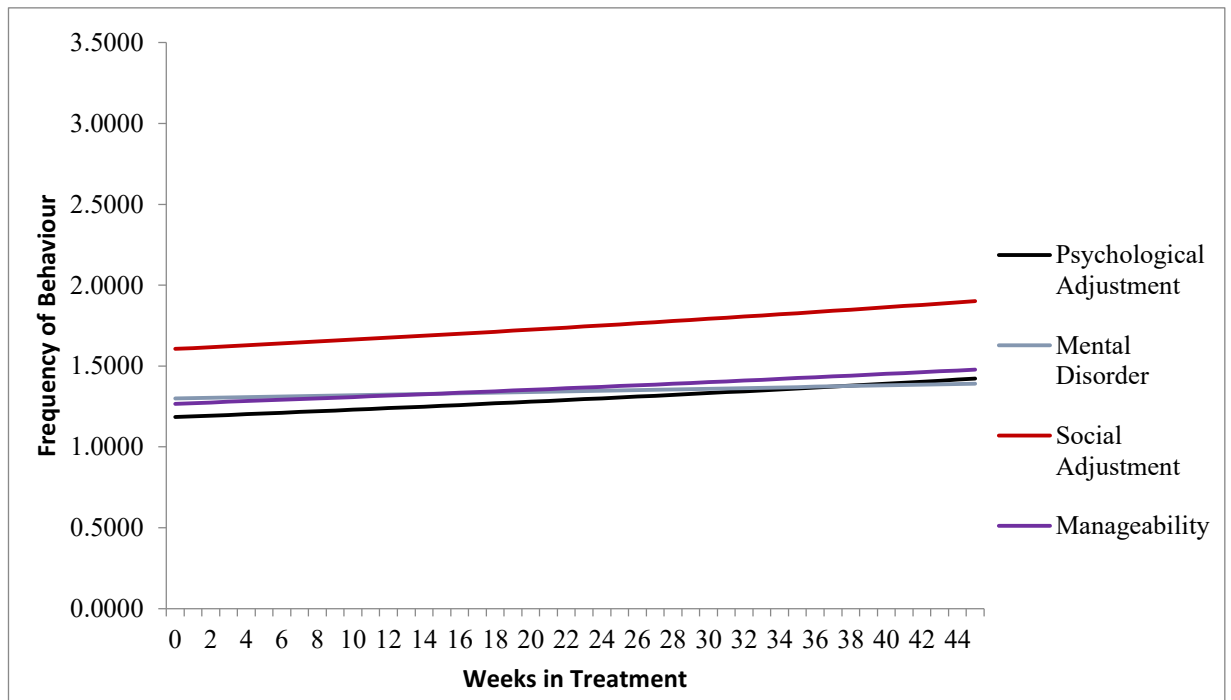


Figure 4. Custodial officer observations of group-level change in each positive domain throughout treatment.

Some participant information variables were included in the regression models as covariates, to determine whether they had an impact on changes in behavioural frequencies over time in treatment. As will be described, within some domains, some of these covariates affected behavioural change as observed by custodial officers (see Table 31). Specifically, participants' victim type had an impact on behavioural change in treatment within the domains Mental Disorder and Manageability. On average, participants who offended against adult victims had a greater frequency of negative behaviours in this domain as compared with those who offended against children. The impact of victim type was also time-dependent. For participants with adult victims, there was a decrease in negative behaviours over time in treatment. On the other hand, for participants who offended against children, there was an increase in negative behaviours as they progressed through treatment. Within positive behaviours in both Social Adjustment and Manageability, the effect of the number of days between ERD and treatment commencement moderated the overall positive impact of treatment over time.

Table 31. Custodial Officer Observations of Group-Level Change throughout Treatment

Domain	<i>b</i>	Overall <i>p</i>	Added variable 1	Added variable 1 constant <i>p</i>	<i>b</i>	Added variable 1 by week <i>p</i>	Added variable 2	Added variable 2 constant <i>p</i>	<i>b</i>	Added variable 2 by week <i>p</i>
PAN	< -0.01	0.46	Age	0.33	< -0.01	0.37				
MDN	< -0.01	0.05*	Static	0.06	< -0.01	0.33	Victim	< 0.01**	-0.09	< 0.01
SAN	< -0.01	0.08								
MNN	< -0.01	0.30	Static	0.16	< -0.01	0.82	Victim	0.01*	-0.05	< 0.01
PAP	< -0.01	< 0.01**	ERD to start	0.13	< -0.01	0.77				0.04*
MDP	< -0.01	0.01*	ERD to start	0.12	< -0.01	0.39				
SAP	< -0.01	< 0.01**	ERD to start	0.13	< -0.01	0.01**				
MNP	< -0.01	< 0.01**	ERD to start	0.37	< -0.01	0.01**				

Note. PAN = Psychological Adjustment Negative; MDN = Mental Disorder Negative; SAN = Social Adjustment Negative; MNN = Manageability Negative; PAP = Psychological Adjustment Positive; MDP = Mental Disorder Positive; SAP = Social Adjustment Positive; MNP = Manageability Positive; These coefficients are based on the reciprocal value of the quartile as described in the Method section; Added variable = covariate included in the model based on exploratory analysis; Age = age at time of treatment commencement; Static = STATIC-99R score; ERD to start = number of days between Earliest Release Date and treatment commencement; Victim = victim of current offence child or adult; *b* = coefficient, based on the reciprocal value of the quartile as described in the Method section; * = $p < 0.05$; ** = $p < 0.01$.

4.3.1.4 Participant self-report versus custodial officer observations. Overall, participant self-report and custodial officer observations of the frequency of behaviours within treatment did not statistically differ from each other in the majority of the domains (see Table 32). The only area with a significant difference observed between participants' and custodial officers' reports was positive behaviour within the Psychological Adjustment domain, such that on average, participants reported engaging in more positive behaviours in this domain than custodial officers reported observing.

Regression analyses revealed that there were changes over time in the difference between participant self-report and custodial officer observations within the positive behaviours in three of the domains (see Table 32). Within Psychological Adjustment, Mental Disorder and Manageability, initially participant self-report and custodial officer observations were similar to each other. Over time, this difference increased significantly. Participants initially reported engaging in positive behaviours within these domains with greater frequency than custodial officers observed. As time in treatment progressed, the frequency with which participants self-reported engaging in positive behaviour increased significantly more than did custodial officers' observations of these behaviours.

For Psychological Adjustment, the initial difference between participants and custodial officers was 0.65 of a quartile. Using the average length of treatment, 45 weeks, the difference between reported observations increased to 1.14 quartiles at the completion of treatment. For Mental Disorder, the initial difference between participants and custodial officers was 0.21 of a quartile. The difference between reported observations increased to 0.54 of a quartile at the completion of 45 weeks of treatment. For Manageability, the initial difference between participants and custodial officers was 0.87 of a quartile. The difference between reported observations increased to 1.55 quartiles at the completion of 45 weeks of treatment.

Table 32. *Comparison between Participant Self-Report and Custodial Officer Observations of Behaviour*

Domain	Coefficient (<i>b</i>)	Overall difference <i>p</i>	Coefficient (<i>b</i>)	Difference over time <i>p</i>
Psychological Adjustment Negative	0.06	0.31	< -0.01	0.26
Mental Disorder Negative	< -0.01	0.92	< -0.01	0.91
Social Adjustment Negative	0.09	0.10	< -0.01	0.45
Manageability Negative	0.02	0.49	< -0.01	0.69
Psychological Adjustment Positive	0.13	0.04*	< 0.01	< 0.01**
Mental Disorder Positive	0.02	0.64	< 0.01	< 0.01**
Social Adjustment Positive	0.11	0.12	< 0.01	0.11
Manageability Positive	0.10	0.08	< 0.01	< 0.01**

Note. These coefficients are based on the natural log transformation of the quartile as described in the method; * = $p < 0.05$; ** = $p < 0.01$.

4.3.1.5 Individual-level self-reported and observed change. To gain a more detailed understanding of change over time in treatment, each participant's behavioural change was analysed through self-report data and custodial officer observations. Due to the multiple comparisons conducted within these analyses, there is an increased likelihood for the occurrence of Type I error.

Participants more commonly reported behavioural changes than custodial officers observed. In addition, positive behaviours were reported to change more than were negative behaviours. Some individual participants changed across all domains while others made fewer changes. The results can be seen in Tables 33-36.

Within self-reported negative behaviour (see Table 33), one participant out of the 17 for whom analyses could be conducted reported an increase in behavioural frequency across all domains. No participants reported a decrease in all domains. Within self-reported positive behaviour (Table 34), three participants reported an increase in all domains and two participants reported a decrease in all domains. There were no participants who self-reported changes across all negative and positive domains. Seven participants self-reported no change within any negative domains, while five participants self-reported no change within any positive domains. Three of these participants were common across the negative and positive domains; that is, three

participants from the seven who reported no change in the negative domains similarly reported no change in the positive domains.

The domain with the largest number of participants self-reporting an increase in the frequency of negative behaviour was Manageability ($n = 3$); this domain also had the lowest number of participants self-reporting a decrease in the frequency of positive behaviour ($n = 2$). The remaining three domains had three participants each who self-reported a decrease in positive behaviour; the three participants whose negative behaviour increased in Manageability were not the same participants as those whose positive behaviour decreased in the other three domains. There was overlap within those participants whose positive behaviour decreased across the three domains; two of these participants were those who reported a decrease across all positive domains, and one participant reported a decrease in Psychological Adjustment and Social Adjustment.

Table 33. *Self-Reported Individual-Level Change in Each Negative Domain throughout Treatment*

Participant	Psychological Adjustment Negative		Mental Disorder Negative		Social Adjustment Negative		Manageability Negative	
	Change	Significance	Change	Significance	Change	Significance	Change	Significance
1	Dec [^]	< 0.01**	Dec [^]	0.01**	No	0.12	No	0.25
2	No	0.10	No	0.44	No	0.18	No	0.48
3	No	0.98	No	0.78	No	0.21	No	0.98
4	Dec [^]	< 0.01**	No	N/A	No	N/A	No	N/A
5	Inc	0.01**	Inc	0.01**	Inc	< 0.01**	Inc	0.02*
6	No	0.23	No	0.23	No	0.23	No	N/A
7	No	0.30	No	N/A	No	N/A	No	0.81
8	Dec [^]	0.01**	Dec [^]	0.02*	No	0.31	Dec [^]	< 0.01**
9	No	0.31	No	0.06	No	0.83	No	0.65
10	No	0.49	No	0.07	No	0.36	Inc	0.02*
11	Dec [^]	< 0.01**	No	0.24	No	0.74	No	0.46
12	No	0.27	No	0.40	No	0.58	Inc	< 0.01**
13	No	0.06	Inc	< 0.01**	No	0.18	No	0.10
14	No	0.22	No	0.24	No	0.33	No	0.54
15	No	0.43	No	0.78	No	0.25	No	0.58
16	Dec [^]	0.04*	No	0.09	Dec [^]	0.01*	No	0.07
17	Dec [^]	0.03*	No	0.15	No	0.13	Dec [^]	0.04*
18	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
19	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note. Dec[^] = significant decrease in frequency of negative behaviours over time; Inc[^] = significant increase in frequency of positive behaviours over time; Dec = significant decrease in frequency of positive behaviours over time; Inc = significant increase in frequency of negative behaviours over time; No = no significant change; N/A = error in analysis due to being constant or insufficient data; * = $p < 0.05$; ** = $p < 0.01$.

Table 34. *Self-Reported Individual-Level Change in Each Positive Domain throughout Treatment*

Participant	Psychological Adjustment Positive		Mental Disorder Positive		Social Adjustment Positive		Manageability Positive	
	Change	Significance	Change	Significance	Change	Significance	Change	Significance
1	No	0.30	No	0.81	No	0.10	No	0.37
2	No	0.86	No	0.57	No	0.25	No	0.34
3	Inc [^]	<0.01**	Inc [^]	<0.01**	Inc [^]	<0.01**	Inc [^]	<0.01**
4	Inc [^]	<0.01**	No	0.06	No	0.07	Inc [^]	<0.01**
5	No	0.16	No	0.17	No	0.48	No	0.23
6	No	0.92	No	0.93	No	0.65	No	0.59
7	No	0.42	No	0.52	No	0.95	No	0.44
8	Dec	<0.01**	Dec	<0.01**	Dec	0.01*	Dec	0.05*
9	Dec	0.03*	No	0.96	Dec	0.02*	No	0.51
10	Inc [^]	<0.01**	No	0.06	Inc [^]	0.02*	Inc [^]	<0.01**
11	No	0.85	No	0.07	Inc [^]	0.01**	Inc [^]	0.05*
12	Inc [^]	<0.01**	Inc [^]	<0.01**	Inc [^]	0.01**	Inc [^]	0.01*
13	No	0.35	Dec	0.01**	No	0.27	No	0.09
14	Inc [^]	<0.01**	Inc [^]	<0.01**	No	0.25	Inc [^]	<0.01**
15	Inc [^]	<0.01**	Inc [^]	<0.01**	Inc [^]	<0.01**	Inc [^]	<0.01**
16	Dec	<0.01**	Dec	<0.01**	Dec	<0.01**	Dec	<0.01**
17	Inc [^]	0.01*	No	0.73	No	0.29	No	0.09
18	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
19	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Note. Dec[^] = significant decrease in frequency of negative behaviours over time; Inc[^] = significant increase in frequency of positive behaviours over time; Dec = significant decrease in frequency of positive behaviours over time; Inc = significant increase in frequency of negative behaviours over time; No = no significant change; N/A = error in analysis due to being constant or insufficient data; * = $p < 0.05$; ** = $p < 0.01$.

Custodial officers did not report any of the 19 participants to increase or decrease their negative behaviours across all domains (Table 35). One participant was observed to increase his positive behaviours in all domains, with no participants observed to decrease these positive behaviours in all domains (Table 36).

According to custodial officer observers, 10 participants did not make any significant changes in negative behaviours across all domains over treatment. Similarly, 10 participants did not make significant changes in positive behaviour across domains over treatment; six of these participants were common across both the negative and positive domains, such that the same six participants did not make significant changes in any negative or positive domains.

According to custodial officers' observations, three participants increased their negative behaviour in Social Adjustment. Three participants decreased their positive behaviour in Manageability; one of these participants had also self-reported an increase in negative behaviour in this domain.

Table 35. Custodial Officer Observations of Individual-Level Change in Each Negative Domain throughout Treatment

Participant	Psychological Adjustment Negative		Mental Disorder Negative		Social Adjustment Negative		Manageability Negative	
	Change	Significance	Change	Significance	Change	Significance	Change	Significance
1	No	0.12	No	0.72	No	0.93	No	0.99
2	Dec [^]	< 0.01**	No	0.12	No	0.38	No	0.09
3	No	0.23	No	0.48	No	0.62	No	0.16
4	No	0.78	No	N/A	Inc	0.01**	No	0.98
5	No	0.33	No	0.95	No	0.30	No	0.09
6	No	0.33	No	0.32	Inc	0.01*	No	0.68
7	No	0.07	No	0.06	No	0.25	No	0.23
8	No	0.39	No	N/A	No	0.89	No	0.10
9	No	0.61	No	0.32	No	0.73	No	0.16
10	No	0.65	No	N/A	No	0.73	No	0.52
11	No	0.53	Inc	< 0.01**	No	0.78	No	0.06
12	No	0.28	Dec [^]	0.01**	No	0.54	No	0.30
13	Inc	< 0.01**	No	0.74	No	0.71	No	0.13
14	No	0.57	No	0.08	No	0.68	Inc	< 0.01**
15	No	0.79	No	0.54	No	0.25	No	N/A
16	No	0.55	No	0.42	No	0.17	No	0.89
17	No	0.08	No	0.22	Inc	0.03*	No	0.17
18	Dec [^]	< 0.01**	No	0.52	Dec [^]	< 0.01**	No	N/A
19	No	0.92	No	0.11	No	0.91	No	0.49

Note. Dec[^] = significant decrease in frequency of negative behaviours over time; Inc[^] = significant increase in frequency of positive behaviours over time; Dec = significant decrease in frequency of positive behaviours over time; Inc = significant increase in frequency of negative behaviours over time; No = no significant change; N/A = error in analysis due to being constant or insufficient data; * = $p < 0.05$; ** = $p < 0.01$.

Table 36. Custodial Officer Observations of Individual-Level Change in Each Positive Domain throughout Treatment

Participant	Psychological Adjustment		Mental Disorder Positive		Social Adjustment Positive		Manageability Positive	
	Change	Significance	Change	Significance	Change	Significance	Change	Significance
1	No	0.37	No	0.72	Inc [^]	0.02*	No	0.23
2	No	0.24	Dec	0.02*	No	0.05	Dec	0.04*
3	No	0.84	No	0.42	No	0.16	No	0.89
4	No	0.35	No	0.80	No	0.84	No	0.48
5	No	0.38	Inc [^]	0.04*	No	0.22	No	0.50
6	No	0.10	No	0.31	No	0.50	No	0.18
7	No	0.64	No	0.09	No	0.82	No	0.17
8	No	0.33	No	0.85	No	0.90	No	0.93
9	No	0.10	No	0.91	No	0.09	No	0.38
10	No	0.55	No	0.90	No	0.42	No	0.28
11	Dec	0.04*	No	0.80	No	0.40	Dec	0.05*
12	No	0.71	No	0.07	No	0.10	Dec	0.04
13	No	0.19	No	0.47	No	0.16	No	0.47
14	No	0.20	No	0.74	No	0.29	No	0.67
15	Inc [^]	0.01**	No	0.59	Inc [^]	0.03*	No	0.16
16	Inc [^]	< 0.01**	No	0.57	Inc [^]	< 0.01**	Inc [^]	0.01**
17	No	0.21	No	0.70	No	0.44	Inc [^]	0.04*
18	Inc [^]	< 0.01**	Inc [^]	< 0.01**	Inc [^]	< 0.01**	Inc [^]	< 0.01**
19	No	0.48	No	0.26	No	0.13	No	0.29

Note. Dec[^] = significant decrease in frequency of negative behaviours over time; Inc[^] = significant increase in frequency of positive behaviours over time; Dec = significant decrease in frequency of positive behaviours over time; Inc = significant increase in frequency of negative behaviours over time; No = no significant change; N/A = error in analysis due to being constant or insufficient data; * = $p < 0.05$; ** = $p < 0.01$.

4.3.2 Psychometric Measures

The pre- and post-treatment psychometric measures were analysed at both the group- and individual-level. Specifically, in reference to Hypothesis 3, these analyses were intended to assist with determining the relation between the various measures of change.

4.3.2.1 Descriptive statistics. Pre- and post-treatment descriptive statistics for the psychometric measures are presented in Table 37. The highest MC-SDS score was 27 (range = 1-27, $M = 13.83$, $SD = 7.16$); therefore, no participants were excluded based on a tendency for positive impression management.

Table 37. Descriptive Statistics for Pre- and Post-Treatment Psychometric Measures

Psychological domains and associated psychometric measure	Possible range	Cronbach's alpha		Pre-Treatment		Post-Treatment	
		Pre	Post	<i>M</i> (<i>SD</i>)	Min-Max	<i>M</i> (<i>SD</i>)	Min-Max
Psychological Adjustment							
Bumby Rape Scale	36-144	0.86	0.86	50.61 (8.87)	38.00-70.01	44.75 (6.91)	36.00-56.00
Bumby Molest Scale	38-152	0.96	0.94	52.78 (12.83)	38.00-82.00	47.44 (10.59)	38.00-75.00
CISS Task-Oriented Coping	16-90	0.96	0.90	58.93 (13.34)	22.00-75.00	63.10 (8.39)	48.00-76.00
CISS Emotion-Oriented Coping	16-90	0.75	0.80	45.64 (6.86)	30.00-55.00	42.10 (7.76)	25.00-53.00
CISS Avoidance-Oriented Coping	16-90	0.49	0.63	44.58 (11.49)	27.00-72.00	49.49 (11.39)	27.00-77.00
CISS Avoidance Distraction	8-40	0.82	0.85	20.97 (6.76)	8.00-37.00	21.27 (6.81)	10.00-38.00
CISS Avoidance Social Diversion	5-25	0.71	0.82	15.22 (4.33)	7.00-23.00	19.03 (3.41)	12.00-25.00
Coping Using Sex Inventory	16-80	0.79	0.71	24.56 (6.82)	16.00-34.00	22.59 (4.52)	16.00-34.00
Mental Disorder							
WSFQ Intimate	10-50	0.92	0.91	23.46 (8.90)	10.00-44.00	25.62 (8.15)	11.00-39.00
WSFQ Impersonal	10-50	0.76	0.66	14.71 (4.73)	10.00-29.00	15.11 (4.12)	10.00-22.00
WSFQ Exploratory	10-50	0.62	0.12	14.99 (4.72)	10.00-25.89	14.18 (2.62)	10.00-19.00
WSFQ Sadomasochism	10-50	0.67	0.88	11.38 (3.44)	10.00-24.00	11.38 (3.44)	10.00-24.00
Social Adjustment							
Loneliness Scale-UCLA	20-80	0.87	0.74	44.44 (11.08)	21.42-63.00	34.45 (5.87)	22.00-46.00
Social Intimacy Scale	17-170	0.87	0.94	131.21 (23.87)	70.75-160.73	130.58 (21.94)	60.00-149.00
Social Self Esteem Inventory	30-180	0.96	0.95	126.23 (27.89)	68.00-170.54	134.76 (20.96)	94.00-173.00
Marlowe Crowne-Social Desirability Scale	0-33	0.34	0.56	13.83 (7.16)	1.00-27.00		

Note. *N* = 18 as data unavailable for one participant; Minimum and Maximum scores for the current sample were calculated from the averaged multiple imputation data, thus creating fractional values; Psychological domains are from the RSVF and each psychometric measure was categorised within a domain, with no measures categorised into the domain Manageability (see Method); CISS = Coping Inventory for Stressful Situations; WSFQ = Wilson Sexual Fantasy Questionnaire.

4.3.2.2 Group-level program change. As presented in Tables 38 and 39, scores on the BRS, BMS, LS-UCLA, SSEI, and two CISS subscales (Avoidance and Avoidance Social Diversion) changed significantly over participants' time in treatment. Participants improved on each of these measures from pre- to post-treatment, with the exception of the two CISS subscales, in which participants' use of avoidance-oriented coping increased from pre- to post-treatment.

In relation to the psychological domains into which each psychometric measure was classified, there were some changes within Psychological Adjustment and Social Adjustment but these changes were not common across measures within the domains. There was no significant change in the domain Mental Disorder, as represented through the WSFQ.

Table 38. *Group-Level Change through Pre- to Post-Treatment Psychometric Measures*

Measure	MChange	SD	Std. Error Mean	<i>t</i>	<i>df</i>	Significance
Psychological Adjustment						
Bumby Rape Scale	5.85	6.54	1.58	3.71	17	< 0.01**
CISS Task-Oriented Coping	-4.17	9.80	2.36	-1.77	17	0.08
CISS Emotion-Oriented	3.54	8.97	2.20	1.61	17	0.11
Coping						
CISS Avoidance-Oriented	-4.90	8.52	2.10	-2.33	17	0.02*
Coping						
CISS Avoidance Distraction	-0.31	8.62	1.40	-0.22	17	0.83
CISS Avoidance Social	-3.81	3.72	0.94	-4.06	17	< 0.01**
Diversion						
Mental Disorder						
WSFQ Intimate	-2.16	8.38	1.99	-1.09	17	0.28
WSFQ Exploratory	0.81	4.25	1.03	0.79	17	0.43
Social Adjustment						
Loneliness Scale-UCLA	9.99	10.93	2.58	3.87	17	< 0.01**
Social Self Esteem Inventory	-8.53	17.89	4.25	-2.01	17	0.05*

Note. These measures were also tested using the Wilcoxon signed rank analysis due to the small sample size. The results remained significant for those measures that gained significance in the paired samples *t*-tests. CISS = Coping Inventory for Stressful Situations; WSFQ = Wilson Sexual Fantasy Questionnaire; * = $p < 0.05$; ** = $p < 0.01$.

Table 39. *Group-Level Change through Pre- to Post-Treatment Psychometric Measures*

Measure	Z	Significance
Psychological Adjustment		
Bumby Molest Scale	-1.97	0.05*
Coping Using Sex Inventory	-0.64	0.52
Mental Disorder		
Wilson Sexual Fantasy Questionnaire (Impersonal subscale)	-0.71	0.48
Wilson Sexual Fantasy Questionnaire (Sadomasochism subscale)	-0.15	0.88
Social Adjustment		
Social Intimacy Scale	-0.72	0.47

Note. * = $p < 0.05$.

4.3.2.3 Individual-level change. The group-level change analyses reflected that overall, participants improved on some measures through treatment, but there was also deterioration observed in others. The most common result amongst these measures was that there was no observed change over time in treatment. A more comprehensive representation of change can be gained through exploring change at the individual-level. A common method by which to explore individual-level change in psychometric measures is to calculate reliable and clinically significant change, as described in section 4.2.5.2.2 (pp. 175-176). Both reliable and clinically significant change results will be described in two forms: by participant for each psychometric measure; and, by psychometric measure.

Indices of reliable and clinically significant change are presented in Table 40, per participant. In other words, the change category each individual participant attained is provided for each psychometric measure. Within each participant's self-report, changes were made in some measures and not in others. This inconsistency was similarly apparent in the psychological domains, such that participants' change was not confined to certain domains. These results will be further outlined below, with some key findings in relation to both reliable and clinically significant change per participant provided. Notably, each participant's change profile was complex and idiosyncratic; this concept will be considered in section 4.4 and implications will be explored.

In relation to both reliable and clinically significant change, participants generally did not improve on the BRS, BMS and CISS subscales. In terms of clinically significant change, for the BRS and BMS, a common pattern was for participants' scores to be in the functional range pre-treatment, whereas for the CISS, participants remained in the dysfunctional range post-treatment. Three participants who improved on the CUSI also

improved on the LS-UCLA. There was a lack of consistency in levels of clinically significant change across the four WSFQ subscales within participants; that is, 15 participants demonstrated different levels of pre- to post-treatment change across the four subscales.

Including each CISS and WSFQ subscale as a distinct measure, the largest number of measures any participant made clinically significant improvements on, was five out of 15 ($n = 3$). One participant improved on four measures, three participants improved on two measures, and six participants improved on one measure. On the other hand, the largest number of measures on which any participant demonstrated deterioration was three out of 15 ($n = 2$). One participant deteriorated on two measures and five participants deteriorated on one measure.

Within these changes, the majority of participants ($n = 14$) demonstrated either only improvement or only deterioration on those measures for which clinically significant change was observed. Three participants demonstrated improvement on one measure and deterioration on another measure, while one participant demonstrated improvement on five measures and deterioration on one measure. One participant demonstrated no reliable or clinically significant change on any measures.

Table 40. *Reliable Change and Clinically Significant Change for Each Participant*

Participant	Psychological Adjustment										Mental Disorder					Social Adjustment		
	BRS	BMS	CISS	CISS	CISS	CISS	CISS	CISS	CISS	CISS	WSFQ	WSFQ	WSFQ	WSFQ	WSFQ	LS-	SIS	SSEI
			Task	Emotion	Avoid	AvoidD	AvoidS				Int	Imp	Expl	Sadom	UCLA			
1																		
RC	No	No	No	Imp	No	No	No	Imp	No	No	No	No	No	No	Imp	No		Imp
CSC	Okay	Okay	UC	Rec	UC	UC	UC	Imp	UC	UC	UC	UC	UC	UC	Rec	UC		Rec
2																		
RC	No	No	No	No	Det	No	Det	No	No	No	No	No	No	No	No	No	No	No
CSC	Okay	UC	Okay	UC	Det	UC	Det	UC	Okay	UC	Okay	UC	Okay	UC	Okay	UC	UC	Okay
3																		
RC	Imp	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Imp	No	No
CSC	Okay	Okay	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	Rec	Rec	UC
4																		
RC	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Imp	No		Imp
CSC	Okay	Okay	Okay	Okay	UC	Okay	UC	UC	Okay	Okay	Okay	Okay	Okay	UC	Rec	UC	UC	Imp
5																		
RC	Imp	Imp	No	No	No	UC	UC	Imp	Imp	No	Imp	No	No	No	Imp	No	No	No
CSC	Rec	Imp	UC	Okay	Okay	UC	UC	Imp	Rec	Rec	UC	UC	UC	UC	Rec	UC	UC	UC
6																		
RC	No	Imp	No	Imp	No	No	No	No	No	No	No	No	No	Imp	Imp	No	Imp	Imp
CSC	Okay	Rec	UC	Rec	Okay	Okay	UC	UC	Okay	UC	Imp	UC	UC	Imp	Rec	Okay	Rec	Rec
7																		
RC	No	No	No	No	No	No	No	No	Det	Det	Det	Det	No	Det	No	No	No	No
CSC	Okay	UC	Okay	UC	UC	UC	UC	UC	Det	Det	Det	Det	UC	Det	Okay	UC	UC	Okay
8																		
RC	No	No	No	No	No	No	No	Imp	No	No	No	No	No	No	Imp	No	No	No
CSC	Okay	UC	Okay	UC	UC	UC	Okay	Imp	Okay	UC	Okay	UC	UC	Imp	Rec	Rec	Okay	Okay
9																		
RC	No	No	No	No	No	No	No	Imp	No	No	No	No	No	No	No	No	No	No

	CSC	Okay	UC	UC	Okay	UC	Okay	UC	Imp	UC	UC	UC	UC	Okay	Okay	Okay
10	RC	No	No	Imp	Det	No	No	No	No	No	No	No	No	No	No	No
	CSC	Okay	Okay	Rec	Det	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC
11	RC	No	No	Imp	Imp	No	No	No	No	No	No	No	No	No	No	No
	CSC	Okay	Okay	Imp	Rec	Okay	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC
12	RC	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
	CSC	Okay	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC
14	RC	No	No	No	No	No	No	Det	No	No	No	No	No	Imp	Rec	Okay
	CSC	Okay	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC
15	RC	No	No	No	No	No	No	Imp	Imp	Imp	Imp	Imp	Imp	No	Det	No
	CSC	Okay	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC
16	RC	Imp	No	No	No	No	No	No	No	No	No	No	No	No	Det	No
	CSC	Rec	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC
17	RC	No	No	No	No	No	No	Imp	No	Imp	Imp	Rec	No	No	No	No
	CSC	Okay	Okay	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC
18	RC	No	No	No	No	No	No	Det	Det	Det	UC	UC	No	Det	No	No
	CSC	Okay	Okay	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC
19	RC	Imp	No	No	No	No	No	Det	No	Det	Det	Det	Det	No	Det	No
	CSC	Okay	Okay	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC
	RC	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
	CSC	Okay	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC	UC

Note. No post-testing was available for Participant 13, so this participant was excluded from the psychometric measure pre- and post-testing analyses. RC = Reliable change; CSC = Clinically significant change; Rec = Recovered; Imp = Improved; UC = Unchanged; Okay = Already

okay; Det = Deteriorated; BRS = Bumby Rape Scale; BMS = Bumby Molest Scale; CUSI = Coping Using Sex Inventory; LS-UCLA = Loneliness Scale; SIS = Social Intimacy Scale; SSEI = Social Self Esteem Inventory; CISS Task = Coping Inventory for Stressful Situations (Task-Oriented Coping subscale); CISS Emotion = Coping Inventory for Stressful Situations (Emotion-Oriented Coping subscale); CISS Avoid = Coping Inventory for Stressful Situations (Avoidance-Oriented Coping subscale); CISS AvoidD = Coping Inventory for Stressful Situations (Avoidance Distraction subscale); CISS AvoidS = Coping Inventory for Stressful Situations (Avoidance Social Diversion subscale); WSFQ Int = Wilson Sexual Fantasy Questionnaire (Intimate subscale); WSFQ Imp = Wilson Sexual Fantasy Questionnaire (Impersonal subscale); WSFQ Expl = Wilson Sexual Fantasy Questionnaire (Exploratory subscale); WSFQ Sadom = Wilson Sexual Fantasy Questionnaire (Sadomasochism subscale).

As indicated previously, reliable and clinically significant change indices were also calculated for each psychometric measure. These results are described below, in relation to the number of participants whose outcomes met the criteria for each change category.

4.3.2.3.1 Reliable change. RCI were calculated for each psychometric measure. As shown in Table 41, more than a quarter of the participants ($n = 5$) demonstrated reliable change in the CUSI and LS-UCLA. None of the participants improved on two of the CISS subscales, Avoidance and Avoidance Distraction; one participant deteriorated in the Avoidance subscale while all 18 participants remained unchanged in the Avoidance Distraction subscale. Deterioration was highest for the SIS and WSFQ Intimate ($n = 4$), while two participants deteriorated according to the WSFQ Sadomasochism and one participant deteriorated according to the CUSI, CISS Emotion, CISS Avoidance, and WSFQ Impersonal.

Table 41. *Number of Participants who Made Reliable Changes per Psychometric Measure*

Measure	Improved n (%)	No Change n (%)	Deteriorated n (%)
Psychological Adjustment			
Bumby Rape Scale	4 (22.22)	14 (77.78)	0 (0)
Bumby Molest Scale	2 (11.11)	16 (88.89)	0 (0)
CISS Task-Oriented Coping	2 (11.11)	16 (88.89)	0 (0)
CISS Emotion-Oriented	3 (16.67)	14 (77.78)	1 (5.56)
Coping			
CISS Avoidance	0 (0)	17 (94.44)	1 (5.56)
CISS Avoidance Distraction	0 (0)	18 (100)	0 (0)
CISS Avoidance SD	2 (11.11)	16 (88.89)	0 (0)
Coping Using Sex Inventory	5 (27.78)	12 (66.67)	1 (5.56)
Mental Disorder			
WSFQ Intimate	3 (16.67)	11 (61.11)	4 (22.22)
WSFQ Impersonal	1 (5.56)	16 (88.89)	1 (5.56)
WSFQ Exploratory	2 (11.11)	16 (88.89)	0 (0)
WSFQ Sadomasochism	2 (11.11)	14 (77.78)	2 (11.11)
Social Adjustment			
Loneliness Scale-UCLA	5 (27.78)	13 (72.22)	0 (0)
Social Intimacy Scale	2 (11.11)	12 (66.67)	4 (22.22)
Social Self Esteem Inventory	3 (16.67)	15 (83.33)	0 (0)

Note. CISS = Coping Inventory for Stressful Situations; Avoidance SD = Avoidance Social Diversion subscale; WSFQ = Wilson Sexual Fantasy Questionnaire.

4.3.2.3.2 Clinically significant change (CSC). Table 42 shows the proportion of participants in each CSC category for each psychometric measure. As outlined in the Method section, the following category labels are used: Improved = reliable change but remained in dysfunctional range; Recovered = moved from dysfunctional to functional range; Unchanged = no reliable change; Deteriorated = post-treatment score significantly worse than pre-treatment and remained in dysfunctional range; and, Already okay = in functional range pre- and post-treatment.

Between one and five participants recovered on nine of the 15 measures pre- to post-treatment. Additionally, between one and three participants deteriorated on eight of the 15 measures over the course of treatment. On average across the measures, approximately nine of the participants remained unchanged (i.e., remained in the dysfunctional range post-treatment) and approximately six were already at a functional level pre-treatment.

In relation to specific measures, almost all participants ($n = 16$) were already in the functional range pre-treatment on the BRS and, therefore, did not demonstrate clinically significant change. Eight participants did not demonstrate clinically significant change on the BMS because they were already in the functional range pre-treatment, with the same number remaining in the dysfunctional range pre- and post-treatment. No participants were in the functional range pre-treatment on the CUSI or WSFQ Sadomasochism subscale.

There was a combination of improvement and deterioration within the WSFQ subscales. However, the five improvements seen in Table 42 were from only two participants, with one participant demonstrating improvement in all four subscales. Two other participants demonstrated that they recovered on two of the subscales (Intimate and Exploratory). Three participants demonstrated that they deteriorated in some of the WSFQ subscales, with all three of these participants deteriorating in the WSFQ Intimate.

Three participants recovered in the CISS Emotion subscale, and one participant improved while another recovered in the CISS Task subscale; however, no participants improved in any of the CISS Avoidance subscales. More detailed information in relation to the RCI results for the CISS Avoidance Distraction subscale (i.e., none of the participants made changes in either direction through treatment) was gained through the CSC results. That is, the absence of change over time for any participants was due to 12

participants remaining in the dysfunctional range pre- to post-treatment and six participants remaining in the functional range pre- to post-treatment.

Five participants recovered in the LS-UCLA, while the majority were either functional pre-treatment ($n = 7$) or remained dysfunctional pre- to post-treatment ($n = 6$). No participants deteriorated on this measure from pre- to post-treatment.

Table 42. Number of Participants in Each Clinically Significant Change Category per Psychometric Measure

Measure	Dysfunctional pre-treatment			Functional pre-treatment		
	Improved <i>n</i> (%)	Recovered <i>n</i> (%)	Unchanged <i>n</i> (%)	Deteriorated <i>n</i> (%)	Already okay <i>n</i> (%)	
Psychological Adjustment						
Bumby Rape Scale	0 (0)	2 (11.11)	0 (0)	0 (0)	16 (88.89)	
Bumby Molest Scale	1 (5.56)	1 (5.56)	8 (44.44)	0 (0)	8 (44.44)	
CISS Task-Oriented Coping	1 (5.56)	1 (5.56)	6 (33.33)	0 (0)	10 (55.56)	
CISS Emotion-Oriented	0 (0)	3 (16.67)	9 (50.00)	1 (5.56)	5 (27.78)	
Coping						
CISS Avoidance-Oriented	0 (0)	0 (0)	14 (77.78)	1 (5.56)	3 (16.67)	
Coping						
CISS Avoidance Distraction	0 (0)	0 (0)	12 (66.67)	0 (0)	6 (33.33)	
CISS Avoidance SD	0 (0)	0 (0)	12 (66.67)	2 (11.11)	4 (22.22)	
Coping Using Sex Inventory	5 (27.78)	0 (0)	12 (66.67)	1 (5.56)	0 (0)	
Mental Disorder						
WSFQ Intimate	1 (5.56)	1 (5.56)	3 (16.67)	3 (16.67)	10 (55.56)	
WSFQ Impersonal	1 (5.56)	0 (0)	13 (72.22)	1 (5.56)	3 (16.67)	
WSFQ Exploratory	1 (5.56)	1 (5.56)	10 (55.56)	0 (0)	6 (33.33)	
WSFQ Sadomasochism	2 (11.11)	0 (0)	14 (77.78)	2 (11.11)	0 (0)	
Social Adjustment						
Loneliness Scale-UCLA	0 (0)	5 (27.78)	7 (38.89)	0 (0)	6 (33.33)	
Social Intimacy Scale	0 (0)	2 (11.11)	8 (44.44)	2 (11.11)	6 (33.33)	
Social Self Esteem Inventory	1 (5.56)	2 (11.11)	6 (33.33)	0 (0)	9 (50.00)	

Note. CISS = Coping Inventory for Stressful Situations; Avoidance SD = Avoidance Social Diversion subscale; WSFQ = Wilson Sexual Fantasy Questionnaire.

4.3.3 Treatment Gain: Short Scale and Relation to Psychometric Measures

Therapists' ratings for participants on the Treatment Gain scale ranged from 4 to 23 ($M = 16.7$, $SD = 5.13$).

Table 43 outlines for each participant, the number of psychometric measures that met the criteria for each CSC category, alongside the Treatment Gain scale score and the number of standard deviations from the mean, for each participant's Treatment Gain score. It identifies the rating provided by both a custodial officer and the participant's therapist on the SBRG, as either Satisfactory or Unsatisfactory. Correlational analyses were conducted to determine the relation between each CSC category and the Treatment Gain scores, which relates to Hypothesis 3. Participants 7 and 13 were excluded from these correlational analyses, as their data were incomplete. There was a significant positive correlation between the number of psychometric measures on which participants were rated as having recovered and their Treatment Gain scores ($r = 0.63$, $p = 0.01$). This correlation demonstrates that a higher Treatment Gain score was strongly associated with having recovered on more psychometric measures. There were no other significant correlations between CSC category and Treatment Gain scores.

Two participants achieved the highest Treatment Gain score in the sample (23 out of 24). These participants demonstrated that they recovered on three and four psychometric measures respectively. The participant who achieved the lowest Treatment Gain score (4 out of 24) improved on one measure, with the remainder of measures falling in the unchanged and already okay categories. Two participants deteriorated on three measures each; the Treatment Gain score was unavailable for one of these participants and the other scored 18 out of 24, which was 0.25 SD above the mean.

One participant, who scored 13 out of 24 on the Treatment Gain scale (0.73 SD below the mean), improved on five measures. This improvement reflected significant positive changes pre- to post-treatment, but he remained in the dysfunctional range for those measures post-treatment. He deteriorated on one measure, with the remainder in the unchanged (seven) and already okay (two) categories.

All participants gained a Satisfactory rating on the SBRG as rated by their therapist. A custodial officer rated one participant as Unsatisfactory. The inter-rater reliability could not be calculated for the SBRG ratings, because all the therapist ratings were the same.

Table 43. Number of Psychometric Measures in Each Clinically Significant Change Category and Treatment Gain Score per Participant

Participant	Improved (% of total number of measures)	Recovered (% of total number of measures)	Unchanged (% of total number of measures)	Already Okay (% of total number of measures)	Deteriorated (% of total number of measures)	Treatment Gain (out of 24)	SD from M	Custodial	SBRG Therapist
1	1 (6.67)	3 (20.00)	9 (60.00)	2 (13.33)	0 (0)	23	1.22	Satisfactory	Satisfactory
2	0 (0)	0 (0)	7 (46.67)	6 (40.00)	2 (13.33)	12	-0.92	Satisfactory	Satisfactory
3	0 (0)	1 (6.67)	11 (73.33)	3 (20.00)	0 (0)	22	1.03	Satisfactory	Satisfactory
4	1 (6.67)	1 (6.67)	5 (33.33)	8 (53.33)	0 (0)	17	0.05	Satisfactory	Satisfactory
5	2 (13.33)	3 (20.00)	8 (53.33)	2 (13.33)	0 (0)	19	0.44	Satisfactory	Satisfactory
6	1 (6.67)	4 (26.67)	5 (33.33)	5 (33.33)	0 (0)	23	1.22	Satisfactory	Satisfactory
7	0 (0)	0 (0)	8 (53.33)	4 (26.67)	3 (20.00)	N/A	N/A	N/A	N/A
8	1 (6.67)	1 (6.67)	7 (46.67)	6 (40.00)	0 (0)	22	1.03	Satisfactory	Satisfactory
9	1 (6.67)	0 (0)	8 (53.33)	6 (40.00)	0 (0)	4	-2.48	Satisfactory	Satisfactory
10	0 (0)	1 (6.67)	8 (53.33)	5 (33.33)	1 (6.67)	16	-0.14	Satisfactory	Satisfactory
11	1 (6.67)	1 (6.67)	6 (40.00)	7 (46.67)	0 (0)	20	0.64	Satisfactory	Satisfactory
12	0 (0)	1 (6.67)	10 (66.67)	3 (20.00)	1 (6.67)	17	0.05	Unsatisfactory	Satisfactory
13	N/A	N/A	N/A	N/A	N/A	20	0.64	Satisfactory	Satisfactory
14	5 (33.33)	0 (0)	7 (46.67)	2 (13.33)	1 (6.67)	13	-0.73	Satisfactory	Satisfactory
15	0 (0)	1 (6.67)	8 (53.33)	6 (40.00)	0 (0)	10	-1.31	Satisfactory	Satisfactory
16	0 (0)	1 (6.67)	8 (53.33)	5 (33.33)	1 (6.67)	19	0.44	Satisfactory	Satisfactory
17	0 (0)	0 (0)	3 (20.00)	11 (73.33)	1 (6.67)	15	-0.34	Satisfactory	Satisfactory
18	0 (0)	0 (0)	7 (46.67)	5 (33.33)	3 (20.00)	18	0.25	Satisfactory	Satisfactory
19	0 (0)	0 (0)	10 (66.67)	5 (33.33)	0 (0)	11	-1.12	Satisfactory	Satisfactory

Note. There were 18 participants included in the analyses; the excluded participant from the pre- to post-treatment psychometric measures was a different participant from the participant excluded from the Treatment Gain scale and SBRG. The values in each CSC category include the subscales from the CISS and WSFQ as separate measures.

4.3.4 Correspondence Between Change Measures

Each participant's change across the behavioural checklists, psychometric measures, Treatment Gain scale and SBRG, is described below. This description includes comment on the similarities and differences between the various measures. This process provided further insight into the final aim of the current research, which was to determine whether the frequency of dynamic risk factors and prosocial equivalent behaviours change over the course of treatment and whether changes in these behaviours corresponded with other markers of change, such as reliable and clinically significant change in the psychometric measures used in pre- and post-treatment testing, and the Treatment Gain scale and SBRG. The changes observed within the RSVP domains are provided in relation to the behavioural checklists and psychometric measures, with reports on both the negative and positive behaviours. Comparisons are also drawn between participant self-report and custodial officer observations.

4.3.4.1 Participant 1. In relation to change over time based on the self-report behavioural checklists, Participant 1 demonstrated a decrease in negative behaviours within Psychological Adjustment and Mental Disorder, but no change in the other negative domains or any of the positive domains. Custodial officer observations reflected different results, with an increase in positive behaviours within Social Adjustment.

This participant achieved reliable change in four psychometric measures: CUSI; LS-UCLA; SSEI; and, CISS (Emotion-Oriented Coping subscale). The parallel results for his clinically significant change reflect an improvement in the CUSI and recovered in the LS-UCLA, SSEI and CISS (Emotion-Oriented Coping subscale). He did not achieve reliable change in any other psychometric measures. In relation to the BRS and BMS, this lack of reliable change was consistent with his clinically significant change results, which reflected he was already in the functional range pre-treatment. He did not demonstrate deterioration in any of the psychometric measures, which is also consistent with his self-report and custodial officer observations within the behavioural checklists, such that there were no domains in which he demonstrated significant change in an unexpected direction (i.e., increase in negative behaviour or decrease in positive behaviour).

Comparing the psychometric measure results with the behavioural checklist results in relation to the psychological domains reveals similarities. Changes were observed through both these measures within the domains of Psychological Adjustment

and Social Adjustment. He scored 23 out of 24 on the Treatment Gain scale, which was 1.22 SD above the mean. Both his treating psychologist and a custodial officer provided an overall rating of Satisfactory on the SBRG. There was congruence between his positive changes within the psychometric measures and behavioural checklists, and his ratings on the Treatment Gain scale and SBRG.

4.3.4.2 Participant 2. Participant 2 did not self-report any significant change over time in treatment within any of the psychological domains through completion of the self-report behavioural checklists. Custodial officer observations reflected a decrease in negative behaviours within Psychological Adjustment. There was also a decrease in positive behaviours within Mental Disorder and Manageability. He did not achieve reliable change in a positive direction in any of the psychometric measures. He demonstrated deterioration over time in treatment within two subscales of the CISS (Avoidance-Oriented Coping and Avoidance Social Diversion). His deterioration in these two subscales was also clinically significant. This participant did not demonstrate clinically significant change over time within the BRS, LS-UCLA, CISS (Task-Oriented Coping subscale), WSFQ (Intimate subscale) or WSFQ (Exploratory subscale) as he was already in the functional range pre-treatment.

The results of the behavioural checklists and the psychometric measures can be compared: an apparent difference between these two measures of change was that his attitudes measured within the domain Psychological Adjustment appeared to deteriorate, whereas his behaviour within this domain improved according to custodial officer observations. This participant received a score of 12 out of 24 in the Treatment Gain scale, which was 0.92 SD below the mean. He was rated as Satisfactory on the SBRG by both his treating therapist and a custodial officer. There was some congruence between his psychometric measure results, behavioural checklists, and Treatment Gain scale results; however, these differed from the ratings he received on the SBRG.

4.3.4.3 Participant 3. In relation to the behavioural checklists, participant self-report reflected an increase in positive behaviours within each domain. However, there were no significant decreases in negative behaviours over time. Custodial officers did not report any behavioural changes over time in treatment. Participant 3 demonstrated reliable change in the BRS and SIS pre- to post-treatment. His reliable change in the SIS was also clinically significant, such that he recovered. On the other hand, his reliable change in the BRS did not reflect clinically significant change, as he was already in the functional range pre-treatment. He did not reliably change in any other

psychometric measures. He was already in the functional range pre-treatment within the BMS and CISS (Task-Oriented Coping subscale).

Comparing the behavioural checklist results with the psychometric measures reflected a difference between the changes associated with each psychological domain. The clinically significant change observed through the SIS is related to the Social Adjustment domain; however, no significant change was observed in the other psychometric measures, which suggests attitudes related to the other domains was similarly static over time. This is contrary to his self-reported behavioural change; however, it appears more similar to the custodial officer observations.

Participant 3 scored 22 out of 24 in the Treatment Gain scale, which was 1.03 SD above the mean. Both his therapist and a custodial officer rated him as Satisfactory overall while in treatment according to the SBRG. There was some congruence between his changes within the psychometric measures and behavioural checklists, and his Treatment Gain scale and SBRG ratings.

4.3.4.4 Participant 4. In relation to the behavioural checklists, within the Psychological Adjustment domain, there was a decrease in negative behaviours and a corresponding increase in positive behaviours according to participant self-report. There was no corresponding change in custodial observations of these behaviours. He also self-reported an increase in positive behaviours within Manageability. Custodial officers reported an increase in negative behaviours within Social Adjustment, which provides a contrast with the self-reported attitudes demonstrated through the LS-UCLA and SSEI, both of which were classified as measuring Social Adjustment. That is, Participant 4 demonstrated reliable change in the LS-UCLA and SSEI measures. This reliable change manifested as having recovered in the LS-UCLA (i.e., moving from the dysfunctional range pre-treatment to the functional range post-treatment), whereas he remained in the dysfunctional range post-treatment in the SSEI. He was already in the functional range in several of the measures: BRS; BMS; CISS (Task-Oriented Coping subscale, Emotion-Oriented coping subscale and Avoidance Distraction subscale); and, WSFQ (Intimate, Impersonal, and Exploratory subscales).

Participant 4 received a score of 17 out of 24 on the Treatment Gain scale, which was 0.05 SD above the mean. Both his therapist and a custodial officer rated him as Satisfactory overall on the SBRG. There was some congruence between his changes within the psychometric measures and behavioural checklists, and his Treatment Gain scale and SBRG ratings.

4.3.4.5 Participant 5. Within the behavioural checklists, based on participant self-report, there was an increase in negative behaviours within Mental Disorder, Social Adjustment and Manageability, with no significant changes over time in positive behaviours within any domain. These results are in contrast to the results obtained from the psychometric measures, which demonstrated improvement in the psychological domains of Psychological Adjustment and Mental Disorder. His improvement in Mental Disorder based on the psychometric measures was similar to custodial officers' observations of an increase in positive behaviours within this domain.

Specifically, Participant 5 made reliable change in the following psychometric measures: BRS; BMS; CUSI; LS-UCLA; and, WSFQ (Intimate subscale). These changes were paralleled through clinical significance for the BRS, LS-UCLA, and WSFQ (Intimate subscale). He was already within the functional range in the CISS subscales of Emotion-Oriented Coping and Avoidance-Oriented Coping.

He received a score of 19 out of 24 for the Treatment Gain scale, which was 0.44 SD above the mean. He received an overall rating of Satisfactory on the SBRG from both his treating therapist and a custodial officer. Overall, his self-reported behavioural checklist results were in contrast to the psychometric measure results and the ratings he received for the Treatment Gain scale and SBRG.

4.3.4.6 Participant 6. Within the behavioural checklists, there were no self-reported significant changes over time. This lack of reported change is consistent with the results obtained from some psychometric measures, in which he was already in the functional range pre-treatment. Observations from custodial officers reflected an increase in negative behaviours in Social Adjustment. These behavioural checklist results differ from the psychometric measures, in which there was no observed deterioration pre- to post-treatment. He remained in the dysfunctional range in the remaining psychometric measures. Specifically, in relation to the psychometric measures, Participant 6 demonstrated reliable change in the BMS, with this change similarly clinically significant. Likewise, these results were gained in the LS-UCLA, SSEI and CISS (Emotion-Oriented Coping subscale). He also demonstrated reliable change in the WSFQ (Sadomasochism subscale); however, post-treatment, he remained in the dysfunctional range. He was already in the functional range pre-treatment within the following measures: BRS; SIS; CISS (Avoidance-Oriented Coping and Avoidance Distraction subscales); and, WSFQ (Intimate subscale).

Participant 6 received a score of 23 out of 24 in the Treatment Gain scale, which was 1.22 SD above the mean. Both his treating therapist and a custodial officer provided an overall rating of Satisfactory on the SBRG. His limited change in the psychometric measures and behavioural checklists through self-report, and deterioration in custodial officer observations within the behavioural checklists, are in contrast to the results obtained from the Treatment Gain scale and SBRG.

4.3.4.7 Participant 7. There were no behavioural changes observed within the behavioural checklists either through self-report or custodial officer observations. There were only four weeks of data collection, which decreased the potential for identification of change over time. Consistent with his behavioural checklist results, Participant 7 did not achieve reliable or clinically significant change in the majority of psychometric measures, such that he remained in the dysfunctional range post-treatment in the majority of measures. On the other hand, he was already in the pre-treatment functional range for the following measures: BRS; LS-UCLA; SSEI; and, CISS (Task-Oriented Coping subscale). He deteriorated in three of the WSFQ subscales: Intimate, Impersonal, and Sadomasochism. These changes were clinically significant. The Treatment Gain scale and SBRG were not completed for this participant.

4.3.4.8 Participant 8. Within the behavioural checklists, participant self-report reflected changes in all domains. There was a decrease in both negative and positive behaviours within three of the four domains: Psychological Adjustment, Mental Disorder and Manageability; while there was also a decrease in positive behaviours within Social Adjustment but no change in negative behaviours. Custodial officers did not report changes over time in any of the domains. In relation to the psychometric measures, Participant 8 demonstrated reliable change in the CUSI and LS-UCLA measures. Although he remained in the dysfunctional range in the CUSI, he demonstrated clinically significant change in the LS-UCLA such that he was determined to have recovered pre- to post-treatment. This participant was already in the functional range pre-treatment in the following measures: BRS; SIS; SSEI; CISS (Task-Oriented Coping and Avoidance Social Diversion subscales); and, WSFQ (Intimate subscale). He remained in the dysfunctional range post-treatment in the following measures: BMS; CISS (Emotion-Oriented Coping, Avoidance-Oriented Coping, and Avoidance Distraction subscales); and, WSFQ (Intimate, Impersonal, and Exploratory subscales).

The results of the psychometric measures and the behavioural checklists are in opposition. The improvement observed in the LS-UCLA reflects an improvement in Social Adjustment, a domain in which this participant reported a decrease in positive behaviours through the behavioural checklists. Additionally, the overall limited change portrayed through the psychometric measures was in contrast to the self-reported behavioural checklist results, but similar to custodial officer observations.

Participant 8 scored 22 out of 24 in the Treatment Gain scale, which was 1.03 SD above the mean. Overall, he was rated as Satisfactory on the SBRG by both his treating therapist and a custodial officer. In general, there was limited congruence between the psychometric measures and behavioural checklists, and the Treatment Gain scale and SBRG results.

4.3.4.9 Participant 9. Within the behavioural checklists, there was a self-reported decrease in positive behaviours within Psychological Adjustment and Social Adjustment but no changes observed by custodial officers. Participant 9 demonstrated reliable change in one psychometric measure, the CUSI. Despite this change pre- to post-treatment, he remained in the dysfunctional range at the completion of treatment. Within the BRS, LS-UCLA, SIS, SSEI, CISS (Emotion-Oriented Coping and Avoidance Distraction subscales), his absence of clinically significant change was due to his being in the functional range pre-treatment. Post-treatment, he maintained his pre-treatment level of dysfunction with the other measures.

The limited changes reported through the psychometric measures are consistent with the minimal behavioural changes reported through the behavioural checklists. This participant scored 4 out of 24 in the Treatment Gain scale, which was 2.48 SD below the mean. On the contrary, his overall rating in the SBRG as reported by his treating therapist and a custodial officer was Satisfactory. Overall, his results in the psychometric measures and behavioural checklists were congruent with the Treatment Gain scale; however, these results were inconsistent with the SBRG ratings.

4.3.4.10 Participant 10. Within the behavioural checklists, there was a large amount of missing data in both self-report and custodial officer observations. According to participant self-report, there was an increase in positive behaviours within Psychological Adjustment, Social Adjustment and Manageability; there was also an increase in negative behaviours within Manageability. These changes were in contrast to the minimal self-reported changes through the psychometric measures. On the other hand, the absence of changes observed by custodial officers was more consistent with

the psychometric measures. Specifically, Participant 10 demonstrated both reliable and clinically significant change in a positive direction in the Task-Oriented Coping subscale of the CISS, but a negative change in the Emotion-Oriented Coping subscale of the CISS. In the remainder of the measures, he either remained in the functional range pre- to post-treatment (i.e., BRS, BMS, SIS, WSFQ Intimate and WSFQ Exploratory subscales) or remained in the dysfunctional range (i.e., CUSI, LS-UCLA, SSEI, the three CISS Avoidance-Oriented Coping subscales, and WSFQ Impersonal and Sadomasochism subscales).

He received a score of 16 out of 24 in the Treatment Gain scale, which was 0.14 SD below the mean. Both his treating therapist and a custodial officer reported an overall rating of Satisfactory on the SBRG. There was some congruence between his limited changes within the psychometric measures and custodial officer observations through the behavioural checklists, and his Treatment Gain scale score; however, these results were inconsistent with the SBRG ratings.

4.3.4.11 Participant 11. In relation to the behavioural checklists, there was a self-reported decrease in negative behaviour in Psychological Adjustment, which was in contrast to custodial officers' observations of a decrease in positive behaviour in this domain. There was a self-reported increase in positive behaviour in Social Adjustment. Both self-report and custodial officer observations reflected an increase in positive behaviours in Manageability, but no corresponding decrease in negative behaviours. In contrast, custodial officer observations reflected an increase in negative behaviour in Mental Disorder.

Within the psychometric measures, Participant 11 demonstrated reliable change in the CISS Task- and Emotion-Oriented Coping subscales. His change in the Emotion-Oriented Coping subscale was clinically significant such that he recovered pre- to post-treatment; however, his change in the Task-Oriented Coping subscale was not sufficient to improve beyond the dysfunctional range. He remained in the functional range throughout treatment in the following measures: BRS; BMS; the three Avoidance-Oriented Coping subscales of the CISS; and, the Intimate and Exploratory subscales of the WSFQ. On the other hand, he remained in the dysfunctional range throughout treatment in the following measures: CUSI; LS-UCLA; SIS; SSEI; and, the Impersonal and Sadomasochism subscales of the WSFQ.

The results of the psychometric measures differ from those obtained through the behavioural checklists, when the changes in domains are compared. In general, the

psychometric measures reflective of the Psychological Adjustment domain were either functional already or became functional, which was consistent with his self-reported decrease in negative behaviours measured through the behavioural checklists, but inconsistent with custodial officer observations. On the other hand, there was some consistency between self-report in the psychometric measures and behavioural checklist custodial officer observations in the Mental Disorder domain.

He achieved a score of 20 out of 24 in the Treatment Gain scale, which was 0.64 SD above the mean. He also achieved an overall rating of Satisfactory on the SBRG from both his treating therapist and a custodial officer. Overall, self-reported improvements through the behavioural checklists were the only results congruent with the Treatment Gain scale and SBRG ratings from staff; the minimal changes demonstrated through the psychometric measures and the custodial officer observations reported through the behavioural checklists were inconsistent with the Treatment Gain scale and SBRG ratings.

4.3.4.12 Participant 12. Within the behavioural checklists, there was a self-reported increase in positive behaviour over time in all psychological domains in addition to negative behaviour in Manageability. Custodial officers observed a decrease in negative behaviours in Mental Disorder. Participant 12 demonstrated reliable and clinically significant change in two psychometric measures. He improved in the SIS such that he had recovered pre- to post-treatment. Contrary to this result, his results in the WSFQ Intimate subscale suggested deterioration pre- to post-treatment. He was already in the functional range pre-treatment in three of the measures: BRS; SSEI; and, CISS (Avoidance Social Diversion subscale). He remained in the dysfunctional range post-treatment in the remainder of the measures.

His recovery according to the SIS demonstrates some agreement between the psychometric measurement of change and behavioural checklists. However, in general, there was limited consistency between the results obtained from the psychometric measures and behavioural checklists. For instance, there was deterioration in the WSFQ, which falls in the Mental Disorder domain, whereas custodial officers reported a decrease in negative behaviours in this domain.

His treating therapist reported a score of 17 out of 24 on the Treatment Gain scale, which was 0.05 SD above the mean. This score is consistent with his therapist's report of an overall rating of Satisfactory on the SBRG. However, the custodial officer overall rating was Unsatisfactory, which is in contrast to custodial officer observations reported

through the behavioural checklists. This participant was the only participant to gain an Unsatisfactory rating on the SBRG.

4.3.4.13 Participant 13. Within the behavioural checklists, participant self-report reflected an increase in negative behaviour and a decrease in positive behaviour within Mental Disorder. Contrary to these results, within Manageability, there was a self-reported decrease in negative behaviour and an increase in positive behaviour. Custodial officers reported an increase in negative behaviour within the domain Psychological Adjustment. Results from the psychometric measures for Participant 13 were not available.

In contrast with the reported increases in negative behaviours through the behavioural checklists, his treating therapist reported a score of 20 out of 24 on the Treatment Gain scale, which was 0.64 SD above the mean. He achieved an overall rating of Satisfactory on the SBRG from both his therapist and a custodial officer.

4.3.4.14 Participant 14. The behavioural checklist results reflected a self-reported increase in positive behaviour in all domains but Social Adjustment and no change in negative behaviour in any domain. This lack of change in Social Adjustment is consistent with his remaining in the dysfunctional range in psychometric measures in Social Adjustment. Custodial officers observed an increase in negative behaviour in Manageability. Participant 14 demonstrated reliable change in a positive direction in the CUSI and all WSFQ subscales. However, these changes were not clinically significant, such that he remained in the dysfunctional range post-treatment. He demonstrated a significant level of change in a negative direction in the SIS and he remained in the dysfunctional range pre- to post-treatment. He remained in the functional range throughout treatment in the BRS and CISS (Task-Oriented Coping subscale). He remained in the dysfunctional range throughout treatment in the BMS, LS-UCLA, SSEI, and CISS (Emotion-Oriented Coping and all three Avoidance-Oriented Coping subscales).

In contrast to his self-reported positive changes through the behavioural checklists, he scored 13 out of 24 on the Treatment Gain scale, which was 0.73 SD below the mean. Nevertheless, he gained an overall rating of Satisfactory on the SBRG from both his treating therapist and a custodial officer. Overall, there was limited congruence between the measures of change. The behavioural checklist results from custodial officer observations bore some similarity to the therapist-rated Treatment Gain score, with both of these results in opposition to the SBRG.

4.3.4.15 Participant 15. Within the behavioural checklists, there was a self-reported increase in positive behaviour in all domains. Similarly, custodial officers reported these changes in Psychological Adjustment and Social Adjustment. There were no reported changes in any negative behaviour. Participant 15 demonstrated clinically significant change in the BRS from pre- to post-treatment. He demonstrated reliable change in a negative direction in the SIS; however, he was in the functional range for this measure both pre- and post-treatment. He was similarly in the functional range pre- and post-treatment in the following measures: LS-UCLA; SIS; CISS (Task-Oriented Coping subscale); and, the Intimate, Impersonal, and Exploratory subscales of the WSFQ. On the other hand, Participant 15 remained in the dysfunctional range pre- to post-treatment in the following measures: BMS; CUSI; SSEI; CISS (Emotion-Oriented Coping and all three Avoidance-Oriented Coping subscales); and, the Sadomasochism subscale of the WSFQ.

This participant's reported increase in positive behaviours through the behavioural checklists was not consistent with the results obtained from the psychometric measures. Nevertheless, the stability in negative behaviour is consistent with the psychometric measures in which he remained in the dysfunctional range pre- to post-treatment. It is also consistent with his score of 10 out of 24 on the Treatment Gain scale, which was 1.31 SD below the mean. However, both his treating therapist and a custodial officer provided him with an overall rating of Satisfactory on the SBRG. Overall, there was limited congruence between the measures of change.

4.3.4.16 Participant 16. According to the behavioural checklist results, participant self-report reflected a decrease in positive behaviours within all domains. Contrary to these self-reported changes in positive behaviour, custodial officers reported an increase in positive behaviours in Psychological Adjustment, Social Adjustment and Manageability. There was also a self-reported decrease in negative behaviours within Psychological Adjustment and Social Adjustment.

Within the psychometric measures, Participant 16 demonstrated an improvement in the WSFQ Intimate subscale; however, he was already in the functional range pre-treatment. He similarly demonstrated an improvement in the Exploratory subscale, such that he improved from the dysfunctional to the functional range while in treatment. In the CISS (Avoidance Social Diversion subscale), he deteriorated such that he remained in the dysfunctional range throughout treatment and his post-treatment score was significantly worse than his pre-treatment score. He remained in the functional range

throughout treatment in the following measures: BRS; BMS; SSEI; CISS (Task-Oriented Coping subscale); and, WSFQ (Intimate subscale). He remained in the dysfunctional range throughout treatment in the rest of the measures.

There does not appear to be a consistent pattern of change within the psychological domains between the psychometric measures and the behavioural checklists. He received a score of 19 out of 24 on the Treatment Gain scale, which was 0.44 SD above the mean. Both his treating therapist and a custodial officer provided an overall rating of Satisfactory on the SBRG. His Treatment Gain score and SBRG ratings are consistent with his observed increases in positive behaviours according to the behavioural checklists.

4.3.4.17 Participant 17. The behavioural checklists reflected a self-reported positive change across Psychological Adjustment, such that there was a decrease in negative behaviour and an increase in positive behaviour. Custodial officers reported an increase in negative behaviours in Social Adjustment, with no self-reported significant changes in this domain. There was a self-reported decrease in negative behaviours in Manageability, with an associated increase in positive behaviours in this domain according to custodial officer observations.

Participant 17 demonstrated deterioration pre- to post-treatment in the CUSI, such that he remained in the dysfunctional range. Although he made a reliably significant deterioration in the SIS and WSFQ Intimate subscale, he remained in the functional range within both these measures pre- to post-treatment. He was in the functional range for the majority of the measures both pre- and post-treatment. Those measures in which he remained in the dysfunctional range post-treatment were: CISS (Avoidance-Oriented Coping and Avoidance Social Diversion subscales) and WSFQ (Sadomasochism subscale). His deterioration in the CUSI, which measures characteristics within the Psychological Adjustment domain, was inconsistent with his self-reported behaviours within the behavioural checklist data.

He received a score of 15 out of 24 for the Treatment Gain scale, which was 0.34 SD below the mean. He received a SBRG rating of Satisfactory upon completion of treatment from both his treating therapist and a custodial officer. Overall, he did not report many changes within the psychometric measures but reported generally positive changes through the behavioural checklists; whereas custodial officers reported both positive and negative changes within the behavioural checklists, which is consistent

with receiving both positive and negative ratings from staff in the Treatment Gain scale and SBRG.

4.3.4.18 Participant 18. No self-report data were available in relation to changes within domains through the behavioural checklists, as Participant 18 withdrew from the research too early for changes to be measured. However, based on custodial officer observations, he made positive changes over time. Within both Psychological Adjustment and Social Adjustment, there was a decrease in negative behaviour and an increase in positive behaviour. In addition, there was an increase in positive behaviours within both Mental Disorder and Manageability.

Participant 18 demonstrated significant improvement on the BRS; however, he was already in the functional range pre-treatment. On the other hand, he demonstrated significant deterioration on the SIS and WSFQ (Intimate and Sadomasochism subscales), such that he remained in the dysfunctional range post-treatment. He did not make reliable or clinically significant change and remained in the functional range pre- to post-treatment for the following psychometric measures: BMS; LS-UCLA; SSEI; and, CISS (Task-Oriented Coping subscale). He did not make reliable or clinically significant change and remained in the dysfunctional range pre- to post-treatment for the following measures: CUSI; CISS (Emotion-Oriented Coping and all Avoidance-Oriented Coping subscales); and, WSFQ (Impersonal and Exploratory subscales).

The results for the psychometric measures do not reflect a pattern in relation to the psychological domains. However, his deterioration in the SIS and WSFQ are contrary to the improvements that custodial officers observed within the Social Adjustment and Mental Disorder domains of the behavioural checklists. Consistent with the behavioural checklist results, he received a score of 18 out of 24 on the Treatment Gain scale, which was 0.25 SD above the mean. In addition, he was given an overall Satisfactory rating on the SBRG from both his treating therapist and a custodial officer. Overall, his self-reported results from the psychometric measures were inconsistent with the changes staff members observed through the behavioural checklists, Treatment Gain scale and SBRG.

4.3.4.19 Participant 19. No self-report data were available in relation to changes within domains through the behavioural checklists, as Participant 19 withdrew from the research too early for changes to be measured. There were no significant changes in any domains based on custodial officer observations.

The lack of significant changes over time was also consistent with the psychometric measures. Participant 19 did not demonstrate reliable or clinically significant change in any of the psychometric measures pre- to post-treatment. He remained in the functional range for the following measures: BRS; SSEI; CISS (Emotion-Oriented Coping and Avoidance Social Diversion subscales); and, WSFQ (Intimate subscale). He remained in the dysfunctional range for the following measures: BMS; CUSI; LS-UCLA; SIS; CISS (Task-Oriented Coping, Avoidance-Oriented Coping and Avoidance Distraction subscales); and, WSFQ (Impersonal, Exploratory and Sadomasochism subscales).

He received a score of 11 out of 24 on the Treatment Gain scale, which was 1.12 SD below the mean. At the completion of treatment, he was rated as Satisfactory on the SBRG by both his treating therapist and a custodial officer. Overall, the results from the psychometric measures, observer ratings within the behavioural checklists and the Treatment Gain scale were congruent with each other; however, the SBRG ratings were inconsistent with these measures.

4.4 Discussion

The overarching aims of the current study were to determine how incarcerated sexual offenders' risk factors manifest within a custody-based sexual offender treatment program, change over time, and are measured as a function of self-report and staff observations. Four assessment measures were used to monitor offenders' behaviour and the process of behavioural change: behavioural checklists; pre- to post-treatment psychometric testing; Treatment Gain scale; and, Satisfactory Behaviour Rating Guide (SBRG). These four measures were selected as they are representative of a variety of methods by which change can be assessed.

Interpretation of the results from this study should be made with caution due to the small sample size and a lack of follow-up data (e.g., recidivism). Follow-up data is relevant to determine whether the changes made using these methods correspond with an absence in recidivism. In addition, the behavioural checklist was a new measure with no prior empirical testing or validation. While the results should be regarded as tentative, they provide important information for future research into the use and further development of change measures in a custodial treatment program context. The results also have important clinical implications.

First, an overview of the results from this study will be provided, followed by more detailed interpretations in relation to each of the change measures. Practical implications will be outlined, with some suggestions for practitioners. Challenges and limitations of the current research will follow, after which, potential future research will be discussed.

4.4.1 Overview of Results

4.4.1.1 Group-level change. As per the first aim of the study, both participants and custodial officers reported the presence of behavioural manifestations of participants' risk factors through completion of the weekly behavioural checklists. Some behaviour was more evident than other behaviour; for instance, custodial officers tended not to report negative behavioural manifestations of Sexual Deviance (e.g., inappropriate sexual arousal). In addition, on average, participants were more likely to report the presence of positive behavioural manifestations of dynamic risk factors than were custodial officers.

As per the second aim, the behavioural checklists provided evidence of behavioural change over time in treatment, generally in the expected directions. Both participant self-report and custodial officer observations revealed an increase in positive behaviour over time within all four psychological domains. There was less consistency in relation to negative behaviour recorded through this measure. Further, contrary to the hypothesis, participants' and custodial officers' reports of positive behaviours within three domains became less consistent with each other over time.

As per the third aim, the psychometric measures overall elicited varying group-level change results. There did not appear to be consistency between the psychometric measures that reflected change over time and the RSVP domains into which the measures were classified; that is, within the same domains, participants made changes in some measures but not in other measures. Similarly, there did not appear to be discernible patterns in the representation of change across the four measures. While the Treatment Gain scale scores varied considerably between participants, therapists rated all participants as Satisfactory on the SBRG and custodial officers rated all but one participant this way. Of the four measures, the behavioural checklists provided the greatest depth of information about specific behavioural manifestations of risk factors and change over time in treatment.

4.4.1.2 Individual-level change. Reflecting on the different levels of analysis conducted, it was apparent that analysing data only at the group-level led to the loss of

information about individual offenders in treatment. It is likely that individuals make changes in their attitudes and behaviours in a range of different, perhaps nuanced ways (for instance, within different risk areas, captured by different change measures), for a variety of reasons, and at different speeds. This is consistent with prior suggestions that the differential impact of treatment on individuals is evident through some offenders reoffending post-treatment whereas others desist (Beggs, 2010).

A task that clinicians can face on a daily basis is to seek to understand whether a particular offender has changed. However, individual-level change is generally complex and often non-linear; therefore, identifying meaningful change is difficult. Davies et al. (2010) noted the general absence of detailed individual evaluation in the forensic field, with researchers tending to focus on the group-level. These authors suggested a number of basic requirements to determine change in an individual, including: clear specification of what is to be measured; measurement from multiple perspectives (e.g., psychometrics, behaviour, informant-report, self-report); using different methods (e.g., frequency counts, quality judgements); and, using atheoretical measures (i.e., not uniquely related to a specific model of therapy), which can demonstrate patterns of change over time. These requirements were fulfilled within the current study.

The individual-level results provided a complex representation of offenders' change over time. There were differences in the consistency with which changes were captured within and between the measures. These differences were apparent within individual participants' progression over time. For instance, some participants deteriorated on some measures while they improved on others.

4.4.2 Behavioural Checklists

The behavioural checklist will be explored first. Within this section, discussion will cover interpretations of the results relating to: associations between participants' and custodial officers' observations of each risk factor and associations between the two behavioural manifestations of each risk factor; group-level change over time, through participant self-report and custodial officer observations; differences between self- and observer-report over time; and, individual-level change.

4.4.2.1 Comparison between reports of behaviours. The results supported the first hypothesis, such that the behavioural manifestations of risk factors and their prosocial alternative behaviours were evident throughout the treatment period. There was a significant association between the frequencies with which participants and custodial officers reported the presence or absence of the risk factor behavioural

manifestations. That is, there was a greater likelihood of both participants and custodial officers similarly ‘never’ or ‘ever’ reporting behaviours, than of conflicting reports. Correspondingly, within the comparison between the two behavioural manifestations of each risk factor, there was a greater likelihood that the two behaviours were similarly reported as ‘never’ or ‘ever’ being present, than of one behaviour but not the other being reported.

Further investigation of the individual comparisons revealed that there were no significant associations between the participants’ and custodial officers’ reports of participants ‘never’ or ‘ever’ engaging in the behaviours. Prior to applying the correction for multiple comparisons, there were some statistically significant associations when comparing the two behavioural manifestations of each risk factor. However, the effects were small, meaning and a larger sample size would be required to attain statistical significance.

The results from the associations between the two behavioural manifestations of each risk factor suggested that for some of the risk factors, the two negative or two positive behaviours were observed with similar frequencies. That is, if one of the two behaviours was ‘ever’ reported, it was likely that the other behaviour was similarly ‘ever’ reported, and vice versa for ‘never’ reported. For example, within Problems with Stress or Coping, if ‘fluctuating emotions’ was ‘ever’ reported, it was more likely that ‘withdrawal/isolation’ was similarly ‘ever’ reported as opposed to ‘never’ reported. One implication is that the two examples of behavioural manifestations may not both be required in the measurement tool. As such, the behavioural checklist could be condensed in the future, which may be more acceptable to those routinely completing it (i.e., both staff and offenders).

As reported in section 4.3.1.1 (p. 180), several contingency table analyses could not be computed because at least one of the two variables was a constant, such that all instances were labelled as either ‘never’ or ‘ever’ reported. These analyses most commonly involved participants self-reporting that they ‘ever’ engaged in the risk factors’ positive behavioural manifestations (refer to Tables 27 and 28, pp. 181 and 184 respectively). This result suggests that participants were more likely to report engaging in positive behaviours than never engaging in these behaviours. Similarly, it suggests that participants tended to record positive behaviours to a greater extent than did custodial officers. This finding might indicate participants’ tendency to over-emphasise their positive behaviours.

Notably, analyses could not be computed for any associations between participants' or custodial officers' reports for either the negative or the positive behavioural manifestations of Problems with Stress or Coping. Within the negative behaviours, for both participants and custodial officers, there were no reports of 'never' engaging in the behaviours. This finding suggests that fluctuating emotions, withdrawal/isolation, talking about problems with supports, and good problem solving were commonly engaged in and were observed by custodial officers. Participants had no reports of 'never' talking about problems with supports or using good problem solving; on the other hand, there were some custodial officer reports of participants 'never' engaging in the behaviours. Again, this finding may suggest participants' tendency to report more positive behaviours than custodial officers.

There was overlap between behavioural manifestations of separate risk factors (e.g., withdrawal/isolation was a manifestation of both Problems with Stress or Coping and Problems with Non-Intimate Relationships). The examples provided on the checklists for each item were distinct from each other and were specific to the risk factor; however, this delineation was not provided on the checklists. Therefore, reporting engagement in a behaviour assigned to one risk factor might in fact have been an indicator of another risk factor with the same behavioural manifestation. The overlap in behavioural manifestations on the checklists was a result of the behaviours that participants identified as indicative of risk factors in Study 1; that is, there was similarity between some of the behaviours that participants suggested across different risk factors. This intersection between the behaviours has implications for use of such a checklist in the risk assessment process and evaluation of change. While functional analysis for each behaviour at each time point may decrease the measure's practicality, as it would become more cumbersome and custodial officers may not have access to this information or be able to conduct the functional analysis, it might be useful to determine the relevance of specific behaviours in certain situations (Vess, 2008).

There were some differences in the patterns with which risk factors were reported as present or absent. There were two risk factors for which the negative behaviours were never reported: custodial officers' observations of the two negative behavioural manifestations of Sexual Deviance (collecting inappropriate images and inappropriate sexual arousal); and, participants' self-report and custodial officers' observations of one of the two negative behavioural manifestations of Problems with Substance Abuse (positive urinalysis).

There are various potential explanations for these results. One explanation is that the risk factor is not relevant for offenders in this context. Alternatively, the risk factor itself might be relevant, but some behaviours might not be accurate manifestations of the risk factor. On the other hand, the behaviour might be an accurate manifestation of the risk factor, but it is not commonly observed within this context and, therefore, may not be a useful measure of change. In relation to custodial officers' observations of behavioural manifestations of Sexual Deviance, it is likely that although the behaviours are accurate manifestations and are commonly reflected in this context, they are less observable by others and there is more reliance on self-report (see Hanson & Harris, 2001).

4.4.2.2 Domains related to behavioural checklists. Mann, Thornton, et al. (2010) indicated that for sexual offenders, the range of behaviours with the potential to be offence paralleling, and which should therefore be monitored, is almost inexhaustible. This idea can similarly be applied to OABs and ORBs. In practice, this challenge was highlighted through the length of the behavioural checklist, which was created in an attempt to include a broad range of risk-relevant behaviours to adequately guide observation for custodial officers who are not necessarily trained in this task.

The quantity of data produced as a result of the length of the behavioural checklist had implications for the methods by which it could be meaningfully analysed. Therefore, the behavioural checklist data were condensed further to facilitate analyses of change over time (see section 4.2.5.1.1, p. 164). While the large amount of data and small sample size was not conducive to useful data analysis, it should be noted that consideration of the more detailed data might elicit more nuanced information in the evaluation of offenders' change over time. One limitation associated with the reductionist approach to analysis is that it is unknown whether there were particular behaviours within the domains, which influenced the results. However, use of grouped domains has been reported in prior research (Thornton, 2002) so the approach taken is considered valid and meaningful. In addition, Casey (2016) asserted that in prior studies, the effect sizes were typically stronger in relation to the predictive value of domain composite scores than individual test scores. This is also consistent with meta-analytic findings demonstrating single risk factors were not strongly enough correlated with sexual recidivism to be sole predictors (e.g., Hanson and Bussière, 1998; Hanson and Morton-Bourgon, 2005; Mann et al., 2010). Based on this combined evidence, the aggregation of data into domains was considered an appropriate method within the

current research. The risk factors were aggregated into the four RSVP domains: Psychological Adjustment (coping and problem solving); Mental Disorder; Social Adjustment (relationships and social influences); and, Manageability (cooperation with supervision). Throughout this discussion, the RSVP domains will be referred to by the previously bracketed descriptions of their underlying psychological constructs. Note that the Mental Disorder domain encompasses risk factors related to offenders' mental health and functioning, including sexual deviance, major mental illness, and personality disorder. As this domain label succinctly describes the underlying construct, this domain will continue to be referred to as mental disorder.

Results related to change over time in treatment will now be explored in reference to the first measure of change, the behavioural checklists. First, group-level results will be discussed, including participant self-report, custodial officer observations, and the differential change over time between the two reports. Individual-level results will then be considered.

4.4.2.2.1 Participant self-report group-level analyses. Overall, participants tended to report a greater increase in positive behaviours than a decrease in negative behaviours over time, as reflected through the figures depicting change over time (see section 4.3.1.2, p. 185). These results may indicate that participants tended to exaggerate their positive behaviour and minimise their negative behaviour. It is possible that a contributing factor to this finding was that participants had greater scope to report changes in positive behaviours, through continued increase in frequency over time. On the contrary, there may have been less potential to decrease negative behaviours and demonstrate change if participants were initially reporting behaviours infrequently.

In contrast to this overall pattern, participants reported both a significant increase and decrease in positive and negative behaviours respectively in relation to coping and problem solving. The following interpretation relates to the negative behaviours, as this change was in contrast to the negative behaviours within the remaining domains. This result may reflect that the relevant behaviours were more sensitive to change, or easier for participants to monitor. Behaviours such as fluctuating emotions, withdrawal/isolation, and verbal aggression might have initially been reported at greater frequencies compared with negative behaviours in the other domains; therefore, there was greater capacity for these behaviours to decrease over time. It is possible that participants felt more comfortable acknowledging these negative behaviours during the

earlier stages of treatment. Alternatively, these behaviours may be more commonly triggered within the custodial context (Mann, Hanson, et al., 2010).

The other three domains demonstrated an increase in positive behaviours but no significant changes in negative behaviours. There are various possible interpretations for this result. For instance, positive behaviours may have been more sensitive to change than negative behaviours, either due to the nature of the behavioural change process or through treatment more effectively targeting them. In other words, treatment might have a larger impact on developing positive behaviours compared with facilitating a decrease in negative behaviours. It is also possible that the positive behaviours included in this measure were better indicators of change than were the negative behaviours. While the positive behaviours increased over time, the negative behaviours may have remained. Note that some negative behaviours would be expected to decrease to a certain point but remain at low frequencies. Simultaneously, the related prosocial behaviours would continue to increase as ongoing management strategies are implemented. This process is commonly observed in treatment as offenders develop and practise new positive skills to manage risk-related behaviour.

4.4.2.2.2 Custodial officer group-level analyses. Further information can be gained about offenders' behavioural change through observer data. Results revealed a significant difference between the responses of the custodial officers who completed the checklists. The three custodial officers who demonstrated differences in the fewest domains completed the fewest checklists. The significant differences may suggest: methods of measuring weekly frequencies varied between custodial officers; some custodial officers may have focused on certain behaviours to the detriment of observing other behaviours; or, the differing quantity and quality of interactions with participants may have altered their ability to observe behaviours.

Consistent with participant self-report, custodial officers observed an increase in participants' positive behaviour in all domains over time in treatment. There was no observed improvement (i.e., decrease) in negative behaviours in any of the domains. Contrary to expectations, there was deterioration in mental disorder, such that custodial officers observed an increase in negative behaviours over time. As previously noted, custodial officers did not observe negative behaviours within Sexual Deviance (i.e., collecting inappropriate images and inappropriate sexual arousal) across the majority of time points. There was a combination of 'ever' observed and 'never' observed for the other three relevant risk factors: Psychopathic Personality Disorder; Major Mental

Illness; and, Problems with Substance Abuse. However, since this result does not reveal changes over time, the contribution of these risk factors to the increase in negative behaviours over time remains unknown.

There was an impact of covariates on custodial officers' observations. Participants with adult victims engaged in a greater number of negative behaviours overall in relation to mental disorder and cooperation with supervision. However, this variable was time-dependent, such that there was a decrease in negative behaviours for those with adult victims and an increase in negative behaviours for those with child victims. This result might suggest that treatment more effectively targets these risk factors for offenders with adult victims. Alternatively, it might relate to observer bias (Serin, Chadwick, et al., 2016). For instance, while offenders with child victims generally present as less visibly antisocial (Mills, Anderson, & Kroner, 2001), custodial officers might be biased against them due to their victim type (see Levenson et al., 2007) and focus more on their negative behaviours over time.

Further, in the behaviours related to cooperation with supervision, participants' number of days between their Earliest Release Date (ERD) and treatment commencement moderated the positive effect of treatment over time. That is, more negative behaviours were recorded for participants who commenced treatment further into their parole period. It is possible that these participants were less compliant due to increasing frustration, as a function of their desire for release, or with the demands of program participation.

4.4.2.2.3 Participant self-report versus custodial officer observations. It was hypothesised that initially, there would be a significant difference between participant self-report and custodial officer observations within all domains, with the two forms of reporting expected to converge as participants gained insight through treatment and were more engaged with staff in the therapeutic community. The results revealed a significant difference between the two reports in the rates of change over time, within the positive behaviours related to coping and problem solving, mental disorder and cooperation with supervision. However, contrary to the hypothesis, the difference increased, rather than decreased, over time. Although both participants and custodial officers reported an increase in the frequency of these behaviours, participants reported a significantly larger increase than custodial officers observed. Potential overarching explanations for these results will now be explored, after which, these results for each individual domain will be discussed.

It could be assumed that neither participants nor custodial officers were familiar with monitoring positive behaviour. Within an institutional context, negative behaviour tends to be the focus of attention, due to the risk-averse and punitive nature of the organisation (Polaschek, 2017). However, the treatment environment may have encouraged participants to increase their awareness of positive behaviours such that they were more adept at identifying and further developing them. Additionally, previous research has suggested that pre-treatment psychometric scores might be more predictive of future reoffending than post-treatment scores for various reasons, including the increased likelihood of offenders providing socially desirable responses as they progress through treatment (Barnett et al., 2012). Therefore, the results might have been related to participants' desire to present themselves in a positive light by reporting increasingly frequent positive behaviour.

Observing and reporting positive behaviours are often neglected within the custodial context (Mooney & Daffern, 2013). The behavioural checklists should have served to focus custodial officers' attention to positive behaviours. Their observations of participants' positive behaviours did increase over time in treatment; however, the slower rate of increase compared with participant self-report might have been a function of custodial officers' roles within custody. Due to their primary role including the maintenance of the safety and security of the prison, they may have continued to focus more on negative behaviour, to the detriment of observing positive behaviour. Alternatively, custodial officers' reports of positive behaviour may have been a more realistic representation of behavioural change, which was not affected by socially desirable responding. The general agreement between participants' and custodial officers' reports of negative behavioural change over time may reflect greater accuracy in the identification of negative behaviour compared with positive behaviour.

It is of interest that the exception to these results was behaviours related to relationships and social influences. The individual behaviours and risk factors subsumed by the domains might provide further information to explain these results.

Starting with coping and problem solving, the behavioural manifestations of risk factors within this domain included both internal and external characteristics. For instance, 'victim awareness (understanding the harm of offending)' from the risk factor Attitudes that Support or Condone Sexual Violence and 'self awareness related to impact of past abuse/trauma (e.g., discusses emotional reactions and impact on his life)' from the risk factor Problems Resulting from Child Abuse, represent internal factors,

which custodial officers may be unable to observe. Custodial officers would have had more awareness of behaviours within Problems with Stress or Coping, such as ‘talking about problems with supports’ and ‘good problem solving’. Therefore, participant self-report included opportunities to provide information about a wider range of attitudes and behaviours, which may have led to a greater increase in reported behavioural frequencies compared with custodial officers.

Turning to mental disorder; this domain included the risk factor Sexual Deviance. While observers rarely reported the behavioural manifestations of this risk factor, participants’ awareness of these behaviours would likely have been greater. As participants progressed through treatment, their development of skills to manage problematic behaviours related to sexual self-regulation would likely have increased (e.g., using coping strategies to manage sexual deviance, such as replacing deviant thoughts with appropriate thoughts and managing sexual arousal in a healthy way). However, custodial officers may have remained unaware of such changes. Therefore, changes in this risk factor likely contributed to these results.

Of the three domains demonstrating these differential changes between participants’ and custodial officers’ reports, cooperation with supervision revealed the largest difference. Polaschek (2017) indicated that some items in protective factor scales refer to characteristics that may reflect offenders’ responsiveness to the influence of interventions, or their willingness to develop a positive relationship with staff, which would facilitate their influence on the offender. It is possible that this concept is relevant for the positive behaviours related to cooperation with supervision. That is, custodial officers’ perceptions of certain participants might have become more positive over time, through an increase in these participants’ willingness or ability to engage appropriately with staff. The presence of these responsivity factors might have required less effort from staff members in their engagement with participants, which further increased custodial officers’ positive regard for those individuals. Not only would these responsivity factors lead to higher frequencies of positive behaviours, but perhaps positive interactions led to positive biases such that custodial officers were more likely to report these adaptive behaviours.

Finally, it is also important to explore the only positive behaviour domain that did not reach significance; that is, relationships and social influences. It is possible that participants and custodial officers reported comparatively similar increases in these positive behaviours, as they were equally aware of the relevant behaviours. For

instance, there were behaviours related to employment, which was operationalised within treatment as completion of domestic duties. Custodial officers were generally more attuned to this behaviour as it formed a central component of their role within the wing; that is, overseeing the completion of domestic duties. Further, with respect to relationships, custodial officers might have had greater exposure to participants' communication and interaction styles. This result might suggest that custodial officers' observations of these behaviours are more useful in custody.

4.4.2.3 Individual-level self-reported and observed change. The current results revealed that it was more common for participants to report behavioural change than for custodial officers to observe change. There were also individual differences in the patterns of behavioural change across domains, both according to self-report and observers. For instance, some participants made changes in all domains, whereas others demonstrated less widespread change. These differences between participants' behavioural change over treatment highlights the importance of examining individual patterns of change, rather than relying on group-level changes to draw conclusions about treatment processes and the use of measures of change.

Further, either internal or external processes that might differentially affect individuals can be monitored through these analyses. For instance, each offender in CUBIT is assigned a domestic duty. However, some jobs are completed on a weekly basis whereas others are required daily, with some multiple times a day. Therefore, any change in the number of times this behaviour was engaged in during the preceding week might be indicative of the regular reassignment of domestic duties, rather than risk-related behavioural change. Information about individual offenders' fluctuations over time would be more beneficial in these situations than would observed changes in the group as a whole. Additional measures of change should also be evaluated to assist with the interpretation of the information gained.

4.4.3 Psychometric Measures

The second form of change assessment in this study consists of the pre- to post-treatment psychometric measures. Group-level change will be explored, after which, individual-level change will be considered through the reliable and clinically significant change analyses. Subsequently, discussion will turn to the application of the RSVP domains to the psychometric measures, to facilitate comparison with the behavioural checklist results. The general use of psychometric measures as assessments for treatment change will be examined, with reference to the results of the current study.

4.4.3.1 Group-level change. At the group-level, the BRS and BMS reflected participants' improvement from pre- to post-treatment, as did the LS-UCLA and SSEI. The BRS and BMS results revealed that throughout treatment, participants developed their prosocial attitudes and beliefs about sexual offending against women and children. Participants' expressed attitudes might have reflected genuine change through gaining insight into the problematic nature of prior attitudes. Alternatively, these attitudes might have reflected greater insight into socially acceptable attitudes, which would be consistent with prior research suggesting post-treatment psychometric results are more susceptible to positive impression management (e.g., Cording et al., 2016). In order to determine whether the expressed changes over time are meaningful or related to participants' desire to present themselves in an overly positive light, it might be beneficial to compare these results with those from other measures. It also highlights the potential use of individual-level analysis, as the reasons underlying the findings might differ between offenders.

The LS-UCLA and SSEI results revealed that there was an improvement in participants' sense of intimacy with others, and in their self-concept/sense of self-worth. It would be expected that those whose self-concept improved, would be more adept at engaging more meaningfully with others, which may decrease their feelings of loneliness and increase their sense of connectedness.

Less promising results were revealed through the CISS. Both the Avoidance-Oriented Coping and Avoidance Social Diversion subscales demonstrated participants deteriorated from pre- to post-treatment. This result will be discussed in more detail in relation to the individual-level analyses (see section 4.4.3.2, p. 241).

As Nunes et al. (2011) noted, group-level analyses are not sensitive to the presence of dysfunction post-treatment, which may lead to more positive results than those revealed through individual-level analyses; therefore, there are limits to the conclusions that can be drawn in relation to change through treatment. If participants remain in the dysfunctional range post-treatment, it would be indicative that additional changes were required. Further, group-level analyses do not reflect pre-treatment functioning and could similarly lead to overly positive conclusions about treatment change. For instance, if participants commenced treatment in the functional range, further changes are less critical than if they commenced in the dysfunctional range. This concept demonstrates the importance of having clarity about the aims and purpose of research being conducted. The current research involved measurement of change over

time and a comparison of various change measures, rather than treatment effectiveness. The changes remain important even if dysfunction remains evident, as it reflects the offender has greater ability to manage problematic attitudes and behaviours. This concept is relevant for the analysis of change at the individual-level.

4.4.3.2 Individual-level change. There is no validated method for integrating disparate findings from clinically significant change calculations between different psychometric measures. Therefore, when offenders achieve clinically significant change on some measures but not on others, difficulties remain in the combined interpretation of these results (Daffern & Ogloff, 2017). The most common outcome within the psychometric measures was for participants to demonstrate no reliable or clinically significant change pre- to post-treatment. These results are comparable with some prior research but in contrast with other research. For instance, Wakeling et al. (2013) reported that in measuring clinically significant change, on the majority of measures, the largest proportion of offenders remained unchanged (15.9-63.3%). On the contrary, within Beggs' (2010) meta-analysis, within-treatment change research revealed significant levels of improvement, assessed through pre- to post-treatment psychometric measures. Ordinarily, pooled results across multiple studies, such as within a meta-analysis, would be considered a more accurate representation of the outcome than would a single study such as Wakeling et al. However, the reported studies in Beggs' meta-analysis used various methods by which to measure pre- to post-treatment change and many did not use clinically significant change. For instance, some studies calculated change through subtracting post-treatment raw scores from pre-treatment raw scores (e.g., Hudson et al., 2002). Use of raw scores can be problematic due to pre-treatment variance (Beggs, 2010). Therefore, the overall results from this meta-analysis are perhaps less useful as a comparison with the current study.

In the current research, participants' limited change over time revealed through the pre- to post-treatment psychometric measures was due to a combination of two primary factors: on some measures, many participants commenced treatment in the functional range; and, on other measures, many participants commenced in the dysfunctional range and remained at this level. Some possible reasons and implications for these results will now be explored.

The majority of participants were already in the functional range pre-treatment in the BRS. In the BMS, participants were generally either in the functional range pre-treatment or remained in the dysfunctional range post-treatment. These results are

consistent with Nunes et al. (2016), who found that the majority of offenders were already in the functional range pre-treatment in both the BRS and BMS. Such results raise questions regarding the benefits of these measures as indicators of change through treatment. Although some prior research has supported the use of these scales (e.g., Blumenthal et al., 1999) and the majority of prior studies have used them to assess offenders' attitudes about sexual behaviour with women or children, these measures have been criticised for their transparency and susceptibility to response bias (Grady et al., 2011).

Contrary to the criticism that the BRS and BMS are susceptible to socially desirable responses, the MC-SDS results revealed no indication of participants providing such responses. Therefore, participants may in fact have already been in the functional range pre-treatment. If so, it might suggest one of two potential scenarios: either these measures are not appropriate indicators of risk for individuals deemed eligible for this particular treatment program, as they do not properly measure the factors that identify an individual as requiring treatment within CUBIT; or, these individuals do not have needs in these areas and, therefore, are unnecessarily receiving treatment in relation to these factors. Although there is evidence for the validity of these measures, there may be a discrepancy between the constructs being measured and the determination of treatment suitability within CUBIT. This may reflect the need for careful consideration of the particular psychometric measures used within treatment programs, to ensure the relevant risk factors are targeted.

Within the CISS, participants tended to remain in the dysfunctional range post-treatment in the Avoidance-Oriented Coping and Avoidance Social Diversion subscales. There is no clear interpretation of these results. It is possible that they suggest participants were more truthful in their responses, as compared with the BRS and BMS. The CISS might be less transparent, such that participants were less sensitive to the messages contained in the items. Alternatively, participants may have felt more comfortable discussing problems with coping. Participants might have perceived the potential for less negative outcomes through reporting negative behaviours within this measure, as the items may have been viewed as less directly related to sexual offending and consequently less confronting.

Coping is often a core treatment-target and the observation of relevant changes would be expected; however, perhaps there is a discrepancy between the skills developed in treatment and the relevance or validity of the specific questions within this

measure. That is, participants might have developed their coping skills, but these skills might not have been consistent with their measurement in the CISS. For instance, offenders in CUBIT develop coping skills such as thought challenging, distress tolerance, assertive communication, and consequential thinking. There is also discussion regarding the use of distraction strategies (e.g., watch television, speak to a friend), which can be beneficial in the short-term, prior to focusing on the problem. Within the CISS, items related to avoidance-focused coping include activities such as watching television, calling a friend, and spending time with others. The CISS does not capture the function of these strategies and the context within which they may be used. Therefore, if participants endorsed a larger number of these strategies post-treatment, it might not reflect deterioration; rather, it might reflect their use of these activities as an intermediate strategy prior to implementing solution-focused coping skills.

More generally, participants' apparent deterioration might have been a function of the treatment process. For instance, treatment may have facilitated participants' development of a greater awareness of problematic attitudes. Their responses may have reflected this increased awareness, rather than a deterioration in their attitudes. Alternatively, participants' comfort in reporting unhelpful attitudes may have increased, similarly as a function of discussing them throughout treatment.

Although there were no significant differences at the group-level for the WSFQ, at the individual-level, there was a larger proportion of participants who deteriorated on the Intimate subscale than on any other measure. It is possible that participants were more likely to provide non-socially desirable responses on this scale compared with the other WSFQ subscales because the items on this subscale are less confronting and may be perceived as less aligned with sexual offending (i.e., less 'deviant'). It may demonstrate parallels with prior research. For instance, Baumgartner (2002) reported that sexual offenders gained higher scores on the Exploratory and Intimate subscales than did non-sexual offenders, while there were no significant differences on the Impersonal and Sadomasochism subscales.

4.4.3.3 Domains in relation to psychometric measures. As none of the psychometric measures were considered suitable for classification into cooperation with supervision, there were some limitations in the domain comparisons between the behavioural checklist and psychometric measure results. However, prior research has reported the use of domains when measuring results from a psychometric battery: Distorted Attitudes; Sexual Interests; Socio-Affective Functioning; and Self-

Management (see Thornton, 2002). Allan et al. (2007) reported that they used these domains despite lacking measures within one of the domains, as their measures were closely related to the remaining three domains. They indicated that some, but not all, domains contributed additional variance beyond the STATIC-99 in predicting sexual recidivism; therefore, there may be value in using domain scores within this context.

In the current study, there were some group-level changes in relation to coping and problem solving, and relationships and social influences; however, these changes were not consistent across psychometric measures within the domains. In other words, there were some measures within these domains that revealed significant changes over time, while other measures did not reveal changes. There were no changes at the group-level within mental disorder, which was measured only through the WSFQ. It is worth exploring potential reasons for the inconsistent results between the psychometric measures within domains. First, it is possible that some of the psychometric measures did not target the factors relevant for those domains into which they were classified. Further, it may indicate that the treatment program does not target the specific factors measured by the particular psychological test; perhaps treatment more effectively targets factors in some of the measures than in others, thus providing inconsistent results within and between the domains.

4.4.3.3.1 Treatment targets. Grady et al. (2011) indicated that without valid instruments, it is challenging to determine whether sexual offenders who completed treatment actually demonstrated changes in the core target areas. As outlined in section 4.2.2.2 (p. 147), there is evidence for the validity of the psychometric measures used in this study. However, the current study was not designed to assess treatment effectiveness; therefore, there was no differentiation between change in core treatment target areas and change in general risk areas for sexual offending. Despite this lack of focus on treatment targets, since CUBIT participants routinely complete the pre- and post-treatment psychometric measures as part of the program, it would be expected that the measures are related to the treatment components. However, it is possible that not all the psychometric measures are compatible with specific treatment targets. Further, at the individual-level, not all offenders will necessarily have particular treatment needs in areas targeted by each psychometric measure.

4.4.3.3.2 Group- vs individual-level results. The diversity within domains according to the group-level results was not commensurate with all of the individual-level results, which is consistent with prior research in which group- and individual-

level analyses revealed disparate results (e.g., Nunes et al., 2011). As group-level results can obscure individual differences in the change process, the benefits of exploring individual-level results are highlighted (Serin et al., 2013).

In the current study, at the individual-level, there were more consistent findings between the psychometric measures associated with relationships and social influences. For instance, there was a higher proportion of participants who recovered post-treatment across the LS-UCLA, SIS, and SSEI, as compared with the measures in the other domains. This outcome is contrary to prior research, in which intimacy deficits were not observed to change significantly through treatment, specifically in relation to the LS-UCLA and SIS (Harkins, Abracen, Looman, & Maillet, 2011). Of note is that Harkins et al. (2011) reported exclusively on group-level analyses. The current results might suggest that the CUBIT program more effectively targets relationship skills. On the other hand, the behavioural checklist results did not reflect this outcome within relationships and social influences. These inconsistent results might indicate there were differences between the ways in which the behavioural checklist and the psychometric measures assessed specific risk factors.

4.4.3.4 Measuring dynamic risk factors. One potential question in relation to the results of the psychometric measures is whether these tools are in fact measuring dynamic risk. If there is limited change pre- to post-treatment in a particular measure, one underlying reason could be that the factors being measured are static rather than dynamic (Cording et al., 2016); this may suggest the measures should not be used in change assessment.

Alternatively, while the factors being measured may be dynamic, they may not be predictive of recidivism. Prior research has noted that psychometric scores could predict outcome as effectively as other risk assessment tools, provided that the psychometric measures used were designed to assess constructs that were empirically related to risk of recidivism (Walters, 2006). This suggestion has implications for the current research. The pre- to post-treatment psychometric measures do not exclusively assess constructs related to risk of recidivism. For instance, the SSEI measures self esteem, which does not directly predict sexual offending (Hanson & Morton-Bourgon, 2005). However, as Olver and Wong (2013) argued, targeting non-criminogenic needs in tandem with primary criminogenic needs may be indirectly beneficial for the reduction in problematic behaviour. Although there is limited evidence for self esteem predicting sexual recidivism, it may be an important area to address in treatment to increase

offenders' motivation for treatment and to facilitate the development of further skills (e.g., relationship development).

4.4.3.5 Use of normative data. As outlined in section 4.2.5.2.2.2 (p. 175), the calculations for clinically significant change require use of functional and dysfunctional normative data (norms) for each psychometric measure. In forensic psychology literature, these norms are gained from the non-offender ('normal') and offender populations respectively. The cut-off score and consequently the use of clinically significant change as a method of analysing individual-level change, is only as effective as the norms on which the analyses are based (Jacobson & Truax, 1991). Finding representative norms is straightforward for more general and well-established measures (e.g., LS-UCLA); however, it can be more challenging for measures designed for specific offender groups such as sexual offenders (Nunes et al., 2011).

Some of the offender norms used in the current research had more functional scores than the normal population norms (i.e., BMS and WSFQ Intimate, Exploratory, and Sadomasochistic subscales). Consistent with prior research (e.g., Allan et al., 2007; Baumgartner et al., 2002), some of the means calculated for the measures in the current study were similarly more functional than the normative data. Therefore, there are implications for the use of these measures as indicators of treatment need, and the use of post-treatment change outcomes; that is, it might not be appropriate to use these measures for such purposes if there is no treatment need in the relevant areas.

These counterintuitive scores might relate to Wakeling et al.'s (2013) suggestion that some psychometric measures (e.g., loneliness, social intimacy) are not assessing offender population constructs. Nunes et al. (2011) noted in their research that some of the norms used may more closely resemble the populations of interest than others; for instance, some normative samples have consisted of university students, while others were custodial officers. These samples might be differentially representative of dysfunctional or functional populations. While the BMS was developed for use within sexual offender populations, the WSFQ was not specifically developed for this population.

Additionally, the development of the clinically significant change method of analysis was based on the assumption that the normative data for both the functional and dysfunctional populations had normal distributions (Jacobson & Truax, 1991). However, some of the data distributions in the current research were skewed. It is possible that this was also the case for the data from some of the populations used for

the cut-off scores; however, there was no information available about the data distribution in those studies. Evidently, further validation work is required with measures routinely used to assess sexual offenders, and normal population norms are needed if the clinically significant change method of change analysis is to be used.

The difficulty interpreting some results provides rationale for using multiple change measurement tools. Similarly, the transparency of some self-report psychometric measures and the potential challenges associated with social desirability, highlight the benefits of using additional change measures rather than relying solely on one format. However, while using multiple measures allows for a comparison between them, the difficulties with combined interpretation must be considered to develop a more comprehensive representation of change.

4.4.4 Integrating Change Measure Results

The third hypothesis related to the convergence of the four change measures. The four measures had different properties, measured different attitudes and behaviours, and measured change in different ways. Therefore, without an established method for the process, comparison between the four measures is challenging. The convergence between the group-level results will be discussed, as will the differences, which will be explored through the concept of fluctuation in changes over time. Discussion of the individual-level results will include examples in relation to specific participants.

Olver et al. (2007) suggested that not all dynamic risk factors are equally dynamic, such that some may be more resistant to change than others. Therefore, the changes in intensity and frequency of behavioural manifestations of risk factors might similarly demonstrate differences. This differential ability to change might provide an explanation for the varied results in the current research, with some behaviours, risk factors and domains changing to a greater extent than others. Risk factors' differential ability to undergo changes over time might be relevant in clinicians' attempts to evaluate clients' changes subsequent to intervention or prior to release from custody.

Further, Mann et al. (2010) suggested that the strength of a behavioural manifestation is dependent on the strength of the underlying propensity, but it is also dependent on the extent to which the environment triggers that particular propensity. For instance, due to the context, sexual interest solely in children would presumably be triggered to a lesser degree in custody than in the community, with fewer observable behaviours evident. Subjective sexual arousal might also be a manifestation of this propensity, but in the custodial environment it would likely be less observable to others.

Therefore, although the risk factor may remain active, custodial officers would have relatively few opportunities to directly observe its manifestations. Greater reliance might be placed on offender self-report for these behaviours and it is important for clinicians to consider the potential limitations (e.g., the offender might seek to present himself in an overly positive light through reporting only appropriate sexual thoughts and behaviours, or he may have minimal insight into his behaviour). There may be an argument for reducing some structure within treatment environments in custody, such that they more closely resemble the community, in which there are more opportunities for behavioural observation of risk factors related to sexual self-regulation (e.g., visits that can be monitored by custodial officers within the therapeutic environment). If there are more available opportunities, it may be easier to determine whether an offender is improving, as his responses to these potential opportunities could be monitored. However, a balance would be required to ensure that the safety and security of the institution are maintained.

Different measures might evidence change in different ways, such that there may appear to be changes in a particular risk factor through one measure but not another measure. This difference between measures was observed in the current research. For instance, as described previously, there were inconsistent results between the behavioural checklists and psychometric measures with respect to relationships and social influences. This difference might be related to the ways in which the risk factors are measured.

In addition, in their assessment and evaluation, some risk factors might be emphasised over others. Generally, the relative weight assigned to specific behavioural manifestations will be determined based on recidivism data, which provides information about the risk factors that have the strongest association with subsequent offending. For instance, factors related to sexual self-regulation (e.g., sexual deviance) are the greatest predictors of recidivism (Hanson & Morton-Bourgon, 2005). However, as Hanson and Harris (2000) noted, dynamic risk factors must be observable in order to be useful to therapists or supervising officers. Unreported behaviours will not provide useful information about change over time. One practical challenge for risk assessment and change measurement is that some risk factors tend towards either more or less observable behavioural manifestations (Hanson & Harris, 2001). In the current study, custodial officers tended to report more observable behaviours such as completion of

domestic duties, abiding by rules, and withdrawal/isolation, rather than less observable behaviours such as those related to sexual self-regulation.

Prior research has provided information regarding the process of change and the tendency for offenders' behaviour to fluctuate as they make changes. For instance, Polaschek (2016) reflected that desistance appears to be a fluctuating process. Serin and Lloyd (2009) indicated it should be expected that individuals who desire to make changes will experience occasional hindrances and decreases in motivation due to external and internal barriers they encounter. Similarly, Davies et al. (2010) suggested that some changes may 'come and go' or may be affected by daily external (e.g., environmental) or internal (e.g., mental state) fluctuations. They noted that once change appears to have gained stability over time in the same context, the generalisability of the change can be further explored. It is possible that different change measures might differentially reveal these setbacks, which would subsequently be observed through different change patterns between measures.

This fluctuation over time highlights the complexities associated with change measurement. There will likely be a combination of improvement, deterioration, and stagnation throughout offenders' attempts to achieve change over time, which may be evidenced across and within different measures of change. In practice, the magnitude of this fluctuation can be monitored and forms part of a clinician's overall assessment of the offender post-treatment; for instance, progression over time would be reflected by diminished fluctuation even if stability is not achieved.

The change in the magnitude of an offender's fluctuation requires a sensitive measurement tool, which is elusive. A more effective evaluation of this fluctuation might be gained through observation of behavioural frequencies. That is, although behaviours might be present at various times throughout treatment, their frequency might increase or decrease over time. As change is a gradual process rather than a sudden and discrete event (see Serin & Lloyd, 2009), it is expected that negative behaviours will be present over an extended period of time but their strength and frequency might decrease. Capturing this fluctuation is beneficial within a clinical context, as it provides a richer representation of offenders' change through treatment. The behavioural checklist provides this information, whereas the pre- to post-treatment psychometric measures, Treatment Gain scale and SBRG provide only an overall evaluation of change.

Further, protective factors can co-exist with risk factors, which can lead to variability in outcome in a group of offenders with similar risk levels (Serin, Chadwick, et al., 2016). It could be anticipated that as the change process progresses, this co-existence of risk and protective factors increases, prior to adaptive attitudes and behaviours finally replacing risk-related attitudes and behaviours. This co-occurrence of behavioural manifestations has implications for change over time within a single risk area/domain, such that it would be possible for both the negative and positive domain scores to increase or decrease over time.

These ideas could provide an explanation for apparently disparate results between and within change measures in the current research. Within the behavioural checklists, custodial officers observed an increase in both negative and positive behaviours associated with mental disorder, which provides support for the co-existence of risk and protective factors. Within this domain, it would be possible for the item 'extreme mood states' to co-occur with 'medication compliance'. For instance, the behaviours associated with extreme mood states may not be a result of medication non-compliance. Similarly, the behaviours related to low or elevated mood may have alternative underlying functions, such as expression of anger, demand avoidance (i.e., an attempt to avoid a direction to engage in an activity), or attention-seeking. Alternatively, participants' medication might not be effective and could lead to an increase in the frequency of extreme mood states. Mental disorder can also be explored through the WSFQ results. There were varying results with respect to the four WSFQ subscales, which may parallel custodial officers' observations within the behavioural checklists and may provide further support for the co-existence of negative and positive attitudes and behaviours. It exacerbates the difficulties associated with the assessment of change over a relatively short period of time (e.g., throughout treatment), as changes might be concealed through the co-existence of problematic and positive behaviours.

This fluctuation in change over time and the co-existence of risk-related and adaptive behaviours raises questions about how clinicians and decision-makers establish which behaviours and observations of behavioural change to prioritise over others when seeking to determine whether an offender has changed. The examples that have been discussed provide further support for the evaluation of change at the individual-level, including whether it is a function of the offender's skills and decisions, or a function of external factors.

The complexities around this decision-making process were also evidenced through the discrepancy between sanctions imposed in treatment and the final ratings on the Treatment Gain scale and the SBRG. For instance, all three participants who were suspended from treatment for misconduct attained a Satisfactory rating on the SBRG. One participant's Treatment Gain score was above the mean, another was unavailable and the third was below the mean. It is worth noting that the participant whose Treatment Gain score was above the mean, was generally compliant and motivated within treatment, while the participant whose score was below the mean, was suspended on multiple occasions and experienced significant ongoing difficulties within treatment. It is possible that in parallel with the concept that poor behaviour leads to greater room for improvement, an offender could display behavioural problems leading to suspension from treatment, but subsequently demonstrate improvements upon return to treatment, which elicit a final positive rating from staff. This possibility was highlighted through the results for the participant described above who received a Treatment Gain score that was below the mean, but for whom a positive rating was provided in the SBRG. Although it was beyond the scope of the current research, it might be of interest to gain greater insight into the decisions therapists and custodial officers made in providing their overall ratings in the Treatment Gain scale and the SBRG.

Results were previously discussed in relation to the difference between participants' and custodial officers' reports of change over time in the positive behaviours associated with cooperation with supervision (section 4.4.2.2.3, p. 238), which relates to offenders' motivation and engagement with staff. Consistent with Polaschek's (2017) suggestions regarding the impact of offenders' responsivity and willingness to engage with staff, participants' Treatment Gain scores could be related to changes in their cooperation with supervision. If this relation were present, it would be expected that more positive change in cooperation with supervision would be associated with higher Treatment Gain scores. Nevertheless, there did not appear to be consistency between participants' and custodial officers' reports of behavioural change and therapists' Treatment Gain scale ratings. This disparity between observations is reflective of the general difficulties clinicians might encounter when attempting to gain an overall understanding of an offender's change processes through treatment.

There was a positive correlation between the psychometric measures on which participants recovered and the Treatment Gain scores. Although this result does not distinguish between the individual psychometric measures, it might be a reflection of

overall treatment engagement. These results may suggest there is a relation between improvement on psychometric measures generally and the treating therapist's assessment of change. Further, most participants received overall Satisfactory evaluations from their therapist and a custodial officer on the SBRG in spite of other more negative results. Taken together, these results may indicate that the self-report measures are as accurate as a clinician's rating; alternatively, it may suggest that clinicians' judgements are heavily influenced by offenders' verbal accounts of their progress. Again, to assist with drawing conclusions regarding the implications of these results, it may be useful to gain information about the key factors therapists and custodial officers used to provide their Treatment Gain scale and SBRG ratings.

4.4.4.1 Individual-level change. Rather than remaining reliant on group-level change, Davies et al. (2010) suggested that individual and group approaches should be viewed as complementary. It is also beneficial to consider the integration of the different change measures at the individual-level. The following subsection will explore this process, which will lead into a case study that outlines the process a clinician might undertake in order to assess an offender's change when multiple measures are used.

Davies et al. (2007) identified that single case approaches are based on a small number of guiding principles, which relate to a focus on the individual, the use of systematic methods to gather and analyse data, and the importance of change over time. They noted that in an individual case, the therapist or researcher might conclude that a client improved following treatment; however, the observations used as evidence might be unsystematic. It has been suggested that the evaluation of change requires the consideration of various factors: risk-related behaviour; prosocial behaviour; detection evasion skills; and, unique custodial reactions (i.e., behaviours that would not be expected in the community). It has been asserted that behavioural improvement, leading to a reduction in risk, should only be concluded when there is evidence of both a reduction in risk-related behaviours and the development of prosocial alternative behaviours. If negative behavioural manifestations persist in the absence of positive behaviours, it should be assumed that no change in risk level has occurred. Further, if there appears to be a reduction in negative behaviours in the absence of positive behaviours, it might be indicative of detection evasion skills. If an individual's repertoire of problematic behaviour appears to be increasing through persistent or increased risk-related behaviour, this individual might be exhibiting signs of deterioration, which could lead to increased risk (Jones, Daffern, & Shine, 2010). The

current research highlighted the need for further research to determine the most appropriate method by which to systematically gather and analyse data from multiple measures.

The evaluation of individual-level change allows for additional insight into the process clinicians undertake to determine change. In determining how to use the available information about individual offenders, the focus must be on each offender's primary areas of risk. Davies and colleagues identified an important question that must be answered when evaluating an offender's change: "Is the change significant, meaningful and relevant?" Prior to the assessment of individual change, the factors that would be expected to change through treatment should be identified, with a focus on how the change might affect the observed behaviours. Klepfisz et al. (2017) similarly noted the importance of considering the impact or relevance of a risk or protective factor, rather than simply whether it is present or not. This idea highlights the added value of individualised assessments.

When assessing change over time, it is important for clinicians to remain mindful of offenders' pre-treatment level of functioning in any given risk factor. An offender's pre-treatment functioning will impact the potential amount of change, which might be achieved through treatment. One offender might make a significant change in an area in which he started from a highly dysfunctional level. Despite this change, he might remain high risk post-treatment (dysfunctional) on that variable. On the other hand, another offender might be at a lower risk level pre-treatment in the same area, in which he makes a smaller amount of change through treatment. Post-treatment, he might be classified as low risk on that variable, despite having made a smaller amount of change than the previously described offender. Although the first offender made a larger amount of change, his post-treatment risk level has implications for further treatment/supervision. Similarly, one individual might demonstrate these differential changes on various measures; therefore, clinicians must determine which areas to prioritise in making judgements about overall change and risk levels.

Another difficulty in the evaluation of individual change is that although counterintuitive, for some risk factors, a decrease in negative behaviours (or a consistent absence) would similarly be reflected by fewer observations of the associated positive behaviour. If the positive behaviour was related to management of an associated negative behaviour, but the negative behaviour was not relevant for the individual, there would similarly be no positive behaviour; in these instances, the negative behaviour

would not require management. For example, an offender would not be expected to use coping strategies to manage deviance if he did not experience sexually deviant thoughts.

On the other hand, during the earlier stages of treatment, as offenders develop more adaptive behaviours to manage risk-related behaviours, both negative and positive behaviours could be observed. Through treatment, more consistent implementation of adaptive behaviours would be expected. In some risk areas (e.g., coping), these management strategies would be implemented such that the risk-related behaviours are no longer exhibited. In these situations, towards the latter stages of treatment, the positive behaviours (e.g., talking about problems with supports) would be expected despite the absence of the associated negative behaviours (e.g., withdrawal/isolation). The individual's specific criminogenic needs (and associated OAB/ORBs) would need to be assessed to determine which scenario would be expected (i.e., absence of negative and positive behaviour or absence of negative but presence of positive behaviour).

Another consideration is that there might be a fluctuation in an offender's behaviour in response to external factors, such as the content of group sessions. For example, in CUBIT, towards the start of treatment, offenders are asked to disclose to the group their offences. An offender's response to this task might have an impact on his behaviours recorded on the behavioural checklist. For instance, after presenting this task, he might report an increase in negative behaviours such as withdrawal/isolation, compared with the previous week. These behaviours might have been a result of: (1) spending time reflecting on the task work to the detriment of interacting with others; and/or, (2) his emotional response to this process. The tendency to spend time engaging with treatment work rather than social interactions with others might be observed throughout treatment, such that behavioural frequencies appear to change at irregular intervals; however, these changes might not represent actual change, rather, it might be a function of external events. On the other hand, if the offender's behaviour is a function of his emotional response to the treatment work, his behaviour may change as he develops more effective coping skills to deal with stressors. Therefore, these fluctuations would be an expected representation of actual change. These two explanations may be differentially observed such that the first might manifest through ongoing fluctuation over time, while the second might manifest through a decrease in negative responses over time. While the behavioural checklist may capture this information, the other change measures would not be sensitive enough to offenders' specific behaviours or fluctuations over time. For instance, the pre- to post-treatment

psychometric measures could only be used to determine overall change in coping and relationship skills. It may be useful to determine this overall change through the pre- to post-treatment measures and further investigate individual offenders' weekly fluctuations through the behavioural checklist results.

In the current research, individual-level change in the behavioural checklists revealed differences between individual offenders' treatment outcomes. These different findings may reflect some measures of change are more appropriate for some but not all offenders, or that treatment gain is observed in different areas for different offenders based on additional factors. Some of these factors may not be known, but some may be similar to those included in this research, such as victim type or risk level. Further, in relation to individual-level change processes, Daffern and Ogloff (2017) noted the tendency for non-significant associations between reliable and clinically significant change, and recidivism outcomes. They asserted that while single psychological tests are narrowly focused, criminal behaviour is "multiply determined" (p. 89). That is, offending occurs through the interaction between several factors, such that change in one risk factor is unlikely to lead to overall change due to the impact of other risk factors, which might have remained stable through treatment. It is likely that the most effective measurement tools vary for different risk factors; however, there remains limited research to date, which guides the selection of assessments for specific risk factors within a certain type of offending. Some examples of specific patterns observed for individual participants in the current research are discussed below.

One participant in the current study demonstrated a disparity within his pattern of change in cooperation with supervision, such that he exhibited a significant decrease in the frequency of both negative and positive behaviours. This participant was in treatment for the longest duration out of all the participants, at 62 weeks. It is possible that over time, his behaviour became less noticeable, for reasons such as: he became more comfortable in the environment; or, he chose to maintain a lower profile in the wing, with the hope that it would accelerate his treatment progress and his release.

Based on the requirements Jones et al. (2010) suggested in the evaluation of change, this participant would not be assessed as having made positive changes over treatment, because although his negative behaviours decreased, so too did his positive behaviours. Jones and colleagues did not provide clarity around the integration of differing change results for the one individual. A decrease in positive behaviours associated with cooperation with supervision might not exclude the possibility of an

overall determination of improvement over time. The various change patterns within and between each measure used must be evaluated and combined. For instance, this participant's Treatment Gain score was below the mean, which may suggest his difficulties with cooperation with supervision had a negative impact on his treatment participation, which his therapist subsequently rated poorly. He did not demonstrate improvement on any of the psychometric measures, the majority of which remained in the dysfunctional range pre- to post-treatment. His absence of change on several of these measures could be interpreted alongside his cooperation with supervision results. That is, his poor motivation might have been a barrier to his skill development, which was reflected through the absence of change within the psychometric measures. Therefore, through consideration of his results from multiple measures, it may be concluded that he did not evidence change through treatment.

Another participant's results provided an interesting comparison between self-report and observed behaviour. In relation to the behavioural checklists, this participant reported an increase in all positive domains but no change in the negative domains. His MC-SDS score was not elevated, which suggests that his increase in positive domains may have been an accurate reflection of his behaviour. However, he may have been more likely to present himself in an overly positive light with respect to negative behaviours, such that he reported low frequencies of these behaviours throughout treatment. This argument might also relate to his psychometric results. He was already in the functional range pre-treatment for several psychometric measures, which might suggest his attempt to conceal problematic attitudes. In relation to observer results, custodial officers did not observe any changes within the negative or positive domains in the behavioural checklists; however, his Treatment Gain score was above the mean and he was rated as Satisfactory on the SBRG by both his therapist and a custodial officer. With the greater level of information included in the behavioural checklist compared with the Treatment Gain scale and the SBRG, in a clinical context it may be useful to further explore this participant's behavioural checklist results to gain greater insight into his level of change over time.

Another participant reported a decrease in positive behaviours in each behavioural checklist domain. The psychometric measure results were examined to determine whether this deterioration within the behavioural checklists was more generalised. He remained in the dysfunctional range pre- to post-treatment for seven out of the 15 psychometric measures but notably, he did not demonstrate deterioration in any

psychometric measures. He also displayed positive change within the behavioural checklists, through a decrease in negative behaviours in coping and problem solving, mental disorder, and cooperation with supervision. However, this improvement was not widely observed in the psychometric measures, with improvement in one coping and problem solving measure and recovery in one relationship and social influences measure. Similar to the previous participant described, this participant's ratings in the Treatment Gain scale and the SBRG reflected positive reports from observers, which was in contrast to the absence of any improvements observed by custodial officers in the behavioural checklists. In a clinical setting, an overall evaluation regarding change in treatment would require the clinician to assess these results in tandem with this offender's primary treatment needs and relevant risk factors.

Mann et al.'s (2010) propensities framework might assist with the interpretation of some results in the current research. For instance, one of the participants deteriorated in the psychometric measures related to coping and problem solving. However, custodial officers observed an improvement in this participant's positive behaviour within this domain. It is possible that this participant's propensity for poor coping was strong (reflected through the psychometric measures), but the relevant behavioural manifestations were suppressed in the custodial environment (reflected through the behavioural checklists). There are two potential explanations for his improved behaviour in this domain. First, the improvement might have been irrelevant to his risk (i.e., the specific risk factors were not relevant criminogenic needs). Alternatively, these behaviours might have been related to the development of detection evasion skills. Detection evasion skills may have masked his problematic attitudes; therefore, his behaviour appeared to improve because his previously observed problematic behaviours were no longer obvious.

These results emphasise the importance of ensuring clinicians do not rely solely on one measure or assessment process when determining offenders' change over time. Although there may be differing levels of consistency between measures, greater depth of information can be gained through multiple measures. Any differences between and within measures should be explored to gain a greater understanding of the individual's change process.

There is also a need to determine how a clinician should incorporate the data gained from an offender and custodial officers into a change assessment. Various pieces of information must be integrated in addition to the results of the measures, such as: the

offender's tendency towards positive impression management (e.g., measured through the MC-SDS); clinicians' observations within the treatment context in relation to the relevant risk areas; and, custodial officers' level of interaction with the offender. Based on the results of the current study, it is recommended that clinicians consider all the data to which they have access and use the available information to initiate further investigation of the changes relevant to each individual offender. It is important for clinicians to recognise that offenders might emphasise their positive behaviour, while custodial officers may not be cognisant of some behavioural manifestations of risk factors (e.g., related to sexual self-regulation).

4.4.5 Case Study Outlining a Clinician's Decisions

As previously indicated, the integration of results from different change measures can be a challenging task, particularly when the results appear to be contradictory. Currently, there is no validated, structured method by which to overcome this challenge. However, it is a necessary task in the clinical field to facilitate decision-making. An example of the way in which a clinician might interpret the results of the various measures for an individual will be provided using one of the current participants as a case study: Mr R.

Mr R's scores on the BRS and BMS were already functional pre-treatment. Despite their possible transparency, these measures have been validated for use with sexual offenders. If Mr R had no treatment needs in this area, he would likely spend minimal time in treatment in relation to changing these attitudes and beliefs. As a result, consistent with his current results, his post-treatment scores would be expected to remain unchanged. His pre-treatment scores on the psychometric measures that assessed coping (e.g., coping using sex and emotion-oriented coping), relationships, and self esteem, were in the dysfunctional range. Therefore, Mr R's core treatment needs would likely have been in these areas if they were causally related to his offending. Further to this assumption, these needs should have received the focus in treatment. They should also be the priority in assessing pre- to post-treatment change. The results revealed that Mr R achieved clinically significant change in these measures.

Mr R's behavioural checklist results reflected a decrease in poor coping and problem solving, and negative behaviours related to mental disorder. This result suggests his treatment work may have been effective in decreasing negative coping behaviour. Sexual deviance is also related to mental disorder; however, there was no similar change in the WFSQ. Mr R's offences were against children; therefore, it is

possible that he focused on developing management strategies for sexual thoughts about children, rather than targeting the WSFQ areas, which are predominantly focused on adults.

Within the behavioural checklists, his self-report reflected high frequencies of positive behaviour from the commencement of treatment. He chose to use the scale provided on the checklists rather than indicating a specific numerical response. Recall that the scale was from zero to 50. His responses for positive behaviours were generally between 40 and 50; therefore, his potential for self-reported improvement was limited if he were to continue using the scale rather than writing a response in the space provided. These frequencies were higher than many other participants' self-reported responses, which might reflect a tendency to exaggerate his behavioural frequency. Based on Mr R's self-report, he was engaging in both positive and negative behaviours from the commencement of treatment. As previously discussed, positive and negative behaviours can co-exist throughout treatment. Over his treatment engagement, he reported using more positive behaviours.

Custodial officers' observations reflected Mr R's improvement in positive behaviours associated with relationships and social influences, another of his demonstrated clinically significant change areas. While in treatment, he appeared not to deliberately develop relationships with other offenders and his contact with personal supports in the community was limited. As a result, his self-reported behaviour in this area likely reflected his attitudes about relationships in the custodial environment. However, his development of skills was observable to custodial officers. Mr R received positive ratings from staff through the Treatment Gain scale and the SBRG, which suggests that his self-report was consistent with behaviours observed within the treatment environment. Integration of these change measures reflects Mr R's improvements in several areas relevant to his criminogenic needs. Using Davies et al.'s (2010) and Jones et al.'s (2010) guidelines within this assessment, it is likely that a clinician would evaluate Mr R as having made positive changes throughout treatment.

4.4.6 Limitations and Challenges

The current study highlighted a number of practical challenges within this research area. These challenges paralleled the clinical process of collecting information about offenders' behavioural manifestations of risk factors in order to modify risk assessments and determine whether changes have been made over a specified period of

time. The challenges comprised both limitations of the research and difficulties arising from the nature of the task.

4.4.6.1 Limitations. The primary limitation was the small sample size, which further diminished upon the withdrawal of participants from the self-report process. The small sample size and quantity of missing data weaken the conclusions that can be drawn from the results of the study. While a larger sample was intended and attempts were made to recruit additional participants, external factors prevented further recruitment (e.g., operational changes within the program leading to lower numbers of offenders commencing treatment). Therefore, data collection was discontinued after two and a half years. Further, change research requires the evaluation of both intermediate and long-term outcomes, to evaluate whether the changes observed in the structured custodial environment are maintained upon release to the community. Within the current study, intermediate outcomes were assessed; however, longer-term follow-up to capture recidivism and other indicators of improvement or deterioration post-release would allow for further evaluation of the changes observed through treatment. Due to the small sample size in the current study, this follow-up would provide less value than it would with a larger sample.

Data collection from custodial officers was reliant upon their availability, in addition to their sustained commitment to the research. There was missing data and further, the reliability of the data was unknown because the degree to which the custodial officers and participants interacted within the CUBIT community varied between custodial officers.

Another limitation of the current study was the absence of a multi-item structured professional judgement measure as a comparison with the included measures. For instance, the VRS:SO has been validated as a change measure (e.g., Olver et al., 2007). However, there was no consistent use of a validated structured professional judgement measure in CUBIT at the time of data collection, against which the results of the other measures could be compared.

4.4.6.2 Practical issues and implications. Several challenges arose through conducting this study. These challenges highlighted the difficulties that can be encountered when completing research in a custodial context. Some of these challenges also highlighted the difficulties clinicians may face in the assessment of offenders as they progress through a treatment program. These challenges will now be discussed.

4.4.6.2.1 Checklist development. In order to effectively identify predicted behavioural manifestations of underlying propensities, as required within both the OPB and OAB frameworks, it is important to have clarity surrounding the overarching concept of dynamic risk factors. However, as described in section 1.1.7 (p. 16) and section 1.1.9 (p. 27), the definition and conceptualisation of risk and protective factors remains the subject of ongoing debate. Prior research has suggested that risk factors are complex constructs, which contain multiple levels (Klepysz et al., 2016; Thornton, 2013; Ward & Beech, 2015). This multi-layer complexity was evident in the behavioural checklists in the current study. As previously indicated, there was overlap between risk factors and behavioural manifestations in the behavioural checklist items, such that some manifestations were relevant for more than one risk factor. This overlap had implications for the checklist completion and the interpretation of the responses.

Although each item on the checklist was related to a particular risk factor, the risk factors were not specified on the checklist. For instance, based on the psychologists' survey response codes, emotional instability was a behavioural manifestation of the risk factor Violent or Suicidal Ideation. However, participants in the current research were unaware of the specific context of the behavioural manifestation. Therefore, it is likely that their responses were related to whether they were emotionally stable over the week, irrespective of violent or suicidal ideation. Consequently, the results may have been distorted if the behaviours for one risk factor were reported as occurring when in fact they were related to a different risk factor (e.g., Problems with Stress or Coping). The impact of this potential problem was reduced through the aggregation of the risk factors into domains, such that risk factors with similar manifestations were likely classified into the same domain. Nevertheless, there are some behaviours that could be placed in a different domain depending on the underlying function. For instance, talking to supports could be a manifestation of a risk factor within relationships and social influences, or within coping and problem solving. As such, a more individualised approach to developing checklists of predicted behavioural manifestations, such that the levels are better defined and the risk factors are clearly operationalised through the manifestations, might allow for the development of a more sensitive measure of change.

Further, it is important to note that participants and custodial staff reported the frequency of each behaviour on the checklist, regardless of the function or the context of the behavioural expression. As a result, each behaviour being monitored was not necessarily a criminogenic need specific to the individual. For instance, although a

participant might have engaged in a behaviour such as withdrawal/isolation from others, it might have been a function of the custodial environment and the participant's response to this environment (e.g., safety concerns from other offenders, or spending time completing treatment task work), rather than a specific expression of a behavioural manifestation of a relevant risk factor (i.e., an OAB of Problems with Stress or Coping).

From both a research and treatment perspective, it would have been beneficial to discuss the processes of behavioural monitoring and change, with participants, in conjunction with their treatment. This process was not possible in the current research. In the CUBIT program, there is discussion about the function of behaviour and its relation to risk. Often, offenders have difficulty identifying the function of their behaviour or they are unaware that their behaviour is problematic (Hoberman & Jackson, 2016), as this process requires insight and consistent practise. Implementation of the behavioural checklist in clinical settings would benefit from discussing the function of behaviour during checklist completion. A greater understanding of the relevance of behaviours to the individuals would further facilitate the data capture and evaluation of relevant change.

4.4.6.2.2 Data collection. The data collection process for the behavioural checklists was a representation of both the complexities involved with research in this context and the clinical reality of the process. For instance, challenges were faced from the initial recruitment of participants. There were decreased numbers of treatment participants commencing in CUBIT during the course of the study due to staff changes and uncontrollable operational factors. As a result, there were fewer offenders available to recruit for the research, which limited the possible sample size.

Ongoing challenges were faced throughout the data collection process. Participants were in treatment for differing lengths of time, depending on various internal and external factors such as: earliest release dates (one had a Court Based Release date, which placed an end point on his treatment that was not necessarily related to his performance); difficulties experienced while in treatment; therapists' clinical judgement; and, therapist changes part-way through treatment. These differing lengths of time in treatment may have had implications for the change processes.

As previously discussed, the behavioural checklist included a large number of items. One consequence of the checklist length was that it became cumbersome for participants to complete, as evidenced through the withdrawal of some participants from the research. It also became more difficult for custodial officers' awareness and

effective monitoring of all the behaviours, which may have had a detrimental impact on the validity of the measure.

4.4.6.2.2.1 *Participants*. Data collection issues specific to participants will now be discussed. Participants' withdrawal from the research prior to treatment completion may have been related to their limited motivation. For instance, some participants who engaged in the research for the duration of their treatment stated that they used the checklists to reflect on the behaviours they engaged in through a more concrete manner than would often be used. However, participants who withdrew from the research made comments such as "I can't be bothered" and identified that they no longer wished to take the time to engage in the process. This limited motivation may have served as a barrier to their consideration of the longer-term impact of their behaviour, such that they ignored or remained unaware of potential benefits they could gain through ongoing participation in the research. Similarly, limited motivation is a practical challenge faced in the clinical context, which has implications for the level of change that offenders will attain through treatment.

The withdrawal of participants led to missing self-report data. If there were extraneous factors that led certain participants to withdraw from the research compared with others, it might have skewed the results. For instance, lower motivation might be due to poor insight, poor coping and problem solving skills, and more antisocial attitudes (e.g., hostility towards individuals in authority, negativity surrounding assisting others with no perceived personal gain). The remaining participants might have displayed similar characteristics to each other, such as generally engaging in more positive behaviour or alternatively, the desire for staff members to view them as motivated and engaged. The small sample size prevented the use of statistical analyses to measure these differences between participants who completed all measures and those who did not.

Additionally, for the data that was collected, there was a difference in the methods by which participants reported information (i.e., actual frequency versus percentage of the time, or an approximate representation). Although all participants were provided with instructions and their understanding was confirmed, some consistently engaged in an alternative process for completing the checklists, as evidenced through their comments and the responses they provided. This difference in reporting methods reflects the difficulty with which accurate data are collected in research and the impact of individuals' interpretation of instructions on the results. Additionally, participants'

differing methods of responding may have been related to difficulties with memory of, or insight into, specific examples of behaviours that were identified in the checklists.

Parallels can be drawn between participants' differing interpretations of instructions within the research, and similar challenges in a clinical context. While in treatment, offenders are provided with tasks and their responses have implications for their progress through treatment, and for the analysis of change and risk of reoffending. It highlights the potential value in using multiple methods of assessment, both in the clinical context and in change research.

4.4.6.2.2.2 *Staff observations*. Previous research has provided evidence for the potential benefits to using staff observations of offenders' behaviour in order to enrich the risk assessment and change measurement process (e.g., Atkinson & Mann, 2012; Gordon & Wong, 2010; McDougall et al., 2013; Pearson & McDougall, 2017). It provides an alternative to offender self-report and uses an existing resource; custodial officers, who observe offenders' behaviour on a daily basis. Pearson and McDougall (2017) noted that offenders' daily life in custody as monitored by custodial officers, might be 'opportune' for the observation of behaviours with fewer attempts by offenders to evade detection. Some issues arose with respect to this process in the current research, which highlight the limitations of the current form of research.

Within the current study, the inconsistencies in custodial staff within CUBIT meant that at times, there were no regular custodial officers rostered on for the week. Consequently, no observer checklists could be completed for these weeks. In addition, if a regular officer was subsequently working in CUBIT the following week, he/she would not have sufficient information to provide accurate data regarding the participants' behaviours during the preceding week. As a result, the absence of regular officers in a given week often resulted in two weeks' missing data. Non-regular custodial officers were unfamiliar with the participants and often were only in CUBIT for one or two days; therefore, they were unable to provide accurate data. There were efforts made to overcome these challenges, such as seeking regular custodial officers to complete the checklists a day or two after the scheduled day.

A further challenge was that throughout the data collection period, there were differences between the ways in which custodial officers completed the behavioural checklists. For instance, one officer tended to record "0" for the majority, or all, of the behaviours on the checklist. This reporting pattern might indicate the officer was not

engaged with participants and could not provide accurate observations; alternatively, it might indicate the officer did not wish to place effort into the checklist completion.

In order to be effective, the process of behavioural measurement relies on custodial officers' regular engagement with offenders to facilitate behavioural observation. Within the CUBIT environment, there are differences between custodial officers in their level of engagement with offenders. Some custodial officers spend more time interacting with offenders, while others spend the majority of their time in the wing office and therefore have less opportunity to observe and monitor offenders' behaviour. This difference relates to both their attitudes towards offenders and towards their role within the environment.

These differing levels of engagement with offenders might provide an explanation for the differences between custodial officers' observations in the behavioural checklists. Often, custodial officers completing the checklists had limited awareness of behaviours exhibited, unless it was directly related to behaviour they consistently monitored as part of their primary role (i.e., related to maintaining safety and security in the gaol). This may be related to Atkinson and Mann's (2012) findings that one reason custodial officers did not report problematic behaviour was that they did not consider observation and monitoring of potentially risky behaviour as part of their role.

Similar issues have been raised in previous research (Clarke et al., 1993; Mooney & Daffern, 2013), which has highlighted the practical challenges associated with the implementation of behaviour monitoring in custody. That is, the behaviour in the measure may not be consistent with the behaviour custodial staff seek to monitor for other purposes (e.g., adjudicating misconduct). Therefore, it may be difficult to record behaviour in a consistent and objective manner, and the individualised assessment and monitoring of behaviour may be time intensive. An indication that some custodial officers in the current study were less engaged in the accurate completion of the checklists was that some reported all participants to be compliant with their medication despite some participants not being prescribed medication. This suggests they did not consider the individual offender for every item, which may have been an attempt to simplify the process and decrease the time spent completing the checklists. It may be related to Harris and Hanson's (2010) suggestion that once observers' motivation is low, they are more prone to biased and careless assessments.

One clinical implication is that there may be problems with the validity of custodial officers' responses, which impact the conclusions that can be drawn regarding

treatment-related change in incarcerated offenders. Based on the candidate's clinical experience, some custodial officers' attitudes about sexual offending (e.g., feeling disgusted) also hinder their efforts in gaining information about risk-relevant behaviour. Further, if custodial officers do not spend time interacting with offenders on a meaningful level, one key component of the therapeutic community process is compromised. An important process within the therapeutic community is the sharing of information between staff from different disciplines; however, the quality of the feedback that custodial officers can provide to therapeutic staff will deteriorate if they do not spend time observing and engaging with offenders. It may mean that training is required, in relation to the ways in which custodial officers can engage more fully in the therapeutic community so that they can more effectively observe relevant behaviours. On an operational level, where possible, decisions about which custodial officers are assigned to work within therapeutic programs could be made with greater consideration for the characteristics of a therapeutic community. For instance, it would be of greater benefit in this environment to consistently allocate officers who are motivated to engage with offenders such that they are more likely to influence and measure offenders' behavioural change. If custodial officers are involved more effectively within the therapeutic environment, there may be greater opportunities for important information to be gained about offenders' behaviour and ongoing change processes.

As indicated in the limitations of the study, data collection was reliant on custodial officers' ongoing engagement with the research. However, it was challenging to maintain their interest in the completion of the checklists. Some officers provided reasons for their minimal engagement, such as: the checklist was too long; the checklist was repetitive; they did not want to spend the time completing the checklist every week; and, they could not provide accurate responses for a large proportion of items on the checklist because they did not observe these behaviours.

Staff motivation has been discussed in prior research. For instance, Harris and Hanson (2010) reported their research findings that some community supervising officers demonstrated reduced motivation in the assessment process. They noted that these officers' predictive accuracy was only marginally better than chance. They speculated that the reduced motivation in non-mental health professionals might be due to two reasons: (1) a perception of the assessment as an additional duty, which distracted them from their primary obligations; or, (2) a lack of psychological knowledge leading to their perception of the assessed psychological characteristics as

less complex than they are, thus placing low importance on the assessment. Miller and Maloney (2013) conducted a survey of community corrections staff to measure compliance with risk/needs assessment tools. They reported that some respondents showed relatively low levels of compliance, such as careless tool completion, placing minimal effort in the process, and minimising relevant characteristics. These observations and findings from previous research are relevant for the current research. It is possible that characteristics relevant for community corrections officers are similarly relevant for custodial officers. For instance, discussion about the current research with some custodial officers in CUBIT elicited comments minimising the importance of the behavioural observation process.

Daffern et al. (2007) referred to the implications of observer bias, subjective observation, and accidental observation in the risk assessment process. These factors are important in the measurement of behavioural change. The OPB and OAB frameworks can serve to reduce the impact of these issues through prospective monitoring. Within the current research, the use of a checklist to guide custodial officers sought to target these potential problems. Although it was deemed beneficial to provide custodial officers with a framework such that their biases are decreased and their observations are targeted, it does not necessarily prevent these problems from occurring. This is partially due to the limitations of the context in which treatment was conducted, such as the factors that decrease its effectiveness as a true therapeutic community discussed previously in this section. These factors lessen the ability for the framework to be implemented as a systematic process for behavioural change measurement.

4.4.6.2.2.3 *Multiple change measures.* Use of a variety of change measures may assist assessors to overcome some of the challenges described. In the current research, participants tended to make changes in some areas but not in others. This inconsistency in the change process becomes a clinical dilemma because clinicians must determine which measures to emphasise in an overall assessment.

It should be noted that there were no tests of measurement invariance conducted within the current study. Therefore, it cannot be statistically determined whether the four assessment types used to measure change over time were measuring the same or different constructs at each time-point. However, use of multiple assessment measures has advantages over using fewer assessments. That is, if an assessment has not been validated (e.g., the behavioural checklist), it cannot be definitively concluded that the assessment consistently measures the same constructs in the same way over time in

treatment. Therefore, use of additional validated assessments may serve to minimise the potential impact of measurement error. The outcomes gained from multiple assessment tools can be compared with each other in order to evaluate the validity of the results.

One benefit of implementing several change measures is that each measure has strengths and weaknesses. One limitation of psychometric measures is the potential for biased results due to the self-report process. Self-report can be problematic in the custodial context, particularly during treatment, because many offenders try to create an unrealistically positive impression within treatment, in order to attain earlier release from custody. A strength of psychometric measures is that they can target attitudes and behaviours, which are less observable to assessors and have the potential for concealment by offenders, for instance, those related to sexual deviance (see Hanson & Harris, 2001). Further, there is evidence to suggest these measures are predictive of recidivism, which indicates they can reflect change in important aspects of offenders' psychological profiles (see Barnett et al., 2012).

Response bias may also be relevant for the behavioural checklists. Mistrust of others generally, and specifically towards those in authority, tends to be prevalent within the offender population, due to experiences over their lifespan and within the custodial context, as evidenced through risk factors empirically related to sexual offending such as intimacy deficits and hostility. The behavioural checklists might be susceptible to offenders' mistrust, such that they report behavioural frequencies that they consider to be the least detrimental to their engagement in treatment. In the current research, although the Participant Information Sheet and consent form clearly stated that the information provided within the research would remain confidential, one concern for participants may have been that the information would be used by the candidate, their treating therapist, or custodial officers, for malicious purposes.

Although there are challenges with the routine completion of these checklists, a strength of the behavioural checklists is that they provide a method by which to gain observers' views in addition to self-report. They also provide a richer and more detailed understanding of the behaviours related to certain risk factors and domains of psychological functioning. As discussed in prior research, the reported presence of a risk factor often depends on how well behaviour is monitored by staff and how it is interpreted. The use of checklists, which indicate the most usual types of behaviour in custody associated with each risk factor, might solve this problem for the most commonly occurring risk factors, through provision of a more consistent approach

(McDougall et al., 1995). Further, as Lewis et al. (2013) indicated, assessment of dynamic risk factors at multiple time points might be the most accurate method by which to assess the dynamic nature of the variables. Notably, Douglas and Skeem (2005) suggested that the longer the interval between assessments, the greater the possibility that changes will be missed. Weekly reporting reduces this possibility.

The results of the current research suggested that the SBRG might not be a sensitive measure of change through treatment, as highlighted through its poor discrimination between participants. Only one participant gained a rating of Unsatisfactory, as reported by custodial officers, while no participants were rated as Unsatisfactory by therapists, despite the therapists reporting lower Treatment Gain scores. Discussion with custodial officers suggested that the participant who gained an Unsatisfactory rating on the SBRG may have received this rating in part due to his personality (e.g., being ‘annoying’), rather than specifically engaging in risk-related behaviour.

Beech et al. (2016) indicated that ratings of change may be affected by various factors such as: the relationship the assessor has with the individual in question; the importance the assessor places on various aspects of change; and, their personal opinion of the individual offender. Although the overall rating on the SBRG is related to individual items referencing behavioural areas, it is possible that this measure is prone to biases that lead to subjective ratings rather than specific risk-related behaviours. The SBRG may not be a useful measure of change and at this stage, its use for this purpose is not recommended. Further research is required to understand the reasons for custodial officers’ and therapists’ tendency to rate offenders’ behaviour as Satisfactory despite evidence of problematic behaviour (e.g., institutional misconduct). The ultimate rating provided within the SBRG is likely insensitive and perhaps should not be used. Rather, the Treatment Gain scale may be an alternative measure of satisfactory behaviour within this context, as this measure was more discriminative between participants within the current research.

Similar to the SBRG, the Treatment Gain scale may elicit subjective ratings from clinicians. Related to this point, Davies et al. (2010) asserted that clinicians might produce biased assessments due to their desire for treatment change to have occurred. However, this bias was not apparent through the Treatment Gain scale results, with some participants receiving low scores (e.g., 4 out of 24). While the Treatment Gain scale is seemingly subjective, the results in the current research suggested some

correspondence between this measure and results obtained from other change measures; for instance, the correlation between Treatment Gain scale scores and participants having recovered on psychometric measures. The Treatment Gain scale has previously been validated and can be beneficial for clinicians' use (Sowden & Olver, 2017). Taking into consideration the potential limitations of this measure, it appears useful within the clinical context. Specifically, it is a straightforward measure, which can be used in conjunction with other change measures to further enhance the information available to clinicians regarding offenders' change processes.

In reality, the potential for assessor bias such as that described in relation to the SBRG and Treatment Gain scale, is possible in all assessments. For instance, the structure of the behavioural checklists appears objective as the ratings are based on behavioural frequency rather than a judgement. However, a custodial officer's subjective view of a participant might influence their tendency to focus on certain types of behaviour over other behaviours, such that they are more attuned to positive or negative behaviours from some offenders compared with other offenders. Similarly, it might affect their perception of the behaviour as overly positive or negative and, therefore, the likelihood that they will report it.

These potential biases highlight the value in using multiple measures of change, which are completed by different assessors, to determine consistency and gain greater understanding of the patterns observed. Further, using a combination of several measures allows for the development of a richer depiction of risk and change and the limitations of each approach can be counterbalanced by using another measure. However, as previously noted, it also increases the complexity of the measurement process, as the disparate measures must be synthesised and understood as a cohesive representation of change.

4.4.7 Recommendations for Measuring Change

Overall, considering the advantages and limitations of each of the change measures implemented in the current study, the behavioural checklist is recommended as a potentially useful measure of change. Prior research has identified the use of checklists as a behavioural monitoring method to inform risk assessment prior to the release of high risk offenders (McDougall et al., 2013). The behavioural checklist developed within the current research may provide a greater depth of information about offenders' change over time compared with previously validated measures such as the VRS:SO, due to its weekly completion.

If a behavioural checklist is implemented as a measure of change, clinicians could discuss with offenders their self-reported behaviours on a regular basis. This discussion may facilitate identification of the factors underlying certain behaviours and ongoing modification of case formulations. It would also encourage offenders to continue evaluating their behaviour and might serve as motivation to address problem behaviours when they arise, and for positive behaviours to be acknowledged and reinforced. Since the VRS:SO has previously been validated, it may be beneficial to link the behavioural checklist to the scoring of the VRS:SO. The behavioural checklist may enhance the change assessments conducted through use of the VRS:SO by providing specific behaviours that are associated with each dynamic risk factor.

A challenge associated with implementing the behavioural checklist is the translation of raw data into meaningful information for clinical practice, without the use of statistical analyses. One possible method for clinicians may be similar to the preliminary stages of the current data analyses. The reported frequencies could be reduced to quartiles in order to calculate change scores. The utility of this process may depend on the number of items in the checklist, as it could be time consuming. Clinicians may not have the time to commit to this process. Further research could be conducted to assess the validity, utility and practicality of using this kind of measure within the context of a custody-based treatment program.

The pre- and post-treatment psychometric measures used in this study have been validated. However, as previously discussed, some measures might be less useful due to their transparency (e.g., BRS and BMS; see Grady et al., 2011) and might not provide useful change information. Further, the relevance of each psychometric measure should be evaluated within the context of the specific treatment program. If individual measures are relevant within the treatment program, such that they target the same criminogenic needs, they may be useful adjuncts to the behavioural checklists.

As previously noted, the Treatment Gain scale might be a valuable overall measure of change, through its greater ability to discriminate between offenders, as compared with the SBRG. It could be used as a global assessment in the initial evaluation of change. Subsequently, clinicians could further investigate offenders' specific changes in behaviours, attitudes and beliefs, through the behavioural checklists and psychometric measures, in addition to a structured professional judgement measure. Each offender's criminogenic needs should be the focus of the assessment. It is recommended that clinicians explore any discrepancies between change measures,

based on the individual offender's criminogenic needs. Further, if there are discrepancies between self-report and observer information, the potential underlying reasons for these differences should be assessed (e.g., offenders' attempts to present well to assessors, or observers' limited opportunity to witness behaviour).

4.4.8 Future Research

It is important to view the current study as a progression towards ongoing research to advance knowledge in dynamic risk and change assessment. Therefore, some ideas for future research will now be outlined.

In future research, if a similar process of change assessment was used with a larger number of participants, perhaps the impact of additional factors could be investigated (e.g., whether a participant commences use of anti-libidinal medication while in treatment). In addition, further comparisons could be made between participants, such as the characteristics of those who remain in the research as distinct from those who withdraw consent.

Grady et al. (2011) asserted that given the potential research and clinical knowledge to be gained from collecting data from valid and reliable instruments, more research is required to develop instruments that can be used with sexual offenders. They noted the need for gaining greater understanding of the deficits that have the greatest impact on the recidivism rates of treated offenders. This highlights the value of conducting further research into the behavioural checklists and the importance of follow-up studies. In addition, if the VRS:SO is validated in Australia in the future, this measure could be used as an adjunct to the behavioural checklist, rather than relying on the psychometric test battery used currently.

As Cording et al. (2016) noted, most studies related to the reliability and validity of dynamic risk assessment measures have been conducted with researchers or developers using the tools rather than by professionals in a correctional context. This process raises questions about whether these measures will be similarly effective when they are no longer scored by trained researchers, but instead by staff who may have many other responsibilities and priorities. This idea is similar for change measurement tools. Therefore, the current research provided additional information to the field in relation to the measurement of change in a custodial setting, using custodial officers and therapists rather than researchers, in their application. However, ethics approval constraints prevented the use of therapists' observations in the behavioural checklists. In light of the difficulties custodial officers had observing some behaviours in the

checklists, in the future it may be more effective if therapists also contribute to their completion. This idea is consistent with Gordon and Wong's (2010) suggestion that behavioural observations should be made by all staff members and checklists should be amended throughout treatment. In future, subject to ethics approval, behavioural checklists could be developed in consultation with therapists and modified through treatment. Further, if future research involves several different staff members completing checklists, richer data could be gained. Longer-term follow-up data (e.g., recidivism) could also reveal the differential validity of staff members' responses.

It may be useful to modify the behavioural checklists before implementation as a change measure in custodial environments. For instance, idiographic behaviours could be included for each offender, based on a clinician's case formulation. Further, the checklist could be condensed; for instance, some of the behavioural manifestations of risk factors could be amalgamated, such as those that elicited similar results as each other in the current research, in addition to reducing repetition of behaviours. As a result, it may be more accessible to both staff and offenders. Perhaps the frequency of the checklist completion could also be reduced, although it is important to consider the benefits associated with more frequent assessments, as previously discussed. Prior to its implementation in a modified format, it could be tested through a pilot study. Further, after a pilot study is conducted, a qualitative review could be completed with offenders and custodial officers, in order to gain information that could further improve the checklist.

Gordon and Wong (2013) asserted that while offenders may make some important changes in treatment, these changes may or may not be risk-related or even represent a restoration of healthy or adaptive sexual, interpersonal or social functioning in the community. The content of these changes must be explored and their relation to ongoing change processes must be investigated. Specifically, according to Serin et al. (2013), although measuring treatment targets through intervention is important in order to assess treatment gain, measurement beyond the treatment period is also required to understand change. This requirement is based on the idea that behaviour change is a long-term process, which evolves in a gradual manner. Change is only meaningful if it is maintained post-release. Hence, future research could advance the current study through including follow-up outcome measures to determine ongoing change processes. As previously indicated, the use of recidivism outcomes would be one method by which

to indicate whether changes made on checklists, psychological measures and so forth are actually related to reduced recidivism.

There are limitations to the use of recidivism data as indicators of behavioural change in custody. For instance, it has been suggested that recidivism is not a very sensitive indicator of treatment effect (Schmucker & Lösel, 2015). One contributing factor to its insensitivity is victims' under-reporting, particularly in relation to sexual offending (see Chapter 1, pp. 3-4). However, recidivism outcomes are widely accepted methods by which to evaluate offenders' longer-term behavioural change. While intermediate outcomes are important within the evaluation of offenders' behavioural change, these changes must be maintained over time. Therefore, it is worth highlighting the lack of recidivism data as a limitation within the current study. In responding to this limitation, it is also useful to discuss the ways in which future research could resolve the criticism.

Several criteria should be met in order to conduct robust recidivism research (Schmucker & Lösel, 2015). Importantly, the recidivism study would need a treatment group and a control group not receiving the treatment program. These researchers noted that studies with up to 50 participants included a 'small' sample size. There are limitations to the analyses that can be conducted and the conclusions drawn in relation to recidivism in studies with small sample sizes. Further, there have been questions about the robustness of recidivism research in which the researcher had an affiliation with the treatment program. Schmucker and Lösel (2015) suggested the need for recidivism research that is independently authored.

Two prominent meta-analyses can be used as guidelines for determining the appropriate sample sizes and follow-up times required in order to conduct recidivism studies. Hanson and Morton-Bourgon (2009) conducted a meta-analysis that included 110 studies. The average sample size for these studies was 417 participants. The mean follow-up time post-release was 70 months (5.83 years). Schmucker and Lösel (2015) used 27 studies in their meta-analysis. The average sample size for these studies was 358 participants. The mean follow-up time was 70.26 months (5.9 years). Within the current research, it took two and a half years to recruit 19 participants, several of whom provided only partial data due to their withdrawal from the research. In order to recruit even a fraction of the sample size required for a robust recidivism study (for example, only 100 participants), in addition to the follow-up time required post-release, in order to establish whether change indices are associated with recidivism, the research would

take an additional 18 years to complete. While it is clearly beyond the scope of a PhD to conduct such a recidivism study, it is an important component of research in offenders' behavioural change. Therefore, the field would benefit from this research in the future, if there are researchers who have a sufficient length of time available to them.

4.4.9 Conclusions

This study focused on four measures of change, which can be implemented within a custody-based treatment program. These change measures comprised a variety of modes, including offender self-report, custodial officer observations, and therapist ratings. Results revealed that the behavioural manifestations of sexual offenders' dynamic risk factors can be observed in custody. Change over time in treatment within some behaviours was also evident. However, change as reported by custodial officers was not always consistent with participant self-report. Moreover, there was limited consistency between the changes observed through the behavioural checklists, and the attitudes and beliefs that were represented through the pre- to post-treatment psychometric measures. The challenges that were faced throughout the data collection period have been discussed, with some implications for these challenges further highlighted in the context of clinical practice. Both group- and individual-level analyses were conducted in order to gain greater depth of information regarding offenders' change processes and the ways in which the information from different measures can be integrated. The challenges associated with this integration were discussed, in relation to assessors determining whether or not an offender has made changes.

Through consideration of the measures used in the current research, it can be concluded that there are limitations to all change measurement approaches. However, the use of observations at various time points within structured professional judgement tools may have particular advantages. Within these tools, the relevance and fluctuation of various dynamic risk factors can be captured in an idiographic manner. Scores can be obtained and a change score can be gained over time in treatment, which can then be used to determine changes in risk level. These changes in risk level can be used in decision-making processes related to progression through the criminal justice system, including release into the community.

Chapter 5: General Discussion

5.1.1 Overview of Research

Debate continues in the field of forensic psychology in relation to the definition of dynamic risk factors, their behavioural manifestations in a custodial context, and the most effective methods by which to measure changes in dynamic risk factors over time. These issues are central to the decision-making process authorities face in managing offenders in custody and determining when and if to grant parole. That is, measurement of dynamic risk factors can provide information about offenders' risk, including both the persistence of risk of recidivism, and the imminence of this risk. Risk assessment and change measurement have continued to evolve over the years; however, there remain questions regarding the most effective methods. These processes are particularly challenging within the custodial context, as this environment is not necessarily representative of the environment in which offenders reside in the community.

The current research consisted of three interlinked studies, which together served to increase understanding of sexual (Studies 1-3) and non-sexually violent (Study 2) offenders' behavioural manifestations of risk factors in custody, the observation and monitoring of these behaviours, and whether these behaviours changed over time. An overview of each study will be provided. Subsequent to this overview, discussion will focus on comparisons between the studies' results, prior to an integration of these results. Clinical and policy implications form an important component of this research and will also be explored. An overview of the challenges and limitations of the three studies will precede suggestions for future research as a consequence of these studies.

5.1.1.1 Study 1. Study 1 was a preliminary study to inform a significant aspect of Study 3. The aim of Study 1 was to investigate how dynamic risk factors for sexual offending and prosocial equivalent behaviours manifest in a custodial context, through a survey of psychologists with experience in this area. Participants were asked to provide examples of behavioural manifestations of risk factors and their prosocial equivalent behaviours, in which sexual offenders might be predicted to engage within a custodial environment. Each item in the survey referred to a risk factor from the RSVP and VRS:SO. The results revealed that psychologists could identify behavioural manifestations of relevant risk factors and related prosocial behaviours. Manifestations of some risk factors appeared easier to identify than others; for instance, Problems with Stress or Coping, Interpersonal Aggression, and Substance Abuse elicited more specific and observable behavioural manifestations than did Sexual Deviance. Additionally,

overall, risk-related behaviour appeared easier to identify than prosocial equivalent behaviour. This result was likely due to the tendency for the focus in custody to be on problematic behaviours, rather than positive behaviour (Mooney & Daffern, 2013).

5.1.1.2 Study 2. The behaviours identified in Study 1 included negative and positive behavioural manifestations of risk factors. The negative behaviours encompassed a range of behaviours that included more obvious defiant and antisocial acts including breaches of rules (e.g., avoiding domestic duties), while other behaviours were not necessarily direct rule breaches but were risk-related (e.g., withdrawal or isolation). Rule breaches in custody have been identified as a means by which to evaluate an offender's behaviour (French & Gendreau, 2006), which can subsequently be used to determine readiness for parole. Study 2 was an evaluation of the SBRG. The SBRG was developed for Corrections Victoria (Daffern et al., 2014), to assist with parole decision-making, based on sexual and non-sexually violent offenders' risk-related behaviour within the institutional context. Its use in the current research was related to case workers' (custodial officers') determination of sexual and non-sexually violent offenders' behaviour as Satisfactory, Of Minor Concern, or Of Major Concern.

Overall, the results in Study 2 demonstrated that the final behaviour rating that caseworkers provided was associated with official records of misconduct within the SBRG's four behaviour categories. Increases in both Violent Incidents and Direction/Supervision Incidents were more likely to elicit a more severe overall rating, than were increases in Substance Use/Abuse Incidents or Times Banned from Work. The results demonstrated that at the group-level, ratings of satisfactory behaviour were reliably associated with behaviour in custody. The majority of ratings offenders received were Satisfactory. The primary implication is that officers' ratings of offenders' behaviour on a global scale was consistent with offenders' behaviour more generally, as indicated by official adjudications and work records. Therefore, it appears that custodial officers can provide relevant information about offenders' behaviour.

At the individual-level, there were some inconsistencies observed between offenders' overall ratings and the frequency of recorded incidents. That is, some offenders received a Satisfactory rating despite incurring a large number of institutional misconducts; on the other hand, offenders who received only minimal institutional misconducts gained an Of Minor Concern or Of Major Concern overall rating. Subsequent to the validation of the SBRG, it was included in Study 3. Overall, the

results supported prior research indicating that custodial officers' observations can be useful in behavioural assessment (McDougall et al., 2013)

5.1.1.3 Study 3. Study 3 consisted of an investigation of four measures of sexual offenders' attitudinal and behavioural change through a custody-based treatment program. The primary measure of change was the behavioural checklist developed from the results of Study 1. This behavioural checklist facilitated behavioural monitoring through its weekly completion by both offenders and custodial officers.

The three additional measures were: pre- to post-treatment psychometric measures; the Treatment Gain: Short Scale; and, the SBRG that was validated in Study 2. Together, these four measures represented a variety of methods by which to monitor and evaluate offenders' behaviour in custody, including self-report, custodial officers' observations and ratings, and therapists' ratings. Overall, the results supported prior research that has suggested the benefits of including various assessment methods in the determination of offenders' change over time (Davies et al., 2010).

5.1.1.3.1 *Advantages and limitations of measures.* As described in Chapter 4, each change measurement approach has advantages and limitations. Offender self-report allows for the gathering of information that may not be observable to others, such as that related to thought processes (e.g., sexual thoughts). A limitation of this approach is that it is susceptible to offenders' desire to present themselves in a positive light to assessors. However, as previous research has concluded, offenders who engage in positive impression management may in fact be associated with decreased risk of recidivism (Mills & Kroner, 2005). In Study 3, this limitation is relevant for both the behavioural checklist and the pre- to post-treatment psychometric measures. However, one advantage of the behavioural checklist used in this study was that custodial officers also completed it; therefore, these observations could be compared with participant self-report.

In addition, prior research has criticised methods that rely on a single time point in the evaluation of change, as these methods essentially represent a measure of static risk rather than dynamic risk (Serin, Chadwick, et al., 2016). Further, it has been argued that an insufficient number of time points prevents the adequate measure of change (Douglas & Skeem, 2005). In response to such criticism, Study 3 sought to include the collection of data at multiple time points over the course of treatment. The behavioural checklist facilitated weekly data collection throughout treatment, while the psychometric measures facilitated pre- to post-treatment change through use of these

two time points. On the other hand, the Treatment Gain scale and the SBRG each allowed for an evaluation only at one time point. Advantages of using the Treatment Gain scale and the SBRG included access to evaluations by both therapists (Treatment Gain scale and SBRG) and custodial officers (SBRG), who may have different perspectives in relation to offenders' attitudes and behaviour, and have opportunities to observe offenders in different contexts.

Prior research has identified the potential for observer bias in completing risk and change assessments (Beech et al., 2016; Davies et al., 2010). Through using several methods of change measurement in Study 3, the practical implications for the limitations of each method could be minimised. The primary purpose of using a variety of methods in Study 3 was to gain information from a variety of sources and assist in the identification of the most effective approach, such that it could inform future research into the modification and development of change measures. This research may be valuable for the implementation of change measures in clinical practice.

5.1.1.3.2 Overview of results. The results of the four measures did not consistently correspond with each other. Overall, the behavioural checklist data provided the most complex information, for reasons such as: it included both participant self-report and custodial officer observations; it covered a wide range of dynamic risk factors; and, it facilitated data collection at weekly intervals rather than only one or two time points.

Changes were assessed at both the group-level and individual-level. Group-level results varied across the measures. The behavioural checklists provided the most detail with respect to offenders' change over time in treatment. The results revealed changes over time in treatment in a variety of behavioural manifestations of risk factors, both high- and low-level problematic and prosocial behaviours; however, it may be difficult for clinicians to evaluate overall change using this measure in its current form. As indicated in section 4.4.8 (p. 272), it may be beneficial to revise the checklist and conduct further research into its validity. For instance, in the future, the relevance of these changes for longer term outcomes (e.g., recidivism) could be evaluated in order to determine whether the included behavioural manifestations were meaningful; this process would assist in determining the measure's validity. It may also be beneficial to rationalise items such that a more parsimonious review of behaviours can be undertaken by offenders and officers. While there is inevitably a trade-off between the level of

detail and number of questions included, a more concise checklist may be more palatable when it comes to routine use in these kinds of settings.

Individual-level results provided greater depth of information compared with group-level results, which highlighted the importance of evaluation at this level to assist with determining whether changes have been made, which domains have changed and which appear resistant to change. It is important to gain information about individual offenders' changes across various measures, to facilitate clinicians' ongoing modification of risk assessments. Decision-making occurs on a daily basis in custody in relation to individual offenders, thereby emphasising the need for evaluations at the individual-level rather than only at the group-level. Due to the differences within the individual-level results, it appears there is value in gaining information about various aspects of an offender's change over time, such that the information can be integrated to form the most valid conclusions possible at the time. The results also highlighted the need for further research into methods by which varied results from different change measures can be integrated.

Overall, the behavioural checklist results revealed changes throughout treatment both between and within participants, in the areas of coping and problem solving, mental disorder, relationships and social influences, and cooperation with supervision. There were similarities and differences between participants' and custodial officers' reports of participants' behavioural change patterns using this measure. The pre- to post-treatment psychometric measures revealed some group-level changes. However, at the individual-level, using reliable and clinically significant change indices, the most common outcome was for participants to remain unchanged, either because they commenced treatment in the functional range or because they remained in the dysfunctional range post-treatment. Higher Treatment Gain scale scores were associated with participants having recovered (i.e., moved from the dysfunctional to the functional range pre- to post-treatment) on a larger number of psychometric measures. Therapists rated all participants as Satisfactory on the SBRG, while custodial officers provided this rating for all but one participant.

5.1.2 Comparison Between Study Results

The three studies are linked in various ways. They relate to the identification, observation and monitoring of dynamic risk factors' behavioural manifestations in a custodial context. Both negative and positive behavioural manifestations were included in the research and were measured through various means, as previously described.

Each of the three studies' key results were discussed in the overview in this chapter. Next, the results from each of the studies will be compared and contrasted, prior to their integration. First, a comparison will be presented between the ways in which behaviours were differentially identified and observed in the three studies. Further, comparisons will be drawn in relation to the observation of negative and positive behaviours in Study 1 and Study 3, and between Study 2 and Study 3. Subsequently, the results of the SBRG will be compared between Study 2 and Study 3, to facilitate discussion about its use as an assessment of satisfactory behaviour and behavioural change.

5.1.2.1 Differences between risk factors. Results from Study 1 and Study 2 can be compared in relation to the differential ease with which behavioural manifestations of risk factors can be observed. Within Study 1, psychologists appeared to have greater ease identifying potential behavioural manifestations of some risk factors compared with other risk factors, evidenced through the quantity of examples provided and the relevance of these behaviours to the risk factors. As might be expected, the risk factors that elicited a greater number of relevant behaviours were those that may generally be more observable within a custodial context. They may also be more frequently discussed among multi-disciplinary staff members, which would increase their prominence such that they could be generated more easily within the survey responses. For instance, behaviours associated with Interpersonal Aggression were easier for psychologists to identify compared with behaviours associated with Sexual Deviance. This result can be compared with results obtained in Study 2. Within Study 2, custodial officers appeared to focus on Violent Incidents in their determination of offenders' overall behaviour rating. This result may be partially explained through a similar interpretation. Verbally and physically aggressive behaviours are often the focus of discussion in custody, because they are perceived as more problematic than many other negative behaviours due to their potential impact on others. Further, they are more easily detected within the environment.

These differences between the ease with which behavioural manifestations were observed and monitored across risk factors can also be related to Study 3. To some extent, this comparison is more difficult in relation to Study 3, as a function of the aggregation of risk factors into domains to facilitate data analyses. However, the results revealed that custodial officers were more likely to observe behaviours within some risk factors compared with others, which similarly may be related to their ease of

observation. For instance, custodial officers were less likely to observe behavioural manifestations of Sexual Deviance and more likely to observe behavioural manifestations of Problems with Stress or Coping. Within the domains, it appeared that the risk factors related to coping and problem solving were more likely to elicit behavioural observations compared with those related to mental disorder. Of note, mental disorder included Problems with Substance Abuse. Substance Use/Abuse Incidents were also less likely to elicit a more severe behaviour rating within Study 2. Taken together, these results related to substance use may suggest that custodial officers place less emphasis on this behaviour.

5.1.2.2 Observation of negative and positive behaviours. Risk assessment and change measurement include the observation and monitoring of both negative and positive behaviours. Historically, risk assessment and change measurement generally focused on negative behaviours and ignored the development of positive behaviours as an indication of change and decreased risk of recidivism (Serin, Chadwick, et al., 2016). However, consistent with research highlighting the benefits of identifying and measuring protective factors (de Vries Robbé et al., 2011), in order for a clinician to evaluate an offender as having made positive changes indicative of decreased risk of recidivism, there must be evidence of a combination of a decrease in negative behaviour and an increase in positive behaviour (Gordon & Wong, 2015; Mooney & Daffern, 2013). Therefore, both these types of behaviour must be monitored.

5.1.2.2.1 Study 1 and Study 3. In Study 1, psychologists appeared to have greater difficulty identifying positive behavioural manifestations of risk factors associated with sexual recidivism, compared with negative behavioural manifestations of these same risk factors. This result may have been a product of the focus in this field tending to be on negative behaviour, as the nature of the environment is primarily punitive. However, in Study 3, custodial officers reported that participants' positive behaviour changed more consistently than negative behaviours. Custodial officers' observations suggested that they were attuned to participants' positive behaviour such that they could identify changes in the frequency of these behaviours over time. That is, while Study 1 reflected psychologists' apparent difficulty identifying positive behaviours, which may have been due to the negative focus within the custodial environment; Study 3 reflected custodial officers' ability to observe and monitor positive behaviours despite this negative focus.

It is possible that these results in Study 3 might have been gained because custodial officers were prompted by the behavioural checklist items, such that their

attention was focused towards the specific behavioural manifestations on the checklist, rather than being asked to identify the behaviours spontaneously. This explanation would promote the benefits of providing more structured guidelines to custodial officers and highlighting the relevance of these positive behaviours. When custodial officers are required to monitor offenders' risk-related behaviour, they should be provided with guidelines that include both problematic and positive behaviours. As Gordon and Wong (2015) noted, current functioning assessed through behavioural monitoring can be viewed through "many lenses which are easily coloured by one's professional training, experience, situational demands, political and organisational pressures, not to mention subjective opinions and personal biases" (p. 97). Therefore, a template with examples of behaviours that should be monitored, may focus custodial officers' attention and facilitate objective assessment that includes a range of behaviours, including those that might not ordinarily be considered relevant. These behavioural examples could be used as anchor points for custodial officers to consider relevant behaviour that they observe; the list of examples would be indicative, but not exhaustive.

5.1.2.2.2 Study 2 and Study 3. In Study 2, custodial officers were provided with Likert-type scales within the SBRG; therefore, they could evaluate the offender's behaviour on more of a spectrum from 'antisocial' through to 'prosocial'. The custodial officers were not being asked to regularly monitor offenders' positive behaviour. Although their observations were not being focused on positive behaviour, and despite the presence of offenders' negative behaviour through incurring institutional misconducts, custodial officers tended to regard offenders' behaviour as positive (i.e., rated as Satisfactory). In general, offenders' behaviour was rated as Satisfactory overall, as they were in Study 3. This outcome suggests that custodial officers are more likely to rate offenders' behaviour favourably, which is also consistent with the greater increase in positive behaviour than negative behaviour as observed through the psychological domains in Study 3.

5.1.2.3 SBRG results. Within Study 2, all four categories of the SBRG were used (violent behaviour; substance abuse; attitudes and commitment to employment, education and/or rehabilitation; and, response to direction and supervision) in relation to misconducts, in addition to the overall behaviour rating. On the other hand, the focus of the SBRG in Study 3 was the overall behaviour rating. Nevertheless, it is noteworthy that there were similarities between the results of Studies 2 and 3. Within Study 2, the majority of the overall ratings were Satisfactory or Of Minor Concern. These results

mirrored the SBRG results in Study 3, which revealed that all but one of the behaviour ratings were reported as Satisfactory. Further, in Study 2, some participants who demonstrated poor behaviour, which resulted in reported incidents, were nevertheless assessed as Satisfactory. Similarly, in Study 3, participants who demonstrated minimal positive change or deterioration through other change measures (with some also incurring treatment suspensions due to poor behaviour) received an overall rating of Satisfactory.

The results from these two studies suggest that the SBRG is not a sensitive measure of either satisfactory behaviour or change over time. Based on the results of both Study 2 and Study 3, the SBRG might require significant problematic behaviour in order to elicit an unsatisfactory behaviour rating. Based on the results of Study 3, it also may not be sensitive to behaviour change within a custody-based treatment program for sexual offenders. Rather than a measure of change, it might be more effective as an overall measure of an offender's functioning, which was its intended use.

5.1.3 Integration of Results

One way in which to integrate the results is through the method by which behaviours were categorised to facilitate analysis; that is, the RSVP domains. There are suggestions within prior research regarding the benefits of using domains (Thornton, 2002, 2013). Thornton (2013) asserted that a limitation of considering only individual risk factors rather than integrating them into broader categories is that general patterns within the data might be neglected. Within the current research, the results in relation to the domains can provide information about offenders' manifest dynamic needs and the ways in which these factors change over time. The domains will be explored in this subsection, with specific reference to coping and problem solving, and cooperation with supervision. Further, the current results can be integrated through discussion about the method of data collection; that is, use of observer-report as a means by which to gain information about offenders' behaviour over time. After discussion of the RSVP domains, there will be some discussion about the use of observer data and the practical and logistical challenges within this process, through Studies 2 and 3.

5.1.3.1 Psychological domains. Within Study 3, due to the quantity and complexity of data, to facilitate analysis of the changes in behavioural manifestations of the risk factors, the data were reduced into domains, consistent with the domains assessed using the RSVP. The four RSVP domains can also be retrospectively applied to the SBRG behaviour categories in Study 2. This process may allow for further

comparisons and parallels drawn between the results of Study 2 and Study 3 based on the domain classifications. The behaviour categories could be classified as follows: Violent Incidents = Psychological Adjustment (coping and problem solving); Substance Use = Mental Disorder; Times Banned from Work = Social Adjustment (relationships and social influences); and, Direction/Supervision Incidents = Manageability (cooperation with supervision).

In Study 3, the most consistent results were gained with respect to coping and problem solving. This domain included behaviours such as verbal or physical aggression, withdrawal/isolation, raising issues with supports, focusing on solutions to problems, and engaging in a healthy routine. At the group-level, participants self-reported improvements over treatment through a decrease in negative behaviours and an increase in positive behaviours, while custodial officers reported improvements through an increase in positive behaviours within this domain.

There are various explanations for these results. Changes within this domain may have been observed because it included the most obvious behaviours, which were the easiest for participants and observers to monitor, or it may have had the greatest statistical power as it contained the greatest number of risk factors. Alternatively, the risk factors within this domain may have been the most susceptible to change; prior research has suggested that some risk factors are more dynamic than others, such that they are more susceptible to change (Olver et al., 2007). However, although some of the pre- to post-treatment psychometric measures revealed changes within this domain, this result was not obtained across all the relevant psychometric measures. If these behaviours were the most susceptible to change, given the validity of the psychometric measures, it would be expected that this result would be gained more consistently across the measures within this domain.

Another possible interpretation could be that the changes within this domain were representative of adjustment to the institutional environment rather than risk-related improvement. This concept has been discussed previously; Zamble and Porporino (1990) noted that offenders appeared to cope more effectively in custody as a result of their increased institutional adjustment, as opposed to their development of prosocial skills to counteract risk factors. However, in Study 3, participants' changes included a decrease in negative behaviours and an increase in positive behaviours, which suggests the development of relevant skills rather than merely reflecting institutional adjustment.

As discussed in previous chapters, there are differences in the visibility of behavioural manifestations of some dynamic risk factors compared with others (Mooney & Daffern, 2013), which appeared evident within the results of Study 1. That is, psychologists appeared to have greater ease in identifying predicted behavioural manifestations of risk factors related to coping and problem solving, which tend to be more observable. When psychologists are unable to directly observe the behavioural manifestations, they may predict a broader range of risk-related behaviours. Parallels can be drawn between this concept and the results in Study 3 in relation to coping and problem solving, such that behaviours that are more observable are more likely to both be spontaneously predicted and identified as evident within a custodial context. Therefore, it may be conjectured that the results from Study 1 provided evidence that the change results in Study 3 were due to more obvious behaviours associated with coping and problem solving.

In Study 2, higher frequencies of incidents related to violence and response to supervision/direction were predictive of a more severe overall behaviour rating. As stated previously, Violent Incidents are related to coping and problem solving. These incidents are generally perceived to be the most serious and relevant in custody (Gendreau et al., 1997) and previous research has noted that violence in custody is associated with violence in the community following release (Mooney & Daffern, 2015). This result in Study 2 is similar to some results from Study 3, in which certain risk factors related to coping and problem solving were more likely to be reported as relevant for offenders (e.g., Problems with Stress or Coping and Interpersonal Aggression).

Therefore, in addition to the most relevant and meaningful behavioural manifestations being identified in relation to coping and problem solving through Study 1; within both Study 2 and Study 3, this domain was predictive of Satisfactory behaviour ratings and relevant changes over time in treatment. Overall, the domain results from these three studies revealed some similarities in relation to the types of behaviours identified as relevant for offenders in custody.

A further comparison can be made between Study 2 and Study 3 with reference to cooperation with supervision. In Study 2, Direction/Supervision Incidents revealed the largest association with the final behaviour rating and had the largest contribution to the regression model. Similarly, in Study 3, positive behaviours associated with cooperation with supervision revealed the most prominent change over time. Although the focus of

this domain in Study 2 was negative behaviour, through institutional misconducts, while positive behaviours changed over time in Study 3; the combined results may indicate that behaviours related to cooperation with supervision are sensitive to measurement and predict the assessment of an offender's satisfactory behaviour. Custodial officers might focus on issues related to this domain, as it reflects offenders' interactions with staff and their ability or motivation to abide by rules within the custodial context. Increased attention towards, and awareness of, offenders' capacity to follow direction and instruction might lead to more reliable ratings from custodial officers.

These results suggest that psychological domains encompassing empirically supported dynamic risk factors may be reliable indicators of offenders' behaviour in custody. Researchers have previously discussed this concept. Thornton (2002) developed a framework, which included four risk-relevant domains: Sexual Interests; Distorted Attitudes; Relational Style; and, Self-Management. Each domain included 'subdomains' comprising risk factors and related behaviours, which could be targeted through treatment. Further, Klepfisz et al. (2016) asserted the potential benefits for the use of domains that subsume a number of more specific risk factors. Rather than relying on changes in single causal dynamic risk factors to determine an offender's functioning, Klepfisz et al. suggested the conceptualisation of dynamic risk in terms of broad levels, or domains. They concluded that while greater precision is required in relation to the causal mechanisms underpinning violence at the theoretical level, on the practical level it is crucial to consider how reductions in multiple dynamic risk factors or risk domains might independently and cumulatively be related to a reduction in violence (i.e., change over time). It might be valuable for future research to further investigate risk factor domains in relation to behavioural change. Perhaps change measures in which risk domains are the primary focus could provide clinicians with frameworks within which to measure offenders' progress through treatment.

5.1.3.2 Observer involvement. The purpose of Study 1 was to determine whether psychologists could identify behavioural manifestations of risk factors, which might be observed in a custodial context. Staff members within this context observe and monitor offenders' behaviour for differing reasons and consequently from different perspectives. For instance, psychologists generally do so from a therapeutic perspective to facilitate intervention and rehabilitation, while custodial officers generally do so from a safety and security perspective but in some environments also from a therapeutic perspective. Study 2 and Study 3 provided opportunities for the evaluation of the outcomes of this

observation and monitoring. The use of custodial officers' observations will now be discussed in the context of these two studies.

The current research demonstrated that custodial officers' observations can be used within the assessment of offenders' behaviour. Study 2 demonstrated that custodial officers can reliably rate the satisfactoriness of offenders' behaviour, which are supported by official records of institutional misconducts, in areas of dynamic risk. Study 3 demonstrated that custodial officers can observe changes in offenders' behaviour when provided with a structured checklist. However, in both Study 2 and Study 3, custodial officers appeared disinclined to provide Unsatisfactory behaviour ratings within the SBRG. This result was explored in both Study 2 and Study 3; it will now be reviewed. Subsequently, there will be further discussion related to custodial officers' observations within Study 3, based on prior research.

Atkinson and Mann (2012) discussed the reasons custodial officers provided for deciding not to report problematic behaviour. For instance, if the behaviour is considered 'normal' for the individual or within the custodial context, it might remain unreported. This explanation may be relevant for the current results. If custodial officers perceived participants' behaviour as 'normal', they might have been inclined to suggest the offender's behaviour was Satisfactory. In order for an Unsatisfactory rating, the behaviour may need to be viewed as particularly harmful to others.

Alternatively, these results might reflect custodial officers' motivation to complete assessment measures. For instance, Hanson, Helmus, and Harris (2015) suggested that some community supervising officers may lack the skills and motivation to complete valid risk assessments; it is possible that these characteristics are similar within the custodial environment. This concept might be related to the outcomes Atkinson and Mann (2012) discussed in relation to custodial officers' reports of problematic behaviour. The results from their study revealed that custodial officers do not necessarily report problematic behaviour they observe in the custodial environment. One of the reasons behind this decision not to report behaviour was that of officers' complacency. Decisions not to report problematic behaviour may be similar to decisions not to provide negative behaviour ratings. Custodial officers might be more inclined to provide a rating that could be considered less controversial and less likely to attract attention or discussion. In this instance, custodial officers might perceive a Satisfactory rating as likely to raise fewer questions than would an Unsatisfactory rating.

This potential for limited motivation is consistent with prior research. For instance, Miller and Maloney (2013) conducted a survey of frontline community corrections officers in relation to their compliance and noncompliance with risk/needs tools. One of the research questions was concerned with the range of types of noncompliance with these tools. Within the survey, officers were asked to provide ratings on Likert-type scales, indicating how often they engaged in the specified behaviour. Within the category “Tool completion” were eight items: complete tool when required; complete tool fully; fill out every question; complete/update carelessly; make minimum effort; exaggerate characteristics; minimise characteristics; and, manipulate information. Although the majority of respondents reported compliance with the tools, the results suggested some supervising officers were less compliant. Respondents were classified into one of three groups: substantive compliers (reported generally high levels of compliance); bureaucratic compliers (reported relatively high levels of compliance in completing the tool but lower compliance in relation to decision-making items); and, cynical compliers (reported relatively low levels of compliance across both tool completion and decision-making items). Within the cynical compliers, certain behaviours distinguished them from the other two groups. They were more likely to engage in careless tool completion, make minimum effort, exaggerate characteristics, minimise characteristics, and manipulate information. This result suggests that these respondents tended to complete the tool in a cursory manner, rather than in a manner ensuring accuracy. One potential explanation for this behaviour may be that these respondents lacked the motivation to place effort in the task.

It could be concluded that custodial officers’ limited engagement in the behavioural checklist data collection within Study 3 was similarly related to limited motivation. For instance, not only was there a large amount of missing data through weeks in which custodial officers chose not to complete the checklist, but there were also discrepancies between reported observations of participants’ behaviour and readily accessible information about participants’ actual behaviour. For example, in relation to medication compliance, some participants were reported to comply daily, despite having no prescribed medication. This example might reflect custodial officers’ ‘careless tool completion’ as in Miller and Maloney’s (2013) study described above. That is, a large proportion of treatment participants were required to attend the Justice Health clinic to receive their medication on a daily basis. Therefore, it may have been

easier for custodial officers to report all research participants as compliant with their medication, than to determine which participants had attended the clinic each day.

Overall, custodial officers' interactions with offenders may be influenced by their motivation to meaningfully engage with offenders, which subsequently affects their ability to provide accurate and substantive observations of offenders' behaviour. Prior research has investigated the various characteristics of custodial officers and the impact on their interactions with offenders, with the subsequent development of a typology of officers (Farkas, 2000). "Rule enforcers" was the most common of the five categories, with these officers typically working on posts involving direct offender contact, such as the regular housing units. However, these officers tended to express a preference to work in roles with less offender contact. Therefore, within the current research, it is likely that custodial officers who were less motivated to engage with participants may have had more superficial contributions within the data collection due to their focus on alternative issues such as rule compliance.

There remains the potential for gaining beneficial information from custodial officers regarding offenders' behaviour and related changes over time. However, this information may be more valuable if there are regularly rostered custodial officers within the therapeutic environment, who have therapeutic interests. In practice, if the behavioural checklist is implemented as a measure of change, it will be important for clinicians to be aware that they will likely face difficulties gaining consistent and methodical behavioural reports from custodial officers. There might be benefit in the provision of additional training in relation to behavioural observation and monitoring in a therapeutic community. Overall, it is vital that researchers and clinicians are aware of the potential strengths and limitations to the use of information gained from custodial officers' observations of offenders' behaviour.

5.1.4 Clinical and Policy Implications

The current research was relevant for the development of the field; in particular, through the identification of behavioural manifestations of risk factors and protective factors, and the subsequent measurement of change. These issues will now be discussed.

5.1.4.1 Behavioural manifestations of risk factors. Daffern et al. (2009) noted that monitoring progress, and release decision-making, tend to involve the assessment of observed improvements such as the development of prosocial attitudes, beliefs and behaviours that are incompatible with violence, in addition to indications of persistent problematic behaviour within the institution. The current research provided further

information about the use of OPB/OAB and PAB/ORB as indicators of risk and protective factors in an institutional environment, to assist with the evaluation of change and assessment of current functioning. In determining risk of recidivism, it is beneficial to identify the behavioural manifestations of an offender's criminogenic needs (dynamic risk factors) in custody such that they can be measured to determine change over time and current functioning. It is important to distinguish between the OPB and OAB frameworks. The behavioural manifestations of risk factors, which formed the focus of the current research, could be defined as OAB and ORB as opposed to meeting the criteria for OPB and PAB; that is, they were representative of the behaviours associated with offenders' risk factors for sexual offending, rather than sequences of behaviour relevant within the index offence.

Together, the results of these three studies provided greater insight into the assessment of offenders' behaviour in custody. The studies demonstrated that psychologists are generally aware of the behavioural manifestations of risk factors for incarcerated sexual offenders, and that these behaviours are exhibited and observed in custody. Behavioural manifestations of risk factors are evident through offenders' misconducts in custody, and through their ongoing behaviour within a treatment program. Consistent with McDougall et al. (2013), the behavioural checklist within Study 3 included a variety of behaviours, which were identified from empirically based risk factors but did not necessarily meet the OPB definition. The behaviours included both high-level behaviours, which more closely resembled offending behaviour, and low-level behaviours, which were considered less serious disruptive behaviours that would not necessarily lead to punishment by those in authority.

Polaschek (2017) highlighted that the skills developed through treatment are the intermediate treatment or change goals stemming from a formulation of the offender's dynamic risk factors. This idea relates to Andrews et al.'s (2011) assertion that dynamic risk factors represent intermediate targets of change. Therefore, both risk and protective factors may represent this intermediate goal, which strengthens the argument for monitoring the behavioural manifestations of these factors in treatment. It may be beneficial to conduct research in the future, which further explores relevant frameworks researchers and clinicians can use to assist with the identification of behavioural manifestations of relevant dynamic risk factors.

5.1.4.2 Change measurement. The three studies in the current research used several methods by which to gain information regarding offenders' behaviour in

custody: a survey of psychologists; official records of institutional misconduct; offender self-report; custodial officers' weekly observations; and, clinicians' and custodial officers' overall behaviour ratings. Within these methods, different types of measurement tools were implemented, as previously described.

The results from the various measurement tools revealed that offenders' manifestations of risk factors changed over time in treatment. However, at both the group- and individual-level, the changes within the measures did not necessarily correspond with each other. There are limitations and benefits to all these measures, as previously outlined in section 4.4.6.2.2.3 (p. 267) and reviewed earlier in this chapter (section 5.1.1.3.1, p. 278). It is important for evaluators to consider the limitations when interpreting results, prior to drawing conclusions regarding which measure to use to assess change. The use of multiple measures allows for the impact of these limitations to be minimised. Further, it allows for comparisons to be made between the results. However, this approach increases complexity and the possibility of inconsistent results, which will need to be reconciled. This process highlights the importance of further research into methods by which assessors can reliably integrate the disparate results gained from multiple change measures.

Prior research has suggested that institutional misconduct is a potentially useful source of information (French & Gendreau, 2006). Consistent with this view, results from Study 2 suggested that there remains value in using official records of institutional misconduct to assist in rating offenders' behaviour in custody. That is, overall, greater numbers of institutional misconducts were associated with more severe behaviour ratings, which suggests that offenders' behaviour in custody can be informed by the number of misconducts received. However, reliance on official institutional records to gauge offenders' behaviour and predict future risk can be problematic, as institutional misconduct represents only one of many factors that may inform comprehensive assessments (Mooney & Daffern, 2015). Within Study 3, multiple measures were implemented such that a wider range of factors could be evaluated. Study 3 demonstrated that implementing multiple measures of offenders' change in custody not only provided the opportunity for more factors to inform the assessments, but it also increased the complexity of the assessment process.

The current research provided an opportunity to compare the use of different measures. The SBRG and behavioural checklist will now be compared. Both these measures required respondents to complete a checklist based on particular behaviours.

Results revealed that both provided some value to the assessment of offenders' behaviour. In relation to the SBRG, on a global scale, custodial officers' ratings of offenders' behaviour were consistent with offenders' behaviour more generally, as indicated by official adjudications and work records. This result suggests that custodial officers may have an accurate understanding of offenders' behaviour. In relation to the behavioural checklist, overall, custodial officers could identify and monitor the frequency with which offenders engaged in behavioural manifestations of risk factors.

Notwithstanding the overall SBRG results, there was evidence to suggest that custodial officers tended to evaluate offenders' behaviour as Satisfactory despite receiving institutional misconducts. Therefore, it is possible that for such a complex issue as monitoring and evaluating sexual offenders' behaviour, a simple scale such as the SBRG may be inadequate. Rather, the behavioural checklist appeared to have greater capacity to focus custodial officers' attention on the relevant behavioural manifestations, which may have been related to the more structured framework, such that they can more effectively monitor these behaviours.

5.1.4.2.1 Risk and protective factors. In addition to focusing on risk factors, on the basis of prior research, there may also be benefit in the observation and monitoring of protective factors in the measurement of change over time (Klepfisz et al., 2017). Through containing both risk-related behaviours and prosocial alternative behaviours, the behavioural checklist in Study 3 facilitated further investigation into the relation between risk and protective factors. For instance, through the behavioural manifestations measured in the checklist, there was evidence to suggest that risk factors and their prosocial alternatives can co-exist, with some participants engaging in both types of behaviour simultaneously.

Jones et al. (2015) suggested that the various definitions of protective factors could be amalgamated and broadly considered as strength factors. This use of an umbrella term suggests that risk and protective factors can co-exist. Definitional distinctions have been made between protective and promotive factors. Protective factors moderate the impact of risk factors, such that risk must be present for protective factors to have an effect. On the other hand, promotive factors reduce the probability of offending regardless of the presence of risk (Monahan & Skeem, 2016). There remains room within these definitions for the co-existence of risk factors and their prosocial alternatives. That is, if there is a co-existence, the strength factors may be referred to as protective, while an absence of the risk would suggest the strength factor is promotive.

The current research highlighted further challenges within the definitional properties of protective factors. For instance, whether protective factors are distinct from risk factors' prosocial alternative behaviours. Just as risk factors have behavioural manifestations, it is important to distinguish protective, or strength, factors from their behavioural manifestations, as per the behavioural checklist items. To date, there has been minimal research exploring this distinction, likely due to the research into protective factors remaining formative. As Polaschek (2017) noted, formal investigation of protective factors during intervention remains uncommon. However, the results of the current research revealed that positive behavioural manifestations can be identified and monitored in a sample of incarcerated sexual offenders. Therefore, it may allow for further research into the properties of these factors and their behavioural manifestations in a custodial environment, particularly as offenders progress through treatment. For instance, future research may assist with the determination of whether protective factors within a treatment environment include both the offenders' motivation to use their internal and external resources to desist from offending, and those skills and resources themselves.

Within Study 3, changes related to prosocial behaviours on the checklist were associated with changes within other measures, for some individuals. A preliminary conclusion can be drawn from this association regarding the concurrent validity of the behavioural checklist and the use of positive behaviours in addition to negative behaviours. Further, the ultimate test of the use of protective factors is to assess whether they are related to behaviour change and subsequently a reduction in recidivism.

Polaschek (2017) suggested that conceptually, there is no advantage to using protective factors that are coded using the same information as risk factors, in treatment (that is, identifying protective factors as the opposing poles of risk factors), because the same information is gained. However, based on the suggested behavioural manifestation examples provided by psychologists in Study 1, some of the positive behavioural manifestations in the behavioural checklist developed for Study 3 did represent the opposing poles of the negative behavioural manifestations. The differing change results gained from the negative and positive behavioural domains may suggest that in completing the checklists, there is a distinction between the positive and negative behaviours observed even if they represent the two poles of the same overarching factor. That is, if there was no additional information to be gained through the positive behaviours, the results should have been comparable between the negative and positive

behaviours. The changes over time across these behaviours did not occur at the same rates for the related negative and positive domains. However, the amalgamation of the risk factors into the four domains led to a loss of specificity within the results. Therefore, the individual risk factors that contributed to these results remain unidentified; it is unknown whether the risk factors with the opposing poles contributed to the discrepant results.

Serin, Chadwick, et al. (2016) asserted that dynamic risk factors relate to crime acquisition and expression while protective factors reflect crime desistance. This concept would allow for the prosocial alternative behaviours, which were developed through skill acquisition to counter behavioural manifestations of risk factors, to be considered protective factors. This idea is also supported by prior research asserting that a protective factor interacts with risk, reducing the impact of a negative event, such that the addition of each protective factor further moderates the effect of risk exposure (Lösel & Farrington, 2012). The key distinction for the definition of a protective factor is the interaction between the protective factor and risk (Serin, Chadwick, et al., 2016). The idea that prosocial alternative behaviours might be negatively correlated with recidivism, such that as this behaviour increases, the risk of recidivism decreases, might be more akin to the concept of promotive factors. Unlike protective factors, promotive factors directly reduce reoffending, but are not independent of risk (Loeber et al., 2007). Some promotive factors are the inverse of risk factors (Serin, Chadwick, et al., 2016).

The results from the current research also suggested that behavioural manifestations of risk factors and their prosocial alternatives may change at different rates. For instance, positive behaviours appeared to change to a greater extent over treatment compared with negative behaviours. It is possible that the positive behaviours were developed during the initial stages of treatment but the negative behaviours persisted because the positive skills were not yet sufficiently strengthened. This suggests that these two concepts are not merely the opposites of each other, providing support for the view that some protective factors might have a non-linear relationship with risk factors (Polaschek, 2017). The ways in which risk and protective factors differentially change over time in treatment may constitute valuable future research, which could be conducted using a similar process to that used in the current research such as through use of a behavioural checklist identifying predicted behaviours.

5.1.5 Challenges and Limitations

There were several challenges faced within this research, which reflected the difficulties often encountered when conducting research in this field. The challenges associated with this form of research will again be outlined, with particular reference to Study 3. The challenges were also representative of the difficulties clinicians must overcome in this context when evaluating offenders' change and functioning including current risk level. The implications of these challenges will be discussed. The key limitations for each study will be reiterated, after which, limitations of the research as a whole will be described.

5.1.5.1 Challenges within current research. The current studies highlighted the challenges associated with this form of research. The reliability of the measures in part relies on the skill and motivation of the individuals completing the measures. Study 3 in particular revealed difficulties within the assessment and change arena when using offender self-report over an extended period of time, through the withdrawal of several participants who no longer wished to complete the behavioural checklists. In addition, due to the operational factors described in the previous chapter, data from custodial officers was similarly limited. Due to the missing data, there were implications for the strength of the conclusions that can be drawn from the current research.

Further, the challenges within the research extended to the practical use of the information gained from four different measures. The difficulties clinicians face within this field have been discussed. To reiterate, clinicians must integrate disparate change information to form an opinion about an individual offender. It is important for clinicians to form a coherent account of the individual's change such that it can inform decisions. In order for this account to be formed, the clinician must have a sophisticated understanding of the individual offender's relevant risk factors, such that changes within these factors can be evaluated.

5.1.5.2 Limitations of current research. The key limitations of each study will now be outlined. Limitations of the research as a whole will also be described.

5.1.5.2.1 Study 1. In Study 1, participants were psychologists, who observe offenders from a particular perspective, which is not necessarily consistent with custodial officers' perspectives. Therefore, the suggested behavioural manifestations of risk factors in the survey might have excluded potentially important behaviours on which psychologists do not generally focus. There might have been a discrepancy

between the behaviours on the resulting behavioural checklist and the behaviours custodial officers are able to monitor, as represented in Study 3.

5.1.5.2.2 Study 2. Ethical considerations prevented the differentiation between sexual and non-sexually violent offenders within Study 2. There are two primary issues, which relate to this lack of differentiation. First, there could be no comparison between offence type within the study. Second, there were limitations in the validity of comparisons between the samples in Study 2 and Study 3 in relation to the SBRG results.

Although the SBRG includes the evaluation of both negative and positive behaviour, the focus of Study 2 was offenders' negative behaviour outcomes through institutional misconducts. While case workers were asked to consider all behaviour in their overall behaviour rating, negative outcomes were the focus within the test for validity. That is, official misconducts were used as a comparison measure of offenders' functioning, which may have provided a biased perspective. Further, the SBRG does not necessarily provide information about behavioural change over time, as there is no structured evaluation of a change in frequency of misconducts over time in custody.

5.1.5.2.3 Study 3. It is important to note that the small sample size and quantity of missing data in Study 3 limited the generalisability of the results. It highlighted the challenges inherent in clinical research of this nature. However, it provided the foundations for possible further research to gain greater insight into the most valid and reliable methods by which to measure change over time in a custodial setting.

5.1.5.2.4 Overall limitations. While Study 1 included a combination of international and domestic participants, Study 2 and Study 3 were based on an Australian prison population. Two potential limitations arise from this difference. First, the behaviours identified within Study 1, which were included in the behavioural checklist in Study 3, may not be generalisable from those predicted by professionals internationally, to Australian offenders. Second, the Australian prison population likely differs demographically from that of different countries. Therefore, the results may not be generalisable to offenders internationally.

An important part of change research is determining what caused the change (Jones et al., 2007). It is not possible to draw valid conclusions regarding the causes of changes observed in offenders' behaviour, due to the lack of a control group. There are several potential causes of the changes, with the treatment program presenting just one of these potential causes. However, the current research was not an evaluation of

treatment effectiveness. Rather, the aims were related to observing and identifying the changes displayed through various measurement tools and the ways in which these observations may be interpreted in a clinical context; in particular, in relation to the ways in which the OPB and OAB frameworks can be applied in this context.

5.1.6 Future Research

In the past, custodial officers have been asked to provide assistance in the development of a checklist such that they could monitor incarcerated offenders' behaviour (McDougall et al., 1995). The current research sought to gain information about predicted behaviours as described by psychologists with experience working with sexual offenders in custody. In the future, psychologists and custodial officers working within the custodial environment could be asked to provide information such that combined predictions about behavioural manifestations of risk factors could be made. Such a process may facilitate greater depth and variety of information due to the different contexts within which these individuals observe offenders in custody and, therefore, the different perspectives they hold. Further, additional research could be conducted to facilitate the development of more effective methods by which to gain information from custodial officers regarding offenders' behaviour.

The current research represented a continuation of the early stages of research into change measurement for incarcerated offenders. In the future, the behavioural checklist developed from Study 1 and implemented in Study 3 could be modified based on the current results and the limitations observed within Study 3. For instance, the checklist could be condensed to allow for a less burdensome tool (e.g., reduce perceived repetition of items), and the risk factors associated with the behavioural manifestations could be labelled such that respondents are aware of the distinction between behaviours. These changes to the checklist might facilitate increased engagement from custodial officers and an increase in the reliability and validity of the measure. Further, based on the results of the comparisons between the two behavioural manifestations of each risk factor, this modification of the checklist could include additional analysis to determine the items that provided greater value and removing those items that provided no added benefit. Another potential modification could be the inclusion of idiographic behaviours based on an individual case formulation. Although Harris and Hanson (2010) noted that in their outcome studies using the ACUTE-2000, the option for evaluators to rate unique factors that were important to a particular individual had no relationship with subsequent recidivism; this use of idiographic case formulations in the development and

use of structured professional judgement tools may be beneficial if used correctly. Inclusion of idiographic items could be beneficial for individual-level analyses. Following these modifications, the behavioural checklist could be validated to allow for more widespread use.

Further to the revision of the behavioural checklist, the SBRG could also be modified to produce a more sensitive measure of change. The SBRG was developed as a simple measure to evaluate the satisfactoriness of offenders' behaviour. Currently, this measure is not a valid change assessment tool. If it is modified in the future, rather than only using an overall behaviour rating, this measure could target change over time in its constituent parts and the overall rating could be an evaluation of change.

Although not possible with respect to participants in Study 2 due to ethical considerations, in the future, a follow-up study could be conducted in relation to participants in Study 3. Study 3 involved research into within-treatment change using four measures. Observation of behaviours post-release to the community would facilitate further insight into cross-situational behavioural consistency and use of the measures as predictors for future behaviour. In order to evaluate the change measure that is the most valid, there is a need for recidivism data. However, due to the low base rates of sexual recidivism, this process would require a large database with a long follow-up period. A prospective study such as that completed in the current research would be ideal.

A crucial component of clinical work is the integration of information from a variety of sources in order to make a fully informed evaluation. As described in the previous chapter, the current research was representative of this process in the forensic field, through the combination of change measures and methods by which to gain information about offenders' functioning. Particularly within Study 3, discussion focused on the challenges faced within the custodial context in the observation, monitoring, and subsequent evaluation of offenders' behavioural change over time. The field would greatly benefit from ongoing research within this arena. There is a need for the development of a method by which results from disparate change measures can be integrated to assist clinicians in the evaluation of offenders' change over a specific period of time such as through treatment.

5.1.7 Summary and Conclusions

The definition and measurement of dynamic risk factors have formed the focus of extensive prior research. However, questions remain regarding the manifestations of

dynamic risk factors in incarcerated offenders, the ways in which these manifestations change over time, and the methods by which to measure such changes. The current research facilitated exploration of these pertinent issues.

Importantly, these three studies elaborated on prior research in the area of behavioural monitoring and change measurement in custody. Several components of the current research represented the development of concepts that remain in the preliminary stages of research. For instance: use of professionals in the field to determine current views of incarcerated offenders' behavioural manifestations of risk factors; combining different methods by which change can be measured; validation of a tool for custodial officers to make an overall behaviour rating for offenders; application of the OAB framework to develop a behavioural checklist to aide offenders and custodial officers within a treatment environment; and, use of multiple time points at which to determine offenders' functioning to facilitate the measurement of change over time.

Overall, the current research revealed that incarcerated offenders' behaviour corresponded with predicted risk-related behaviours and their prosocial equivalent behaviours. Subsequent to the development of a behavioural checklist and validation of the SBRG, four measurement tools were implemented to evaluate offenders' change over time in treatment. Offenders' behaviours were observed to change, with some risk domains evidencing greater change than others. Both offenders and custodial officers reported change over time; there was evidence for the value of gaining information from both offenders and custodial officers due to the strengths and limitations of both self- and observer-report.

The current research also provided further evidence for the relevance of measuring change at the individual-level rather than remaining reliant on group-level outcomes. Both Study 2 and Study 3 revealed some discrepancies between group- and individual-level results. In the institutional context, various levels of decision-making (e.g., classification progression, treatment completion, and parole) are conducted based on an individual offender's behaviour and changes over time. There is value at the organisational level in the use of group outcomes to ensure consistency in decision-making and theoretically-based decisions. Nevertheless, the processes and decisions related to individual offenders in custody must be based on idiographic information. As Pearson and McDougall (2017) asserted, it is particularly important when considering progression for serious offenders, to determine whether there is evidence of improvement or deterioration specifically in the present case.

The current research revealed the complexities inherent in change measurement when multiple measures are implemented. Within clinical practice, psychologists are often required to interpret and integrate the results of multiple measures in order to make an evaluation regarding an individual offender's behaviour. In order to do so, psychologists must have a clear understanding of the risk factors that are relevant for the particular offender, such that changes within the behavioural manifestations of these specific risk factors can be determined. This information can form the basis for further investigation and assessment of change. Gaining information from a variety of sources, such as offender self-report and observations from other staff may assist in the collation of information. It may allow the psychologist to determine whether the observations he/she has made, are congruent with other available information to ensure minimisation of potential bias.

The measures used in the current research provided varying levels of structure for the individuals completing them (i.e., custodial officers, therapists, and offenders). There is benefit to engaging a structured process for observations of offenders' behaviour. For instance, McDougall et al. (1995) noted that providing custodial officers with checklists for certain risk factors could provide a more consistent approach to the system of behavioural monitoring. However, at times structure can create rigidity if it is applied strictly and mechanically (Gordon & Wong, 2015), and clinical judgement can be neglected. In the future, in order for clinicians to derive more value from the checklist, information gained from its completion could be incorporated in a structured professional judgement tool, similar to the process implemented in the VRS:SO. If the behavioural checklist in the current research is modified to include individualised items in addition to the current generalised items, it would be a beneficial addition to such an assessment process.

The current research as a whole may represent another development within the change measurement field. Further research questions were raised, which can serve to promote investigation into the most effective methods by which behavioural manifestations of risk factors can be observed and monitored, to facilitate the continued progression of our knowledge of and practice in change measurement.

References

- Adult Parole Board of Victoria. (2015). *Parole Manual* (Fifth ed.). Carlton, Victoria: Adult Parole Board of Victoria.
- Allan, M., Grace, R., Rutherford, B., & Hudson, S. (2007). Psychometric Assessment of Dynamic Risk Factors for Child Molesters. *Sexual Abuse: A Journal of Research and Treatment*, 19, 347-367. doi: 10.1007/s11194-007-9052-5
- Andrews, D., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). Cincinnati, OH: Anderson.
- Andrews, D., & Bonta, J. (2016). *The psychology of criminal conduct* (6th ed.). New York, NY: Taylor & Francis Ltd.
- Andrews, D., Bonta, J., & Hoge, R. (1990). Classification for effective rehabilitation: Rediscovering psychology. *Criminal Justice and Behavior*, 17, 19-52. doi: 10.1177/0093854890017001004
- Andrews, D., Bonta, J., & Wormith, J. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency*, 52, 7-27. doi: 10.1177/0011128705281756
- Andrews, D., Bonta, J., & Wormith, J. (2011). Risk-Need-Responsivity (RNR) Model: Does adding the Good Lives Model contribute to effective crime prevention? *Criminal Justice and Behavior*, 38, 735-755. doi: 10.1177/0093854811406356
- Andrews, P., & Meyer, R. (2003). Marlowe-Crowne Social Desirability Scale and Short Form C: Forensic norms. *Journal of Clinical Psychology*, 59, 483-492. doi: 10.1002/jclp.10136
- Andrews, D., Zinger, I., Hoge, R., Bonta, J., Gendreau, P., & Cullen, F. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28, 369-404. doi: 10.1111/j.1745-9125.1990.tb01330.x
- Australian Bureau of Statistics (2016). Personal Safety Survey. Retrieved from www.abs.gov.au/ausstats/abs@.nsf/mf/4906.0
- Baglivio, M., Wolff, K., Piquero, A., Howell, J., & Greenwald, M. (2017). Risk assessment trajectories of youth during juvenile justice residential placement: Examining risk, promotive, and "buffer" scores. *Criminal Justice and Behavior*, 44, 360-394. doi: 10.1177/0093854816668918
- Bakker, L., Hudson, S., Wales, D., & Riley, D. (1998). *And there was light... Evaluating the Kia Marama treatment programme for New Zealand sex offenders*

against children. Retrieved from New Zealand Department of Corrections website:

https://www.corrections.govt.nz/__data/assets/pdf_file/0004/665635/kiamarama.pdf

- Ballard, R. (1992). Short forms of the Marlowe-Crowne Social Desirability Scale. *Psychological Reports, 71*, 1155-1160. doi: 10.2466/pr0.1992.71.3f.1155
- Barnett, G., Wakeling, H., & Howard, P. (2010). An examination of the predictive validity of the Risk Matrix 2000 in England and Wales. *Sexual Abuse: A Journal of Research and Treatment, 22*, 443-470. doi: 10.1177/1079063210384274
- Barnett, G., Wakeling, H., Mandeville-Norden, R., & Rakestrow, J. (2012). How useful are psychometric scores in predicting recidivism for treated sex offenders? *International Journal of Offender Therapy and Comparative Criminology, 56*, 420-446. doi: 10.1177/0306624X11403125
- Beech, A. (1998). A psychometric typology of child abusers. *International Journal of Offender Therapy and Comparative Criminology, 42*, 319-339. doi: 10.1177/0306624X9804200405
- Beech, A., & Craig, L. (2012). The current status of static and dynamic factors in sexual offender risk assessment. *Journal of Aggression, Conflict and Peace Research, 4*, 169-185. doi: 10.1108/17596591211270671
- Beech, A., & Ward, T. (2004). The integration of etiology and risk in sexual offenders: A theoretical framework. *Aggression and Violent Behavior, 10*, 31-63. doi: 10.1016/j.avb.2003.08.002
- Beggs, S., & Grace, R. (2010). Assessment of dynamic risk factors: An independent validation study of the Violence Risk Scale: Sexual Offender version. *Sexual Abuse: A Journal of Research and Treatment, 22*, 234-251. doi: 10.1177/1079063210369014
- Beitchman, J., Zucker, K., Hood, J., DaCosta, G., & Akman, D. (1991). A review of the short-term effects of child sexual abuse. *Child Abuse and Neglect, 15*, 537-556. doi: 10.1016/0145-2134(91)90038-F
- Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: Logistic regression. *Critical Care, 9*, 112-118. doi: 10.1186/cc3045
- Blagden, N., Winder, B., & Hames, C. (2016). "They treat us like human beings" - Experiencing a therapeutic sex offenders prison: Impact on prisoners and staff

- and implications for treatment. *International Journal of Offender Therapy and Comparative Criminology*, 60, 371-396. doi: 10.1177/0306624X14553227
- Bonham, G., Janeksela, G., & Bardo, J. (1986). Predicting parole decision in Kansas via discriminant analysis. *Journal of Criminal Justice*, 14, 123-133. doi: 10.1016/0047-2352(86)90060-7
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101. doi: 10.1191/1478088706qp063oa
- Braun, V., & Clarke, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners*. London: Sage.
- Brown, S., St. Amand, M., & Zamble, E. (2009). The dynamic prediction of criminal recidivism: A three-wave prospective study. *Law and Human Behavior*, 33, 25-45. doi: 10.1007/s10979-008-9139-7
- Bulmer, M. (1979). *Principles of Statistics*. New York: Dover.
- Canadian Resource Centre for Victims of Crime. Retrieved from <https://crcvc.ca/category/sexual-assault/>
- Canter, D. (2000). Offender profiling and criminal differentiation. *Legal and Criminological Psychology*, 5, 23-46. doi: 10.1348/135532500167958
- Caplan, J. (2007). What factors affect parole: A review of empirical research. *Federal Probation Journal*, 71(1), 16-19. Retrieved from http://www.uscourts.gov/sites/default/files/71_1_3_0.pdf
- Carroll, J., Wiener, R., Coates, D., Galegher, J., & Alibrio, J. (1982). Evaluation, diagnosis, and prediction in parole decision making. *Law & Society Review*, 17, 199-228. doi: 10.2307/3053536
- Carter, A., & Mann, R. (2016). The strengths of treatment for sexual offending. In R. Laws & W. O'Donohue (Eds.), *Treatment of sex offenders: Strengths and weaknesses in assessment and intervention* (pp. 157-174). Cham, Switzerland: Springer International Publishing.
- Casey, S. (2016). Dynamic risk and sexual offending: The conundrum of assessment. *Psychology, Crime & Law*, 22, 104-123. doi: 10.1080/1068316X.2015.1111366
- Christensen, L., & Mendoza, J. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy*, 17, 305-308. doi: 10.1016/s0005-7894(86)80060-0

- Clarke, D., Fisher, M., & McDougall, C. (1993). A new methodology for assessing the level of risk in incarcerated offenders. *British Journal of Criminology*, 33(3), 436-448. Retrieved from <https://www.jstor.org/journal/britjcrim>
- Cochran, J., Mears, D., Bales, W., & Stewart, E. (2012). Does inmate behaviour affect post-release offending? Investigating the misconduct-recidivism relationship among youth and adults. *Justice Quarterly*, 31, 1044-1073. doi: 10.1080/07418825.2012.736526
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen-Mansfield, J., Marx, M., & Rosenthal, A. (1989). A description of agitation in a nursing home. *Journal of Gerontology*, 44, M77-M84. doi: 10.1093/geronj/44.3.M77
- Collaborative Outcome Data Committee. (2007). Sex offender treatment outcome research: Guidelines for evaluation (CODC Guidelines), Part 1: Introduction and overview. Ottawa, Ontario: Public Safety Canada. Retrieved from <https://www.publicsafety.gc.ca/cnt/rsrscs/pblctns/cllbrtv-tcmdt-gdlns/index-en.aspx>
- Cording, J., Beggs Christofferson, S., & Grace, R. (2016). Challenges for the theory and application of dynamic risk factors. *Psychology, Crime & Law*, 22, 84-103. doi: 10.1080/1068316X.2015.1111367
- Corey, G. (2017). *Theory and practice of counseling and psychotherapy*. Boston, MA: Cengage Learning.
- Cortoni, F., & Marshall, W. L. (2001). Sex as a coping strategy and its relationship to juvenile sexual history and intimacy in sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13, 27-43. doi: 10.1177/107906320101300104
- Craig, L., Browne, K., & Stringer, I. (2003a). Risk scales and factors predictive of sexual offence recidivism. *Trauma, Violence, & Abuse*, 4, 45-69. doi: 10.1177/1524838002238945
- Craig, L., Browne, K., & Stringer, I. (2003b). Treatment and sexual offence recidivism. *Trauma, Violence, & Abuse*, 4, 70-89. doi: 10.1177/1524838002238946
- Craissati, J., & Beech, A. (2003). A review of dynamic variables and their relationship to risk prediction in sex offenders. *Journal of Sexual Aggression*, 9, 41-55. doi: 10.1080/355260031000137968

- Crowne, D., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354. doi: 10.1037/h0047358
- Cutajar, M., Mullen, P., Ogloff, J., Thomas, S., Wells, D., & Spataro, J. (2010a). Psychopathology in a large cohort of sexually abused children followed up to 43 years. *Child Abuse and Neglect*, 34, 813-822. doi:10.1016/j.chiabu.2010.04.004
- Cutajar, M., Mullen, P., Ogloff, J., Thomas, S., Wells, D., & Spataro, J. (2010b). Schizophrenia and other psychotic disorders in a cohort of sexually abused children. *Archives of General Psychiatry*, 67, 1114-1119. doi: 10.1001/archgenpsychiatry.2010.147
- Daffern, M. (2010). A structured cognitive behavioural approach to the assessment and treatment of violent offenders using offence paralleling behaviour. In M. Daffern, L. Jones & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 105-120). Chichester: John Wiley and Sons Ltd.
- Daffern, M., Ferguson, M., Ogloff, J., Thomson, L., & Howells, K. (2007). Appropriate treatment targets or products of a demanding environment? The relationship between aggression in a forensic psychiatric hospital with aggressive behaviour preceding admission and violent recidivism. *Psychology, Crime and Law*, 13, 431-441. doi: 10.1080/10683160601060861
- Daffern, M., & Howells, K. (2009). The function of aggression in personality disordered patients. *Journal of Interpersonal Violence*, 24, 586-600. doi: 10.1177/0886260508317178
- Daffern, M., Howells, K., Mannion, A., & Tonkin, M. (2009). A test of methodology intended to assist detection of aggressive offence paralleling behaviour within secure settings. *Legal and Criminological Psychology*, 14, 213-226. doi: 10.1348/135532508X342919
- Daffern, M., Howells, K., Stacey, J., Hogue, T., & Mooney, P. (2008). Is sexually abusive behaviour in personality disordered inpatients analogous to sexual offences committed prior to hospitalization? *Journal of Sexual Aggression*, 14, 123-133. doi: 10.1080/13552600802267080
- Daffern, M., Jones, L., Howells, K., Shine, J., Mikton, C., & Tunbridge, V. (2007). Refining the definition of offence paralleling behaviour. *Criminal Behaviour and Mental Health*, 17, 265-273. doi: 10.1002/cbm.671

- Daffern, M., & Ogloff, J. (2017). Assessment issues in offending populations in secure settings. In J. Ireland, C. Ireland, N. Gredecki, & M. Fisher (Eds.), *The Routledge International Handbook of Forensic Psychology in Secure Settings* (pp. 81-96). Oxon: Routledge Taylor and Francis Group.
- Davies, J., Jones, L., & Howells, K. (2010). Evaluating individual change. In M. Daffern, L. Jones & J. Shine (Eds.), *Offence Paralleling Behaviour: A case formulation approach to offender assessment and intervention* (pp. 287-302). Chichester: John Wiley & Sons, Ltd.
- de Jong, R., Alink, L., Bijleveld, C., Finkenauer, C., & Hendriks, J. (2015). Transition to adulthood of child sexual abuse victims. *Aggression and Violent Behavior*, 24, 175-187. doi: 10.1016/j.avb.2015.04.012
- de Vogel, V., de Ruiter, C., Bouman, Y., & de Vries Robbé, M. (2009). SAPROF. Guidelines for the assessment of protective factors for violence risk. English version. Utrecht, The Netherlands: Forum Educatief.
- de Vries Robbé, M., de Vogel, V., Koster, K., & Bogaerts, S. (2015). Assessing protective factors for sexually violent offending with the SAPROF. *Sexual Abuse: Journal of Research and Treatment*, 27, 51-70. doi: 10.1177/1079063214550168
- de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective Factors for Violence Risk in Forensic Psychiatric Patients: A Retrospective Validation Study of the SAPROF. *International Journal of Forensic Mental Health*, 10, 178-186. doi: 10.1080/14999013.2011.600232
- Dennis, J., Khan, O., Ferriter, M., Huband, N., Powney, M., & Duggan, C. (2012). Psychological interventions for adults who have sexually offended or are at risk of offending (Review). *The Cochrane Collaboration*, 2012, 1-95. doi: 10.1002/14651858.CD007507.pub2
- Deshields, T., Tait, R., Gfeller, J., & Chibnall, J. (1995). Relationship between social desirability and self-report in chronic pain patients. *Clinical Journal of Pain*, 11, 189-193. doi: 10.1097/00002508-199509000-00005
- Desmarais, S., Nicholls, T., Wilson, C., & Brink, J. (2012). Using dynamic risk and protective factors to predict inpatient aggression: Reliability and validity of START assessments. *Psychological Assessment*, 24, 685-700. doi: 10.1037/a0026668

- Douglas, K. S., Hart, S., Webster, C., & Belfrage, H. (2013). HCR-20V3: Assessing risk of violence - User guide. Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.
- Douglas, K., & Skeem, J. (2005). Violence Risk Assessment: Getting Specific About Being Dynamic. *Psychology, Public Policy, and Law*, 11, 347-383. doi: 10.1037/1076-8971.11.3.347
- Doyle, M., & Dolan, M. (2002). Violence risk assessment: combining actuarial and clinical information to structure clinical judgements for the formulation and management of risk. *Journal of Psychiatric and Mental Health Nursing*, 9, 649-657. doi: 10.1046/j.1365-2850.2002.00535.x
- Dvoskin, J., & Heilbrun, K. (2001). Risk assessment and release decision-making: Toward resolving the great debate. *Journal of the American Academy of Psychiatry and the Law*, 29(1), 6-10. Retrieved from <http://jaapl.org/content/29/1/6>
- Edens, J., Buffington, J., Tomicic, T., & Riley, B. (2001). Effects of positive impression management on the Psychopathic Personality Inventory. *Law and Human Behavior*, 25, 235-256. doi: 10.1023/A:1010793810896
- Endler, N., & Parker, J. (1990). Multidimensional assessment of coping: A critical evaluation. *Journal of Personality and Social Psychology*, 58, 844-854. doi: 10.1037/0022-3514.58.5.844
- Ensink, K., Godbout, N., Bigras, N., Lampron, J., Sabourin, S., & Normandin, L. (2018). Persistent and transitory sexualized behavior problems in children. *Child Psychiatry & Human Development*. Advance online publication. doi: 10.1007/s10578-017-0778-0
- Farkas, M. (2000). A typology of correctional officers. *International Journal of Offender Therapy and Comparative Criminology*, 44, 431-449. doi: 10.1177/0306624X00444003
- Federal Bureau of Investigations (2016). *Crime in the United States*. Retrieved from <https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016/topic-pages/tables/table-1>
- Field, A. (2009). *Discovering statistics using SPSS* (Third ed.). London: Sage.
- Fischer, D., & Fick, C. (1993). Measuring social desirability: Short forms of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement*, 53, 417-424. doi: 10.1177/0013164493053002011

- Fisher, D., Beech, A., & Browne, K. (1999). Comparison of sex offenders to nonoffenders on selected psychological measures. *International Journal of Offender Therapy and Comparative Criminology*, 43, 473-491. doi: 10.1177/0306624X99434006
- Fortune, C., Ward, T., & Willis, G. (2012). The rehabilitation of offenders: Reducing risk and promoting better lives. *Psychiatry, Psychology and Law*, 19, 646-661. doi: 10.1080/13218719.2011.615809
- Fraboni, M., & Cooper, D. (1989). Further validation of three short forms of the Marlowe–Crowne scale of social desirability. *Psychological Reports*, 65, 595-600. doi: 10.2466/pr0.1989.65.2.595
- French, S., & Gendreau, P. (2006). Reducing prison misconducts: What works! *Criminal Justice and Behavior*, 33, 185-218. doi: 10.1177/0093854805284406
- Friendship, C., Falshaw, L., & Beech, A. (2003). Measuring the real impact of accredited offending behaviour programmes. *Legal & Criminological Psychology*, 8, 115. doi: 10.1348/135532503762871282
- George, D., & Mallery, M. (2010). *SPSS for Windows step by step: A simple study guide and reference* (10th ed.). Boston: Pearson.
- Gordon, A., & Wong, S. (2009). *The offence analogue and offence reduction behavior rating guide*. Unpublished user manual.
- Gordon, A., & Wong, S. (2010). Offence analogue behaviours as indicators of criminogenic need and treatment progress in custodial settings. In M. Daffern, L. Jones & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 171-183). Chichester: John Wiley and Sons Ltd.
- Gordon, A., & Wong, S. (2015). The use of a structured guide to assess proxies of offending behaviours and change in custodial settings. *Journal of Forensic Practice*, 17, 83-100. doi: 10.1108/JFP-10-2013-0048
- Gottfredson, M. (1979). Parole board decision making: A study of disparity reduction and the impact of institutional behavior. *Journal of Criminal Law and Criminology*, 70(1), 77-88. Retrieved from <https://scholarlycommons.law.northwestern.edu/cgi/viewcontent.cgi?referer=https://www.google.ca/&httpsredir=1&article=6110&context=jclc>
- Grady, M., Edwards Jnr, D., & Pettus-Davis, C. (2017). A longitudinal outcome evaluation of a prison-based sex offender treatment program. *Sexual Abuse: A*

Journal of Research and Treatment, 29, 239-266. doi:
10.1177/1079063215585731

- Graham, J., Olchowski, A., & Gilreath, T. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, 8, 206-213. doi: 10.1007/s11121-007-0070-9
- Grubin, D., Kelly, P., & Brunson, C. (2001). *Linking serious sexual assaults through behaviour*. London: Home Office Research, Development and Statistics Directorate.
- Halstead, I. (2016). Does the Custody-based Intensive Treatment (CUBIT) program for sex offenders reduce re-offending? *Crime and Justice Bulletin* (Vol. 193). Sydney: NSW Bureau of Crime Statistics and Research.
- Hammond, S. (2013). Introducing a common range index of inter-variable similarity for the analysis of radex structures. In D. Youngs (Ed.), *The behavioural analysis of crime: Studies in David Canter's investigative psychology* (pp. 119-140). New York: Ashgate Publishing.
- Hannah-Moffat, K. (2016). A conceptual kaleidoscope: Contemplating 'dynamic structural risk' and an uncoupling of risk from need. *Psychology, Crime & Law*, 22, 33-46. doi: 10.1080/1068316X.2015.1114115
- Hanson, R., Bourgon, G., Helmus, L., & Hodgson, S. (2009). The principles of effective correctional treatment also apply to sexual offenders: A meta-analysis. *Criminal Justice and Behavior*, 36, 865-891. doi: 10.1177/0093854809338545
- Hanson, K., & Bussière, M. (1998). Predicting relapse: A meta-analysis of sexual offender recidivism studies. *Journal of Consulting and Clinical Psychology*, 66, 348-362. doi: 10.1037/0022-006X.66.2.348
- Hanson, K., Gordon, A., Harris, A., Marques, J., Murphy, W., Quinsey, V., & Seto, M. (2002). First report of the collaborative outcome data project on the effectiveness of psychological treatment for sex offenders. *Sexual Abuse: A Journal of Research and Treatment*, 14, 169-194. doi: 10.1177/107906320201400207
- Hanson, K., & Harris, A. (2000). Where should we intervene?: Dynamic predictors of sexual assault recidivism. *Criminal Justice and Behavior*, 27, 6-35. doi: 10.1177/0093854800027001002
- Hanson, R., & Harris, A. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13, 105-122. doi: 10.1023/A:1026600304489

- Hanson, R., Harris, A., Scott, T.-L., & Helmus, L. (2007). Assessing the risk of sexual offenders on community supervision: The Dynamic Supervision Project (Vol. Corrections Research User Report No. 2007-05). Ottawa, Canada: Public Safety Canada.
- Hanson, K., Helmus, L., & Harris, A. (2015). Assessing the risk and needs of supervised sexual offenders: A prospective study using STABLE-2007, Static-99R, and Static-2002R. *Criminal Justice and Behavior*, 42, 1205-1224. doi: 10.1177/0093854815602094
- Hanson, K., & Morton-Bourgon, K. (2005). The characteristics of persistent sexual offenders: A meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, 73, 1154-1163. doi: 10.1037/0022-006X.73.6.1154
- Hanson, K., & Morton-Bourgon, K. (2009). The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies. *Psychological Assessment*, 21, 1-21. doi: 10.1037/a0014421
- Hanson, K., & Thornton, D. (2000). Improving risk assessments for sex offenders: A comparison of three actuarial scales. 24, 119-136. doi: 10.1023/A:1005482921333
- Harkins, L., Abracen, J., Looman, J., & Maillet, G. (2011). *Within-treatment change on dynamic predictors of sexual offense recidivism*. Retrieved from Correctional Service of Canada: www.csc-scc.gc.ca/research/005008-0240-eng.shtml
- Harris, A., & Hanson, K. (2010). Clinical, actuarial and dynamic risk assessment of sexual offenders: Why do things keep changing? *Journal of Sexual Aggression*, 16, 296-310. doi: 10.1080/13552600.2010.494772
- Harris, G., & Rice, M. (2003). Actuarial assessment of risk among sex offenders. *Annals of the New York Academy of Sciences*, 989, 198-210. doi: 10.1111/j.1749-6632.2003.tb07306.x
- Hart, S., Kropp, P., Laws, D., Klaver, J., Logan, C., & Watt, K. (2003). The Risk for Sexual Violence Protocol (RSVP). Burnaby, BC: Simon Fraser University, Mental Health, Law, and Policy Institute.
- Haywood, T., Grossman, L., Kravitz, H., & Wasyliv, O. (1994). Profiling psychological distortion in alleged child molesters. *Psychological Reports*, 75, 915-927. doi: 10.2466/pr0.1994.75.2.915

- Hébert, M., Langevin, R., & Daigneault, I. (2016). The association between peer victimisation, PTSD, and dissociation in child victims of sexual abuse. *Journal of Affective Disorders, 193*, 227-232. doi: 10.1016/j.jad.2015.12.080
- Heffernan, R., & Ward, T. (2015). The conceptualization of dynamic risk factors in child sex offenders: An agency model. *Aggression and Violent Behavior, 24*, 250-260. doi: 10.1016/j.avb.2015.07.001
- Heil, P., Harrison, L., English, K., & Ahlmeyer, S. (2009). Is prison sexual offending indicative of community risk? *Criminal Justice and Behavior, 36*, 892-908. doi: 10.1177/0093854809338989
- Helmets, K., Krantz, D., Merz, C., Klein, J., Kop, W., Gottdiener, J., & Rozanski, A. (1995). Defensive hostility: Relationship to multiple markers of cardiac ischemia in patients with coronary disease. *Health Psychology, 14*, 202-209. doi: 10.1037/0278-6133.14.3.202
- Hildebrand, M., & de Ruiter, C. (2012). Psychopathic traits and change on indicators of dynamic risk factors during inpatient forensic psychiatric treatment. *International Journal of Law and Psychiatry, 35*, 276-288. doi: 10.1016/j.ijlp.2012.04.001
- Hoberman, H. (2016). Forensic psychotherapy for sexual offenders: Has its effectiveness yet been demonstrated? In A. Phenix & H. Hoberman (Eds.), *Sexual offending: Predisposing antecedents, assessments and management* (pp. 605-666). New York: Springer Science+Business Media.
- Hoberman, H., & Jackson, R. (2016). Forensic evaluations of sexual offenders: Principles and practices for almost all sexual offender appraisals. In A. Phenix & H. Hoberman (Eds.), *Sexual offending: Predisposing antecedents, assessments and management* (pp. 353-395). New York: Springer Science+Business Media.
- Honor, G. (2010). Child sexual abuse: Consequences and implications. *Journal of Pediatric Health Care, 24*, 358-364. doi: 10.1016/j.pedhc.2009.07.003
- Hosmer, D., & Lemeshow, S. (2000). *Applied Logistic Regression* (Second ed.). Hoboken, NJ, USA: John Wiley & Sons.
- Hudson, S., Wales, D., Bakker, L., & Ward, T. (2002). Dynamic risk factors: The Kia Marama evaluation. *Sexual Abuse: A Journal of Research and Treatment, 14*, 101-117. doi: 10.1177/107906320201400203
- Jacobson, N., Follette, W., & Revenstorf, D. (1986). Toward a standard definition of clinically significant change. *Behavior Therapy, 17*, 308-311. Retrieved from <https://www.journals.elsevier.com/behavior-therapy/>

- Jacobson, N., Roberts, L., Berns, S., & McGlinchey, J. (1999). Methods for determining the clinical significance of treatment effects: Description, application and alternatives. *Journal of Consulting and Clinical Psychology*, 67, 300-307. doi: 10.1037/0022-006X.67.3.300
- Jacobson, N., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19. doi: 10.1037/0022-006X.59.1.12
- Jones, L. (1997). Developing models for managing treatment integrity and efficacy in a prison based TC: The Max Glatt Centre. In E. Cullen, L. Jones & R. Woodward (Eds.), *Therapeutic Communities for Offenders* (pp. 121-157). Chichester: Wiley.
- Jones, L. (2010). History of the offence paralleling behaviour construct and related concepts. In M. Daffern, L. Jones & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 3-23). Chichester: John Wiley and Sons Ltd.
- Jones, L., Daffern, M., & Shine, J. (2010). Summary and future directions. In M. Daffern, L. Jones, & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 317-329). Chichester: John Wiley & Sons, Ltd.
- Jones, N., Brown, S., Robinson, D., & Frey, D. (2015). Incorporating strengths into quantitative assessments of criminal risk for adult offenders: The service planning instrument. *Criminal Justice and Behavior*, 42, 321-338. doi: 10.1177/0093854814547041
- Kim, B., Benekos, P., & Merlo, A. (2016). Sex offender recidivism revisited: Review of recent meta-analyses on the effects of sex offender treatment. *Trauma, Violence, & Abuse*, 17, 105-117. doi: 10.1177/1524838014566719
- King, L., & Roberts, J. (2017). The complexity of public attitudes toward sex crimes. *Victims & Violent Offenders*, 12, 71-89. doi: 10.1080/15564886.2015.1005266
- Klepfisz, G., Daffern, M., & Day, A. (2016). Understanding dynamic risk factors for violence. *Psychology, Crime & Law*, 22, 124-137. doi: 10.1080/1068316X.2015.1109091
- Klepfisz, G., Daffern, M., & Day, A. (2017). Understanding protective factors for violent reoffending in adults. *Aggression and Violent Behavior*, 32, 80-87. doi: 10.1016/j.avb.2016.12.001

- Langevin, R. (1991). A note of the problem of response set in measuring cognitive distortions. *Annals of Sex Research*, 4, 287-292. doi: 10.1177/107906329100400404
- Lanyon, R., Dannenbaum, S., & Brown, A. (1991). Detection of deliberate denial in child abusers. *Journal of Interpersonal Violence*, 6, 301-309. doi: 10.1177/088626091006003003
- Lawson, J., Marshall, W., & McGrath, P. (1979). The Social Self-Esteem Inventory. *Educational and Psychological Measurement*, 39, 803-811. doi: 10.1177/001316447903900413
- Levenson, J., Brannon, Y., Fortney, T., & Baker, J. (2007). Public perceptions about sex offenders and community protection policies. *Analyses of Social Issues and Public Policy*, 7, 137-161. doi: 10.1111/j.1530-2415.2007.00119.x
- Lewis, K., Olver, M., & Wong, S. (2013). The Violence Risk Scale: predictive validity and linking changes in risk with violent recidivism in a sample of high-risk offenders with psychopathic traits. *Assessment*, 20, 150-164. doi: 10.1177/1073191112441242
- Liamputtong, P. (2012). *Qualitative Research Methods* (Fourth ed.). Australia: Oxford University Press Australia.
- Little, R. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198-1202. Retrieved from <https://www.jstor.org/journal/jamerstatasso>
- Lloyd, C. (2007). *Perceptions of self and desistance: Investigating positive attributes associated with exiting crime*. (Masters thesis), Carleton University, Ottawa, Ontario, Canada, Unpublished thesis.
- Loeber, R., Pardini, D., Stouthamer-Loeber, M., & Raine, A. (2007). Do cognitive, physiological, and psychosocial risk and promotive factors predict desistance from delinquency in males? *Development and Psychopathology*, 19, 867-887. doi: 10.1017/S0954579407000429
- Loo, R., & Thorpe, K. (2000). Confirmatory factor analyses of the full and short versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Social Psychology*, 140, 628-635. doi: 10.1111/j.1559-1816.2004.tb01980.x
- Looman, J., Abracen, J., DiFazio, R., & Maillet, G. (2004). Alcohol and drug abuse among sexual and nonsexual offenders: Relationship to intimacy deficits and

- coping strategy. *Sexual Abuse: A Journal of Research and Treatment*, 16, 177-189. doi: 10.1177/107906320401600301
- Lösel, F., & Farrington, D. (2012). Direct protective and buffering protective factors in the development of youth violence. *American Journal of Preventive Medicine*, 43, S8-S23. doi: 10.1016/j.amepre.2012.04.029
- Lösel, F., & Schmucker, M. (2005). The effectiveness of treatment for sexual offenders: A comprehensive meta-analysis. *Journal of Experimental Criminology*, 1, 117-146. doi: 10.1007/s11292-004-6466-7
- Lösel, F., & Schmucker, M. (2017). Treatment of sex offenders: Concepts and empirical evaluations. In T. Sanders (Ed.), *The Oxford handbook of sex offences and sex offenders*. New York: Oxford University Press.
- MacDonald, I. (2016). *Protective factors in the sexual offending area: Analysis of the concept and a preliminary mode* (Unpublished master's thesis). Victoria University of Wellington, New Zealand.
- MacLaren-Chorney, J., McMurtry, C., Chambers, C., & Bakeman, R. (2014). Developing and modifying behavioral coding schemes in pediatric psychology: A practical guide. *Journal of Pediatric Psychology*, 40, 154-164. doi: 10.1093/jpepsy/jsu099
- Makarios, M., & Latessa, E. (2013). Developing a risk and needs assessment instrument for prison inmates: The issue of outcome. *Criminal Justice and Behavior*, 40, 1-23. doi: 10.1177/0093854813496240
- Manikandan, S. (2010). Data transformation. *Journal of Pharmacology & Pharmacotherapeutics*, 1, 126-127. doi: 10.4103/0976-500X.72373
- Mann, R., Hanson, R., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, 22, 191-217. doi: 10.1177/1079063210366039
- Mann, S., & James, G. (1998). Defensiveness and essential hypertension. *Journal of Psychomatic Research*, 45, 139-148. doi: 10.1016/S0022-3999(97)00293-6
- Mann, R., Thornton, D., Wakama, S., Dyson, M., & Atkinson, D. (2010). Applying the concept of offence paralleling behaviour to sex offender assessment in secure settings. In M. Daffern, L. Jones & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 121-135). Chichester: John Wiley and Sons Ltd.

- Mann, R., Webster, S., Schofield, C., & Marshall, L. (2004). Approach versus avoidance goals in relapse prevention with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 16, 65-75. doi: 10.1177/107906320401600105
- Markovic, M. (2006). Analyzing Qualitative Data: Health Care Experiences of Women with Gynecological Cancer. *Field Methods*, 18, 413-429. doi: 10.1177/1525822X06293126
- Marques, J., Wiederanders, M., Day, D., Nelson, C., & van Ommeren, A. (2005). Effects of a relapse prevention program on sexual recidivism: Final results from California's Sex Offender Treatment and Evaluation Project (SOTEP). *Sexual Abuse: A Journal of Research and Treatment*, 17, 79-107. doi: 10.1177/107906320501700108
- Marshall, W., Anderson, D., & Champagne, F. (1997). Self-esteem and its relationship to sexual offending. *Psychology, Crime & Law*, 3, 161-186. doi: 10.1080/10683169708410811
- Marshall, W., Champagne, F., Brown, C., & Miller, S. (1998). Empathy, intimacy, loneliness, and self-esteem in nonfamilial child molesters: A brief report. *Journal of Child Sexual Abuse*, 6, 87-98. doi: 10.1300/J070v06n03_06
- Marshall, W., Champagne, F., Sturgeon, C., & Bryce, P. (1997). Increasing the self-esteem of child molesters. *Sexual Abuse: A Journal of Research and Treatment*, 9, 321-333. doi: 10.1177/107906329700900405
- Marshall, W., Jones, R., Ward, T., Johnston, P., & Barbaree, H. (1991). Treatment outcome with sex offenders. *Clinical Psychology Review*, 11, 465-485. doi: 10.1016/0272-7358(91)90119-F
- Marshall, W., Marshall, L., Sachdev, S., & Kruger, R.-L. (2003). Distorted attitudes and perceptions, and their relationship with self-esteem and coping in child molesters. *Sexual Abuse: A Journal of Research and Treatment*, 15, 171-181. doi: 10.1177/107906320301500302
- Mews, A., Di Bella, L., & Purver, M. (2017). Impact evaluation of the prison-based Core Sex Offender Treatment Programme *Ministry of Justice Analytical Series*: Ministry of Justice.
- McCrae, R., Kurtz, J., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, 15, 28-50. doi: 10.1177/1088868310366253

- McDonald, J. (2014). *Handbook of Biological Statistics* (3rd ed.). Baltimore, Maryland: Sparky House Publishing.
- McDougall, C., Clarke, D., & Woodward, R. (1995). Application of operational psychology to assessment of inmates. *Psychology, Crime and Law*, 2, 85-99. doi: 10.1080/10683169508409768
- McDougall, C., Pearson, D., Willoughby, H., & Bowles, R. (2013). Evaluation of the ADViSOR project: Cross-situational behaviour monitoring of high-risk offenders in prison and the community. *Legal and Criminological Psychology*, 18, 205-228. doi: 10.1111/j.2044-8333.2012.02059.x
- Miller, R., & Lefcourt, H. (1982). The assessment of social intimacy. *Journal of Personality Assessment*, 46, 514-518. doi: 10.1207/s15327752jpa4605_12
- Mills, J., & Kroner, D. (2005). An investigation into the relationship between socially desirable responding and offender self-report. *Psychological Services*, 2, 70-80. doi: 10.1037/1541-1559.2.1.70
- Miller, J., & Maloney, C. (2013). Practitioner compliance with risk/needs assessment tools: A theoretical and empirical assessment. *Criminal Justice and Behavior*, 40, 716-736. doi: 10.1177/0093854812468883
- Mills, J., Anderson, D., & Kroner, D. (2001). The antisocial attitudes and associates of sex offenders. *Criminal Behaviour and Mental Health*, 14, 134-145. doi: 10.1002/cbm.578
- Ministry of Justice. (2017). *Criminal Justice Statistics quarterly, England and Wales, 2016 (final)*. Retrieved from <https://www.gov.uk/government/statistics/criminal-justice-system-statistics-quarterly-december-2016>
- Monahan, J., & Skeem, J. (2016). Risk assessment in criminal sentencing. *Annual Review of Clinical Psychology*, 12, 489-513. doi: 10.1146/annurev-clinpsy-021815-092945
- Mooney, J., & Daffern, M. (2011). Institutional aggression as a predictor of violent recidivism: Implications for parole decision making. *International Journal of Forensic Mental Health*, 10, 52-63. doi: 10.1080/14999013.2010.550984
- Mooney, J., & Daffern, M. (2013). The Offence Analogue and Offence Reduction Behaviour Rating Guide as a Supplement to Violence Risk Assessment in Incarcerated Offenders. *International Journal of Forensic Mental Health*, 12, 255-264. doi: 10.1080/14999013.2013.867421

- Mooney, J., & Daffern, M. (2015). The relationship between aggressive behaviour in prison and violent offending following release. *Psychology, Crime & Law*, 21, 314-329. doi: 10.1080/1068316X.2014.989163
- National Sexual Violence Resource Center (2010). Impact of Sexual Violence Fact Sheet. Retrieved from <https://www.nsvrc.org/publications/NISVS-2010-summary-report>
- New South Wales Bureau of Crime Statistics and Research (2015). *The progress of sexual offences through the NSW Criminal Justice System*. Retrieved from <http://www.bocsar.nsw.gov.au/Documents/sexualattributiondiagrams.pdf>
- New South Wales Bureau of Crime Statistics and Research (2017). *NSW Recorded crime statistics quarterly update June 2017*. Retrieved from http://www.bocsar.nsw.gov.au/Documents/RCS-Quarterly/NSW_Recorded_Crime_June_2017.pdf
- Nunes, K., Babchishin, K., & Cortoni, F. (2011). Measuring treatment change in sex offenders: Clinical and statistical significance. *Criminal Justice and Behavior*, 38, 157-173. doi: 10.1177/0093854810391054
- Nunes, K., Petterson, C., Hermann, C., Looman, J., & Spape, J. (2016). Does change on the MOLEST and RAPE scales predict sexual recidivism? *Sexual Abuse: A Journal of Research and Treatment*, 28, 427-447. doi: 10.1177/1079063214540725
- O'Donohue, W., Letourneau, E., & Dowling, H. (1997). Development and preliminary validation of Paraphilic Sexual Fantasy Questionnaire. *Sexual Abuse: A Journal of Research and Treatment*, 9, 167-178. doi: 10.1007/BF02675062
- Office for National Statistics. (2018). *Crime in England and Wales: year ending September 2017*. Retrieved from <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingseptember2017>
- Ogloff, J., Cutajar, M., Mann, E., & Mullen, P. (2012). *Child sexual abuse and subsequent offending and victimisation: A 45 year follow-up study* (Report no. 440). Canberra: Australian Institute of Criminology. Retrieved from <https://aic.gov.au/publications/tandi/tandi440>
- Ogloff, J., & Davis, M. (2004). Advances in offender assessment and rehabilitation: Contributions of the risk-needs-responsivity approach. *Psychology, Crime and Law*, 10, 229-242. doi: 10.1080/10683160410001662735

- Olver, M., Christofferson, S., Grace, R., & Wong, S. (2014). Incorporating change information into sexual offender risk assessments using the Violence Risk Scale-Sexual Offender version. *Sexual Abuse: A Journal of Research and Treatment*, 26, 472-499. doi: 10.1177/1079063213502679
- Olver, M., Gordon, A., & Wong, S. (2017). *Offence analogue and offence replacement behavior (OAB/ORB) rating guide for Violence Risk Scale-Sexual Offence version*. Unpublished user manual.
- Olver, M., & Wong, S. (2013). Treatment programs for high risk sexual offenders: Program and offender characteristics, attrition, treatment change and recidivism. *Aggression and Violent Behavior*, 18, 579-591. doi: 10.1016/j.avb.2013.06.002
- Olver, M., & Wong, S. (2016). Assessing sexual violence risk and evaluating change with the Violence Risk Scale-Sexual Offender version. In A. Phenix & H. Hoberman (Eds.), *Sexual Offending: Predisposing antecedents, assessments and management* (pp. 547-569). New York: Springer Science+Business Media.
- Olver, M., Wong, S., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale-Sexual Offender version: Assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment*, 19, 318-329. doi: 10.1037/1040-3590.19.3.318
- O'Shea, L., Picchioni, M., & Dickers, G. (2016). The predictive validity of the short-term assessment of risk and treatability (START) for multiple adverse outcomes in a secure psychiatric inpatient setting. *Assessment*, 23, 150-162. doi: 10.1177/1073191115573301
- Papalia, N., Luebbers, S., Ogloff, J., Cutajar, M., & Mullen, P. (2017). The long-term co-occurrence of psychiatric illness and behavioural problems following child sexual abuse. *Australian & New Zealand Journal of Psychiatry*, 51, 604-613. doi: 10.1177/0004867416667232
- Pearson, D., & McDougall, C. (2017). Serious Offenders: Using Evidence to Predict and Manage the Risk. *Behavioral Sciences & the Law*, 35, 239-252. doi: 10.1002/bsl.2288
- Polaschek, D. (2017). Protective factors, correctional treatment and desistance. *Aggression and Violent Behavior*, 32, 64-70. doi: 10.1016/j.avb.2016.12.005
- Pope, C., Ziebland, S., & Mays, N. (2000). Analysing Qualitative Data. *British Medical Journal*, 320, 114-116. doi: 10.1136/bmj.320.7227.114

- Prochaska, J., DiClemente, C., & Norcross, J. (1992). In search of how people change: Applications to addictive behaviors. *American Psychologist*, 47, 1102-1114. doi: 10.1037/0003-066X.47.9.1102
- QSR NVivo (Version 10) [Computer software]. Melbourne, Australia: QSR International. Retrieved from <http://www.qsrinternational.com/nvivo/nvivo-products>
- Quinsey, V., Harris, G., Rice, M., & Lalumière, M. (1993). Assessing the treatment efficacy in outcome studies of sex offenders. *Journal of Interpersonal Violence*, 8, 512-523. doi: 10.1177/088626093008004006
- Quinsey, V., Khanna, A., & Malcolm, P. (1998). A Retrospective Evaluation of the Regional Treatment Centre Sex Offender Treatment Program. *Journal of Interpersonal Violence*, 13, 621-644. doi: 10.1177/088626098013005005
- Rape, Abuse and Incest National Network. Retrieved from <https://www.rainn.org/articles/adult-survivors-child-sexual-abuse>
- Raudenbush, S., Bryk, A., & Congdon, R. (2011). HLM 7 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International, Inc. Retrieved from <https://www.ssicentral.biz/default.aspx>
- Reynolds, W. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 38, 119-125. doi: 10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I
- Russell, D., Peplau, L., & Cutrona, C. (1980). The Revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology*, 39, 472-480. doi: 10.1037/0022-3514.39.3.472
- Schmucker, M., & Lösel, F. (2015). The effects of sexual offender treatment on recidivism: an international meta-analysis of sound quality evaluations. *Journal of Experimental Criminology*, 11, 597-630. doi: 10.1007/s11292-015-9241-z
- Seager, J., Jellicoe, D., & Dhaliwal, G. (2004). Refusers, dropouts, and completers: Measuring sex offender treatment efficacy. *International Journal of Offender Therapy and Comparative Criminology*, 48, 600-612. doi: 10.1177/0306624X04263885
- Seidman, B., Marshall, W., Hudson, S., & Robertson, P. (1994). An examination of intimacy and loneliness in sex offenders. *Journal of Interpersonal Violence*, 9, 518-534. doi: 10.1177/088626094009004006

- Seifert, K., Boulas, J., Huss, M., & Scalora, M. (2017). Response bias on self-report measures of sexual fantasies among sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, 61, 269-281. doi: 10.1177/0306624X15593748
- Serin, R., Chadwick, N., & Lloyd, C. (2016). Dynamic risk and protective factors. *Psychology, Crime & Law*, 22, 151-170. doi: 10.1080/1068316X.2015.1112013
- Serin, R., Gobeil, R., Lloyd, C., Chadwick, N., Wardrop, K., & Hanby, L. (2016). Using dynamic risk to enhance conditional release decisions in prisoners to improve their outcomes. *Behavioral Sciences & the Law*, 34, 321-336. doi: 10.1002/bsl.2213
- Serin, R., & Lloyd, C. (2009). Examining the process of offender change: The transition to crime desistance. *Psychology, Crime & Law*, 15, 347-364. doi: 10.1080/10683160802261078
- Serin, R., Lloyd, C., Helmus, L., Derkzen, D., & Luong, D. (2013). Does intra-individual change predict offender recidivism? Searching for the Holy Grail in assessing offender change. *Aggression and Violent Behavior*, 18, 32-53. doi: 10.1016/j.avb.2012.09.002
- Serran, G., Moulden, H., Firestone, P., & Marshall, W. (2007). Changes in coping following treatment for child molesters. *Journal of Interpersonal Violence*, 22, 1199-1210. doi: 10.1177/0886260507303733
- Sheldrick, C. (1999). Practitioner review: The assessment and management of risk in adolescents. *Journal of Child Psychology and Psychiatry*, 40, 507-518. doi: 10.1111/1469-7610.00469
- Simourd, D., & Hoge, R. (2000). Criminal psychopathy: A risk-and-need perspective. *Criminal Justice and Behavior*, 27, 256-272. doi: 10.1177/0093854800027002007
- Sowden, J., & Olver, M. (2016). Use of the Violence Risk Scale-Sexual Offender version and the Stable 2007 to assess dynamic sexual violence risk in a sample of treated sexual offenders. *Psychological Assessment*. Advance online publication. doi: 10.1037/pas0000345
- Sowden, J., & Olver, M. (2017). Sexual offender treatment readiness, responsivity, and change: Linkages to treatment completion and recidivism. *Journal of Forensic Nursing*, 13, 97-108. doi: 10.1097/JFN.0000000000000160

- Stevens, C., Tan, L., & Grace, R. (2016). Socially desirable responding and psychometric assessment of dynamic risk in sexual offenders against children. *Psychology, Crime & Law*, 22, 420-434. doi: 10.68316X.2015.1120868
- Stinson, J., Becker, J., & McVay, L. (2017). Treatment progress and behavior following 2 years of inpatient sex offender treatment: A pilot investigation of safe offender strategies. *Sexual Abuse: A Journal of Research and Treatment*, 29, 3-27. doi: 10.1177/1079063215570756
- Strahan, R., & Gerbasi, K. (1972). Short, homogeneous versions of the Marlowe–Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 28, 191-193. doi: 10.1002/1097-4679(197204)28:2<191::AID-JCLP2270280220>3.0.CO;2-G
- Strub, D., Douglas, K., & Nicholls, T. (2014). The Validity of Version 3 of the HCR-20 Violence Risk Assessment Scheme Amongst Offenders and Civil Psychiatric Patients. *International Journal of Forensic Mental Health*, 13, 148-159. doi: 10.1080/14999013.2014.911785
- Sweller, T., Daffern, M., & Warren, N. (2016). Challenges in determining how dynamic risk factors manifest in incarcerated sexual offenders, *Psychiatry, Psychology and Law*, 23, 765-781. doi: 10.1080/13218719.2016.1150144
- Tavakoli, H. (2012). *A dictionary of research methodology and statistics in applied linguistics*. Iran: Rahnama Press.
- Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment*, 14, 139-153. doi: 10.1023/A:1014620214905.
- Thornton, D. (2013). Implications of our developing understanding of risk and protective factors in the treatment of adult male sexual offenders. *International Journal of Behavioral Consultation and Therapy*, 8(3-4), 62-65.
- Thornton, D. (2016). Developing a theory of dynamic risk. *Psychology, Crime & Law*, 22, 138-150. doi: 10.1080/1068316X.2015.1109092
- Tingey, R., Lambert, M., Burlingame, G., & Hansen, N. (1996). Clinically significant change: Practical indicators for evaluating psychotherapy outcome. *Psychotherapy Research*, 6, 144-153. doi: 10.1080/10503309612331331668
- Trulson, C., DeLisi, M., & Marquart, J. (2011). Institutional misconduct, delinquent background, and rearrest frequency among serious and violent delinquent offenders. *Crime and Delinquency*, 57, 709-731. doi: 10.1177/0011128709340224

- Truman, J., & Morgan, R. (2016). *Criminal Victimization, 2015*. Retrieved from U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics: <https://www.bjs.gov>
- van den Berg, J., Smid, W., Schepers, K., Wever, E., van Beek, D., Janssen, E., & Gijs, L. (2017). The predictive properties of dynamic sex offender risk assessment instruments: A meta-analysis. *Psychological Assessment*. Online First Publication.
- Vella-Brodrick, D., & White, V. (1997). Response set of social desirability in relation to the Mental, Physical and Spiritual Well-Being Scale. *Psychological Reports, 81*, 127-130. doi: 10.2466/pr0.1997.81.1.127
- Vess, J. (2008). Risk formulation with sex offenders: Integrating functional analysis and actuarial measures. *The Journal of Behavior Analysis of Offender and Victim Treatment and Prevention, 1*, 29-41. doi: 10.1037/h0100454
- Viljoen, J., Shaffer, C., Gray, A., & Douglas, K. (2017). Are adolescent risk assessment tools sensitive to change? A framework and examination of the SAVRY and the YLS/CMI. *Law and Human Behavior, 41*, 244-257. doi: 10.1037/lhb0000238
- Wagner-Menghin, M. (2005). Binomial Test. In B. Everitt & D. Howell (Eds.), *Encyclopedia of Statistics in Behavioral Science*. New York, NY: John Wiley & Sons, Ltd.
- Wakeling, H., Freemantle, N., Beech, A., & Elliott, I. (2011). Identifying predictors of recidivism in a large sample of United Kingdom sexual offenders: A prognostic model. *Psychological Services, 8*, 307-318. doi: 10.1037/a0025516
- Walters, G. (2006). Risk-appraisal versus self-report in the prediction of criminal justice outcomes. *Criminal Justice and Behavior, 33*, 279-304. doi: 10.1037/a0025516
- Ward, T. (2016). Dynamic risk factors: Scientific kinds or predictive constructs. *Psychology, Crime & Law, 22*, 2-16. doi: 10.1080/1068316X.2015.1109094
- Ward, T. (2017). Prediction and agency: The role of protective factors in correctional rehabilitation and desistance. *Aggression and Violent Behavior, 32*, 19-28. doi: 10.1080/1068316X.2015.1109094
- Ward, T., & Beech, A. (2006). An integrated theory of sexual offending. *Aggression and Violent Behavior, 11*, 44-66. doi:10.1016/j.avb.2005.05.002
- Ward, T., & Beech, A. (2014). Dynamic risk factors: a theoretical dead-end? *Psychology, Crime & Law, 21*, 1-14. doi: 10.1016/j.avb.2016.11.012

- Ward, T., & Brown, M. (2004). The good lives model and conceptual issues in offender rehabilitation. *Psychology, Crime and Law*, 10, 243-257. doi: 10.1080/1068316X.2014.917854
- Ward, T., & Fortune, C.-A. (2016). From dynamic risk factors to causal processes: A methodological framework. *Psychology, Crime & Law*, 22, 190-202. doi: 10.1080/1068316X.2015.1117080
- Ward, T., & Stewart, C. (2003). The treatment of sex offenders: Risk management and good lives. *Professional Psychology: Research and Practice*, 34, 353-360. doi: 10.1037/0735-7028.34.4.353
- Ware, J., Frost, A., & Hoy, A. (2010). A review of the use of therapeutic communities with sexual offenders. *International Journal of Offender Therapy and Comparative Criminology*, 54, 721-742. doi: 10.1177/0306624X09343169
- Ware, J., Mann, R., & Wakeling, H. (2009). Group versus individual treatment: What is the best modality for treating sexual offenders? *Sexual Abuse in Australia and New Zealand*, 1(2), 70-78. Retrieved from <https://www.anzatsa.org>
- Webster, C., Martin, M., Brink, J., Nicholls, T., & Middleton, C. (2004). Short-Term Assessment of Risk and Treatability (START): Clinical guide for evaluation risk and recovery. Ontario, Canada: St. Joseph's Healthcare Hamilton.
- Willis, G., Yates, P., Gannon, T., & Ward, T. (2013). How to integrate the Good Lives Model into treatment programs for sexual offending: An introduction and overview. *Sexual Abuse: A Journal of Research and Treatment*, 25, 123-142. doi: 10.1177/1079063212452618
- Wilson, G. (1978). *The secrets of sexual fantasy*. London: J.M. Dent & Sons.
- Wilson, G. (1988). Measurement of sex fantasy. *Sexual and Marital Therapy*, 3, 45-55. doi: 10.1080/02674658808407692
- Wiseman, J. (2015). *Incidence and Prevalence of Sexual Offending (Part 1)*. Retrieved from National Criminal Justice Reference Service: Office of Justice Programs <https://www.ncjrs.gov/index.html>
- Wong, S., & Gordon, A. (2000). Manual for the Violence Risk Scale. Saskatchewan, Canada: University of Saskatchewan.
- Wong, S., & Gordon, A. (2006). The validity and reliability of the violence risk scale: A treatment-friendly violence risk assessment tool. *Psychology, Public Policy, and Law*, 12, 279-309. doi: 10.1037/1076-8971.12.3.279

- Wong, S., Olver, M., Nicholaichuk, T., & Gordon, A. (2003). The Violence Risk Scale: Sexual Offender Version (VRS-SO). In Regional Psychiatric Centre and University of Saskatchewan, Saskatoon, Saskatchewan, Canada.
- Woodrow, A., & Bright, D. (2011). Effectiveness of a sex offender treatment programme: A risk band analysis. *International Journal of Offender Therapy and Comparative Criminology*, 55, 43-55. doi: 10.1177/0306624X09352162
- Yalom, I. (1985). *The theory and practice of group psychotherapy* (3rd ed.). The University of Michigan: Basic Books.
- Yates, P., & Kingston, D. (2016). The shortcomings of sexual offender treatment: Are we doing something wrong? In R. Laws & W. O'Donohue (Eds.), *Treatment of sex offenders: Strengths and weaknesses in assessment and intervention* (pp. 261-279). Cham, Switzerland: Springer International Publishing.
- Yesberg, J., & Polaschek, D. (2014). Using information from the Violence Risk Scale to understand different patterns of change: An exploratory investigation of intensively treated life-sentenced prisoners. *Journal of Interpersonal Violence*, 29, 1-24. doi: 10.1177/0886260514527824
- Zamble, E. (1992). Behavior and adaptation in long-term prison inmates: Descriptive longitudinal results. *Criminal Justice and Behavior*, 19, 409-425. doi: 10.1177/0093854892019004005
- Zamble, E., & Porporino, F. (1990). Coping, imprisonment and rehabilitation. *Criminal Justice and Behavior*, 17, 53-69. doi: 10.1177/0093854890017001005

Appendices

Appendix A: Published journal article ‘Challenges in determining how dynamic risk factors manifest in incarcerated sexual offenders’

Appendix B: Psychologists’ Survey (Study 1)

Appendix C: Behavioural Checklist (Study 3)

Appendix D: Ethics Approval Notices

Appendix E: Participant Information Sheet and Consent Form (Study 3)

Appendix A: Published Journal Article

Psychiatry, Psychology and Law, 2016
Vol. 23, No. 5, 765–781, <http://dx.doi.org/10.1080/13218719.2016.1150144>



Challenges in Determining How Dynamic Risk Factors Manifest in Incarcerated Sexual Offenders

Tamara Sweller^{a,b}, Michael Daffern^{b,c} and Narelle Warren^d

^aCorrective Services New South Wales, Australia; ^bCentre for Forensic Behavioural Science and Legal Studies, Clifton Hill, Victoria, Australia; ^cVictorian Institute of Forensic Mental Health, Melbourne, Australia; ^dSchool of Psychological Sciences, Monash University, Melbourne, Australia

To determine imprisoned sexual offenders' post-release risk for reoffending and to monitor change during imprisonment, risk-relevant behaviours must be identified and monitored. The aim of this study is to elucidate the behavioural manifestations of empirically-derived dynamic risk factors in incarcerated sexual offenders. A total of 34 psychologists with experience working with sexual offenders in secure settings completed an online survey. They were presented with items drawn from structured risk-assessment instruments – the Risk of Sexual Violence Protocol (RSVP) and the Violence Risk Scale: Sexual Offender version (VRS:SO) – and asked to generate specific behaviours they might observe in custody that are indicative of each of the dynamic risk factors included in the RSVP and the VRS:SO. They were also asked to identify positive variants of each of the dynamic risk factors – pro-social behaviours which may indicate improvement in the offender's behaviour. The participants generated more examples of dynamic risk factors compared with pro-social behaviours. The findings were reflective of the challenges which clinicians and researchers face in the assessment of dynamic risk factors in incarcerated sexual offenders. Practical and theoretical implications of the findings are discussed.

Key words: behavioural observation; change assessment; dynamic risk factors; offence paralleling behaviour; prison; risk assessment; sexual offender.

Risk assessment forms a crucial part of release decision-making and offender management and intervention (Andrews, Bonta, & Wormith, 2006). A difficulty for risk assessors is determining how dynamic risk factors manifest in the restricted and intensely monitored prison environment (Jones, 2010); this difficulty impacts upon the scoring of risk-assessment measures and efforts to measure change. This study seeks to elucidate the behavioural manifestations of dynamic risk for incarcerated sexual offenders. These behaviours may be considered by assessors

when scoring structured risk-assessment instruments, and further, may focus custodial staff members' attention on risk-related behaviours.

Sexual Offender Risk Assessment

Several approaches to risk assessment have been designed to improve the predictive accuracy of an offender's risk level, including both static and dynamic risk factors. Examples of structured risk-assessment tools for sexual offenders that measure both

Correspondence: Michael Daffern, Centre for Forensic Behavioural Science and Legal Studies, 505 Hoddle St, Clifton Hill, Victoria 3068 Australia, Telephone: +61419650917. Email: mdaffern@swin.edu.au

static and dynamic factors include the Risk for Sexual Violence Protocol (RSVP; Hart et al., 2003) and the Violence Risk Scale: Sexual Offender version (VRS:SO; Wong, Olver, Nicholaichuk, & Gordon, 2003). Research exploring these measures is still comparatively underdeveloped, with the result that important questions remain concerning their conceptual foundations, namely whether they target the most relevant risk factors and the extent to which it is possible to associate recidivism rates with specific scores (Mann, Hanson, & Thornton, 2010).

In order to assess risk and needs in a custodial setting appropriately it is important to understand the characteristics and behavioural manifestations of the relevant dynamic risk factors (Simourd & Hoge, 2000). Assessors must also be aware of the consistency with which such behaviours manifest across time and situations (Olver & Wong, 2011).

A central problem for assessments in this setting is that the context of assessment and intervention is significantly different from the context of the initial occurrence of the offending behaviour and the potential recurrence of the behaviour (Jones, 2010); for instance, the preferred victim type may not be available to the offender in prison (McDougall, Clarke, & Woodward, 1995). Additionally, antisocial behaviours that are observable in the community may manifest in different ways in custodial settings (Gordon & Wong, 2010). According to Daffern (2010), an additional consideration is that the prison environment might trigger antisocial behaviour that appears to be a manifestation of risk but is in fact a product of the prison environment (e.g. anti-authoritarian attitudes and interpersonal violence may be highly valued in some prison settings). As a result, assessors must be mindful that they do not misinterpret behaviours as indicators of risk (Mann, Thornton, Wakama, Dyson, & Atkinson, 2010).

Furthermore, some manifestations of dynamic risk factors are more easily observed

by staff than other behaviours. Therefore, there might be a greater rate of observation of these factors (e.g. self-regulation deficits; Hanson & Harris, 2001). Other factors may be more difficult to observe; one risk factor that might be particularly difficult to measure through staff observations is sexual deviance. Such measurement difficulties might decrease clinicians' ability to assess the presence or relevance of behavioural manifestations of certain risk factors (Webster, Müller-Isberner, & Fransson, 2002). It can also be difficult to identify and measure the extent to which attitudes or beliefs are present. For sexual offenders, the presence of offence-supportive attitudes is often inferred from the statements offenders make about their offending (Mann et al., 2010). Although behaviour is generally an outcome of related thoughts, offenders may make proclamations that are contrary to their beliefs due to a wish to impress in a socially desirable manner and they may behave in a manner that differs from these proclamations (Mann et al., 2010). Inferring an offender's attitudes might also allow a greater degree of subjectivity and bias, which likely impacts the reliability and validity of these observations. An additional complexity is that some ratings criteria are based on behaviour that is more likely to manifest in the community; for instance, it might be difficult to assess sexual compulsivity when offenders have limited opportunity to engage in promiscuous behaviour in prison (Olver, Wong, Nicholaichuk, & Gordon, 2007).

Offence Paralleling Behaviour (OPB)

Jones (1997) conceptualised a methodology for identifying behaviours that manifest while offenders are in custody that relate to offence processes, referring to these behavioural sequences as offence paralleling behaviour (OPB). More recently, OPB has been defined as 'a behavioural sequence incorporating overt behaviours (that may be muted by environmental factors), appraisals, expectations,

beliefs, affects, goals and behavioural scripts, all of which may be influenced by the patient's mental disorder, that is functionally similar to behavioural sequences involved in previous criminal acts' (p. 267; Daffern et al., 2007). It should be noted that OPB refers to both overt behaviour and internal processes such as attitudes and beliefs, all of which can be manifestations of dynamic risk factors.

Similarly, Gordon and Wong (2010) use the term offence analogue behaviours (OABs) to represent behaviours in a custodial setting that indicate manifestations of dynamic risk factors. In general, these OABs are idiosyncratic to the individual, linked to their criminogenic needs, and result from an interaction between these criminogenic needs and the immediate environment (Gordon & Wong, 2010). Validation research on the *Offence Analogue and Offence Reduction Behaviour Rating Guide* (Gordon & Wong, 2009) has shown that behavioural manifestations of dynamic risk factors in incarcerated violent offenders are identifiable in a custodial context and that some are related to criminal recidivism (Mooney & Daffern, 2013). Mooney and Daffern (2013) highlighted the importance of such a behaviour guide to increase staff awareness of behaviours that might be indicative of reduced or ongoing risk of recidivism.

Of relevance to the OPB framework is Andrews, Bonta, and Wormith's (2011) assertion that alternative ways of thinking, feeling and acting can counteract an individual's risk factors. Rather than focusing only on manifestations of dynamic risk, the development and measurement of positive behaviour through intervention is similarly important in determining an offender's current risk level (Rogers, 2000). These pro-social behaviours have been termed pro-social alternative behaviours (PABs) by Daffern et al. (2007), and offence reduction behaviours (ORBs) by Gordon and Wong (2010). A decrease in risk level is not only measured by a decrease in OPB but also by an increase in PAB (Daffern et al., 2007). Using the OAB/ORB rating

guide, Mooney and Daffern (2013) demonstrated that pro-social skills (ORBs) were of greater predictive value for violent recidivism than negative behaviours (i.e., OAB).

The assessment of OPB and PAB – and OAB and ORB – involves conducting systematic assessments to reduce subjectivity, observer bias and reliance on accidental observations (Daffern et al., 2007). In a custodial environment, custodial staff have the capacity to make important behavioural observations that contribute to the risk-assessment process. However, they are not ordinarily trained in this complex assessment task. As a result, it may be useful for staff to focus their attention on a list of behaviours that more broadly indicate the presence of common dynamic risk factors (McDougall et al., 1995). The first step in this process is to identify these behavioural manifestations.

The Current Study

Although research has been conducted in relation to the manifestation of violent offenders' criminogenic needs in custody (Mooney & Daffern, 2013), there has been minimal research of a similar nature in relation to the behaviour of incarcerated sexual offenders and the manifestation of dynamic risk factors in custody. The practice of risk assessment based on monitoring behavioural manifestations of dynamic risk factors in custody is therefore limited. In this study, psychologists with experience working with incarcerated sexual offenders were asked how risk factors derived from established structured risk-assessment measures manifest in custody.

Method

Research Design

This qualitative study involved a structured survey completed by psychologists with experience working with sexual offenders. The survey comprised open-ended questions in order to capture the diversity of responses.

Survey responses were analysed using thematic analysis, a process by which qualitative data can be examined and organised into themes (Braun & Clarke, 2013).

Participants

In total, 120 participants commenced the survey. Of these participants, 86 were classified as 'partial completers' because they completed less than 60% of the risk factor questions. These participants were excluded from the data analysis. As a result, 34 participants were included in the sample. All participants were members of Australian and North American organisations or associations that require employees or members to have a minimum level of experience or to be registered as a psychologist. Participants were contacted indirectly by email through their affiliation with national and international professional organisations, including Corrective Services New South Wales, Corrections Victoria, the Australian Psychological Society, and the Association for the Treatment of Sexual Abusers. The introductory email outlined the purpose of the study and the inclusion criteria (i.e. those with experience working in a custodial setting with offenders convicted of a sexual offence). Of the 34 participants, 30 were working with sexual offenders at the time they completed the survey. There were 10 participants with over 10 years' experience, while the majority had at least 5 years' experience ($n = 27$). The median level of experience was 6 years.

Materials

The survey was developed using an online survey tool, SurveyMethods (<https://www.surveymethods.com/>). It consists of 32 questions. The introductory questions are related to participants' clinical background, i.e. whether they currently work or have worked in the past with offenders convicted of a sexual offence, and the depth of their knowledge of OPB. A brief explanation of OPB is

provided prior to the commencement of further questions. Each of the subsequent questions was based on an empirically derived risk factor predictive of sexual recidivism. These risk factors were drawn from two risk-assessment tools: the RSVP and the VRS:SO. For each risk factor, participants were required to respond to two parts: 1) provide examples of behaviours they might expect to observe in offenders if the risk factors were active and 2) provide examples of behaviours they might expect to observe if these risk factors were no longer active and had been replaced with more adaptive pro-social behaviour.

An example of a survey question based on a risk factor is as follows:

8. Problems with stress or coping (i.e., extent to which the person's psychosocial adjustment is unstable or susceptible to external events and occurrences). What are some examples of behaviours/attitudes you might observe if this risk factor manifests in custody? What are some examples of behaviours/attitudes you might observe if the offender has made positive changes and no longer manifests this risk factor?

Procedure

An introductory email was sent to potential participants. The study was explained and the inclusion criteria were outlined. Participants were informed in the email that by clicking on the link to the survey they were consenting to participate in the study. Participants could withdraw consent at any time until they submitted their responses. The survey was anonymous, so consent could not be withdrawn after completion.

Data Analysis

The survey results were imported into a qualitative data management program, QSR NVivo 10 (<http://www.qsrinternational.com>). A database was created in this program and data were analysed using a six-step thematic

analysis approach commonly used in psychological research (Braun & Clarke, 2006, 2013). Repeated readings of the participants' responses to the open-ended survey questions allowed initial and ongoing *data familiarisation* (Step 1). For each participant's response for each survey question, potential emergent codes were noted. This involved assigning a representative or summative word or phrase (a 'code') to a passage of the data ('*forming codes*'; Step 2). Responses were coded systematically and initial codes were *categorised* into potential themes and sub-themes on the basis of any patterns identified (Step 3). A *theme was identified* for the analysis (Step 4) by noting its frequency and/or its relevance to the research questions. In this way, themes were refined to ensure that each had sufficient and meaningful supporting data. This process was largely inductive, drawing on the responses of participants; however, deductive themes were also developed based on the researcher's knowledge of relevant risk factors and behaviour that sexual offenders tend to exhibit as they progress through a treatment programme (Markovic, 2006). The use of deduction was considered appropriate given the applied purpose of the research (Pope, Ziebland, & Mays, 2000).

Each behaviour provided in the survey responses could be coded into one or more of the themes. The number of themes created for each risk factor was flexible and was not predetermined. Themes were collapsed into a manageable number while maintaining their discriminative validity. Each OPB and PAB for each risk factor comprised of between 8 and 15 themes. Themes were then *defined and named* (Step 5) in the context of the broader project (Braun & Clarke, 2006). At this point, frequencies were obtained for each theme. An inter-rater reliability check was conducted on the data. A researcher with no experience of forensic psychology or sexual offenders' behaviour undertook an independent coding of responses for 10% of the data (Liamputtong, 2012). Where there were discrepancies in the coding between coders, discussion was

undertaken and codes were redefined. This process was used to demonstrate that the findings extended beyond the subjective judgements of the researchers (Pope et al., 2000), ensuring the rigour of the analysis. The outcome reflected reliability within the themes, which are *reported* (Step 6) in this paper (Braun & Clarke, 2006, 2013).

Results

Descriptive Statistics

At the time of the survey, 85% ($n = 29$) of participants were currently working with sexual offenders and 15% ($n = 5$) had worked with sexual offenders previously. Levels of experience working with this group varied from between 1 and 20 years. The majority of participants (82%, $n = 28$) reported that they were familiar with the concept of OPB.

Participants' responses for OPB were generally more descriptive than responses for PAB. For instance, when providing examples of behaviour which might indicate that a risk factor is no longer problematic, there were suggestions of 'opposite of above' or 'absence of above'. The interpretation of these comments would have been subjective; therefore, they were coded into a separate theme.

As can be seen in Tables 1 and 2, there was some overlap between risk factors and examples of behavioural manifestations. That is, several behavioural manifestations were provided that were also risk factors. For example, 'problems with substance abuse' was itself a risk factor but was also described as a manifestation of other risk factors (e.g. 'problems with stress or coping'). There were also behavioural manifestations that were suggested for several risk factors (e.g. 'sexual comments to staff' was suggested for risk factors including 'sexual compulsivity' and 'attitudes that condone sexual violence').

Participants provided a greater number of relevant behavioural manifestations for some risk factors than for others. For instance, the overall number of predicted behavioural

Table 1. Risk factors and themes of OPB examples, with examples of behaviours within themes.

<i>Risk factor</i>	<i>Number of different themes of behaviour identified</i>	<i>Most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Second most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Third most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>
Attitudes that support sexual violence	10	Lack of victim empathy, victim blaming = Not accepting responsibility = Broad or questionable relevance	Joking about victims, victim blaming = Justifying behaviour = Poor attitude	Supporting others' unhelpful attitudes	Normalisation of sexual violence, colluding with others	Sexualised or sexually aggressive comments	Sexualised comments to staff
Problems related to child abuse	14	Fear or mistrust	Hyper-vigilant, refusal to discuss childhood	Attachment issues	Poor boundaries with others, keeps physical distance	Poor coping	Isolation, substance abuse, feelings of worthlessness
Problems with stress or coping	13	Physical or verbal aggression	Loss of temper, angry outbursts, irritability in interactions, fighting	Poor problem-solving	Failing to keep appointments, catastrophic thinking	Broad or questionable relevance	Deterioration in behavioural case notes, control issues
Problems with substance abuse	15	Substance use	Glazed eyes, positive urinalysis	Drug-seeking behaviour	Telling the clinic they will self-harm if they do not get medication	Emotional dysregulation = Unhelpful attitude to substance use	Calm one day and aggressive the next = Psychological arousal from discussions about drugs
Problems with treatment	10	Disengagement	Sabotages process, refusal to complete treatment work	Not taking responsibility for treatment	Focusing on others rather than self	Therapy-interfering behaviour	Oppositional attitude, disruptive, distraction to avoid topics
Problems with planning	11	Poor release planning	Unrealistic pre-release plans	Poor problem-solving	Impulsivity, leaving requests to the last minute	Difficulty engaging in routine	Late for muster, not doing homework

(continued)

Table 1. (Continued)

<i>Risk factor</i>	<i>Number of different themes of behaviour identified</i>	<i>Most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Second most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Third most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>
Problems with supervision	9	Not abiding by rules	Institutional misconduct, fighting against rules	Lack of regard for authority and system	Lies to staff, hostility towards staff	Not adhering to staff requests	Refusing orders
Problems with employment	9	Refusing to work or avoiding = Unable to maintain a job	Calling in sick, refusing to do training = Frequently loses job in custody	Lack of motivation	No goals, loses interest	Not taking responsibility	Sense of entitlement, feels owed by system or government
Problems with intimate relationships	10	Communication difficulties	Aggressive phones calls with partner	Poor attachments	Superficial connections, denying the need for others	Avoidance	Isolation, withdrawal from others
Problems with non-intimate relationships	12	Withdrawal or isolation	Avoidance of supports	Difficulties forming and maintaining relationships	Not communicating with more than one other inmate	Aggression	Interpersonal conflict, fighting
Interpersonal aggression	10	Intimidation	Intensive eye contact	Verbal aggression	Raised voice to express hostility	Physical aggression	Violence
Offence planning	9	Manipulating others	Setting up relationships of obligation	Reporting plans to offend	Expressing offence fantasy	Victim contact	Contacts previous victims, makes connections with others to facilitate offending
Extreme minimisation or denial	13	Not accepting responsibility	Downplaying significance of harm to victims, claiming not to remember	Resistance to treatment	Anger at discussing offence, refusal to participate in treatment	Blame shifting	Blaming others

(continued)

Table 1. (Continued)

<i>Risk factor</i>	<i>Number of different themes identified</i>	<i>Most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Second most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Third most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>
Psychopathic personality disorder	14	Manipulation	Playing staff or inmates off against each other	Superficial	Appears to comply but undertones of non-compliance	Lack of empathy	Delight in others' discomfort in group
Mental illness	15	Psychotic symptoms	Talking to oneself	Broad or questionable relevance	Medical records, acting crazy	Mood states	Excessive crying
Non-sexual criminality	12	Institutional misconduct	Theft, rule-breaking	Pro-criminal attitudes	Joking about criminality	Aggression	Fights, stand overs
Physical coercion	11	Physical or verbal aggression	Direct threats to safety	Sexual aggression	Preying on younger inmates	Broad or questionable relevance	Creepy behaviour
Psychological coercion	7	Manipulating for favours	Buying others food in return for going two-out	Stand overs and intimidation	Social shunning, stand overs	Broad or questionable relevance	Sexual offences in custody, poor impulse control
Problems with self-awareness	11	Broad or questionable relevance	Use of drugs or alcohol, poor coping	Not accepting responsibility	Denial that behaviour requires addressing	Not understanding triggers	Talking about returning to old lifestyle
Sexual compulsivity	10	Sexualised comments	Sexual content in conversations, sexual comments to staff	Sexual behaviour in public	Being caught masturbating, talking and acting in sexually explicit ways	Excessive sexual behaviour	Daily masturbation, sex with other prisoners
Sexual deviance	12	Broad or questionable relevance	Watching sexual TV, excessive masturbation	Inappropriate images	Collecting pictures from children's clothing catalogues	Inappropriate sexual arousal	Watching afternoon children's TV programmes
Violent or suicidal ideation	9	Threats or preparation	Weapon-creating	Emotional instability	Irritability, sudden change in presentation, tearfulness	No future focus = Broad or questionable relevance	Giving away possessions, aggressive acting out without regard for personal safety = Periods in segregated custody

Note: '=' is used to denote that multiple themes have the same frequency of responses.

Table 2. Risk factors and themes of PAB examples, with examples of behaviours within themes.

<i>Risk factor</i>	<i>Number of different themes of behaviour identified</i>	<i>Most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Second most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Third most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>
Attitudes that support sexual violence	11	Broad or questionable relevance	Using helpful coping strategies, maturity	Challenging unhelpful attitudes	Challenging others about sexual offending, encouraging others to participate in treatment	Victim empathy	Recognition of harm to the victim
Problems related to child abuse	14	Appropriate attachments	Understanding and respect for personal boundaries	Greater self-awareness	Identifying how abuse history contributed to misbehaviour	Engaging in intervention = Good coping	Willingness to address issues = Discusses emotional reactions to stimuli
Problems with stress or coping	11	Using supports	Seeking out assistance from others	Broad or questionable relevance = Good problem-solving	Coping skills, consistency in behavioural case notes = Positive attitude to challenges, calmer	Appropriate use of services = Manage unhelpful thoughts	Keeping appointments = Withdraws from anger-provoking situations, perspective taking
Problems with substance abuse	11	Engaging in intervention = Negative urinalysis	Compliance with methadone programme	Helpful attitudes about drug use	Discusses problems caused in the past	Risk management	Strategies for managing cravings
Problems with treatment	9	Engagement	Listens and contributes	Takes responsibility for treatment	Addresses own needs, takes advice, active participation	Discusses relevant issues in depth	Gives answers based on own circumstances, volunteers information
Problems with planning	12	Release planning	Communication with outside supports	Good problem-solving	Identify thought pattern before behaviour	Time management	Planning activities, taking control of own timetable
Problems with supervision	9	Compliance with rules	Complies with directions from staff	Engagement	Reliable with tasks, attendance at case meetings	Positive attitudes to supervision	Post-release plans involve continued treatment and supervision

(continued)

Table 2. (Continued)

Risk factor	Number of different themes of behaviour identified		Most frequently reported themes				Second most frequently reported themes		Example(s) of behaviour in the theme		Third most frequently reported themes		Example(s) of behaviour in the theme	
	Number of different themes of behaviour identified		Most frequently reported themes				Second most frequently reported themes		Example(s) of behaviour in the theme		Third most frequently reported themes		Example(s) of behaviour in the theme	
Problems with employment	10		Regular work attendance		Stable employment, establishes and maintains engagement		Positive feedback from work		Good work reports		Positive attitude to work = Seeking training		Pride in job = Active participation in vocational training and education	
Problems with intimate relationships	10		Taking responsibility in relationships		Making contact with loved ones, engagement of partner in release planning		Positive interactions and relationships = Stability in relationships		Positive interactions with family = Desire to maintain relationships		Broad or questionable relevance = Valuing relationships		Pattern of behaviour over time, reflection = creating an equal partnership	
Problems with non-intimate relationships	9		Interacting with others		Mixing with others, working in a team		Stable and meaningful relationships		Developing and maintaining relationships		Assertive communication		Assertive communication with staff	
Interpersonal aggression	11		Appropriate interactions		Calm and able to tolerate disagreement, listens calmly		Assertive communication = Good coping		Open, discussing concerns respectfully = Speaks about not getting something that was wanted in a non-threatening way, self-calming		Absence of behaviour		Reduction or absence of aggression, no fights or assaults	
Offence planning	9		Accepts responsibility for offending		Acknowledges intent, implements relapse strategies		Appropriate interactions and relationships		Encouraging peers to be compliant with rules		Risk management		Challenging distortions, accepts that urges may continue but will pass	
Extreme minimisation or denial	11		Accepting responsibility for behaviour		Accepts own behaviour and feedback without defensiveness		Treatment engagement		Challenge of self and others through feedback, honesty		Empathy		Articulates how behaviour might have affected others	
Psychopathic personality disorder	12		Absence of certain behaviours		Not callous or cynical		Appropriate relationships = Empathy =		Maintaining appropriate boundaries = Kindness when not		Self-reflection and insight		Acknowledging emotional shallowness	

(continued)

Table 2. (Continued)

<i>Risk factor</i>	<i>Number of different themes of behaviour identified</i>	<i>Most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Second most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Third most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>
Mental illness	12	Medication compliance	Regularly taking medication	Treatment engagement Stable mental state	demanding = Willingness to debate actions Stable mood, more spontaneous in responding	Treatment engagement Implements management strategies	
Non-sexual criminality	10	Abides by rules = Absence of certain behaviours	Positive self-regulation, following rules = Absence of incidents or accusations	Appropriate future planning = Pro- social attitudes	Accesses services and planning for future without crime = Improved attitudes to law	Broad or questionable relevance More positive case notes	
Physical coercion	15	Absence of behaviour = Respectful interactions	Absence of threats = Engaging respectfully with staff and inmates	Empathy = Insight into problematic behaviour	Efforts made to genuinely help others = Talking about problems, challenging cognitive distortions	Broad or questionable relevance Engaging in work preparation	
Psychological coercion	10	Respectful interactions	Humility and openness	Absence of behaviour triggers	Lack of manipulation	Assertive communication = Broad or questionable relevance	Openness in communication = Non-coercive
Problems with self awareness	12	Accepting responsibility for behaviour = Realistic risk management = Treatment engagement	Removal of self from risky situations, speaks about actions leading up to offence = Scenario-planning = Identification of	Understanding triggers	Verbalising own risk factors	Broad or questionable relevance Good manners, improved coping	

(continued)

Table 2. (Continued)

<i>Risk factor</i>	<i>Number of different themes of behaviour identified</i>	<i>Most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Second most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>	<i>Third most frequently reported themes</i>	<i>Example(s) of behaviour in the theme</i>
Sexual compulsivity	10	Absence of behaviour	No sexually inappropriate behaviour in custody	Control of sexual behaviour	Delayed gratification, sitting with discomfort	Insight into and intervention in behaviour	Paying attention to thoughts, arousal-management techniques
Sexual deviance	9	Treatment engagement	Openness to discussing difficulties in this area	Coping strategies to manage deviance	Open about sexual thoughts, replaces deviant fantasies	Absence of behaviour	Reduction of these behaviours, absence of above
Violent or suicidal ideation	11	Reduction in threats or acts	Fewer attempts, no talk of violence	Engagement in intervention	Compliance with medication, distress tolerance, seeks help	Appropriate engagement with others = Future focused	Assertive communication, cooperative with others = Consequential thinking, talking of hopeful future

Note: '=' is used to denote that multiple themes have the same frequency of responses.

manifestations was higher for the risk factor 'problems with substance abuse' (100). Conversely, although the overall number of predicted behavioural manifestations for the risk factor 'sexual deviance' was 97, of these responses, 21 were coded as 'broad or questionable relevance' (see Table 1). There was also a difference in the quality of the responses, such that for some risk factors the examples provided were more specific and observable behaviours (e.g. problems with stress or coping, problems with substance abuse, interpersonal aggression). Responses that were less specific and observable were commonly reported for risk factors, such as sexual deviance, attitudes that condone sexual violence, and problems with self-awareness. Further, there were several behavioural manifestations that were uncommon and appeared irrelevant to the risk factor in question (e.g. one participant identified a behavioural manifestation of substance use as 'placement issues'). These responses were coded into a separate theme (i.e. 'broad or questionable relevance').

Risk factors related to sexual self-regulation (e.g. sexual compulsivity, sexual deviance) elicited responses that were commonly based on self-report (e.g. excessive masturbation). Observable behaviour was also described; however, this behaviour would likely be observed only in the most extreme cases (e.g. public masturbation). In addition, such extreme behaviour would often be classified as a sexual offence rather than a risk factor.

Discussion

The aim of this study was to identify behavioural manifestations of sexual offenders' dynamic risk factors in a custodial environment and the potential difficulties that might arise within this process. The identification of these behaviours may assist clinicians in assessing risk, monitoring change during lengthy periods of incarceration, and identifying opportunities for timely intervention

(i.e. intervening when a risk-related behaviour is present and active). Furthermore, this survey illuminates the wide range of factors that assessors focus on when appraising the presence of dynamic risk factors embedded in structured risk-assessment instruments.

Based on the responses provided, participants more readily identified manifestations of dynamic risk factors compared to examples of pro-social equivalent behaviour. Responses were generally more common, detailed and specific in relation to the manifestation of dynamic risk factors. The difficulty in identifying behaviour related to positive change might reflect a range of issues: psychologists are trained in the use of risk-assessment tools, which focus on risk-related factors (de Vries Robbé, de Vogel, & de Spa, 2011; Miller, 2006; Rogers, 2000; Sheldrick, 1999); the culture of a custodial environment is based on punishment or consequences for antisocial behaviour; and when clinicians discuss offenders with custodial staff, more often than not the focus is on observed difficulties or problematic behaviour. As a result, staff in this environment may focus on risk-related behaviour rather than positive behaviour. This finding is also consistent with prior research on the *Offence Analogue and Offence Reduction Behaviour Guide*, which has shown that staff rarely document pro-social behaviour (Mooney & Daffern, 2013). This emphasis on risk factors is likely to contribute to 'professional negativity' and could result in a negative bias against offenders (Rogers, 2000). Rather than limiting the assessment of risk to a determination of the absence of problematic behaviour, it is important to consider behaviour improvement. That is, consistent with research showing the benefits of protective factors (e.g. de Vries Robbé et al., 2011), the development of pro-social skills that replace, or reflect improvement in, a dynamic risk domain may also be relevant to the assessment of risk for recidivism (Mooney & Daffern, 2013).

An alternative explanation for participants' greater difficulty in identifying pro-social behavioural manifestations might be the wording of the survey questions in this study. Participants were asked to identify behaviours they might expect to observe if the risk factor was no longer active. As a result, participants might have focused on the elimination of risk rather than a decrease in risk. Future research could alter the wording of the survey questions such that participants focus on a decrease in the relevance of the risk factor rather than a complete eradication.

For both manifested dynamic risk factors and pro-social alternative behaviour, there was variation between risk factors in the frequency with which participants identified relevant behaviour. For instance, a greater number of relevant behavioural manifestations of interpersonal aggression were identified than those of sexual deviance. This variation might be related to the difference in visibility of behavioural manifestations of some dynamic risk factors compared with others (Mooney & Daffern, 2013).

In this study, some suggested examples of behavioural manifestations were uncommon and appeared unrelated to the risk factor in question. This finding may indicate that some psychologists are uncertain about the behaviours that sexual offenders might exhibit when a particular dynamic risk factor is active; alternatively, there may be idiosyncratic manifestations of these risk factors. Ultimately, careful scrutiny of the behaviour is required before determining that it is related to the risk factor and relevant to the person's offending. Risk factors that are more internal (i.e. related to thoughts and attitudes) produced more uncommon and seemingly irrelevant responses. This suggests that when clinicians cannot directly observe the relevant behaviour they produce a broader range of risk-related behaviours. This issue has important implications for the scoring of structured risk-assessment instruments; if clinicians have difficulty determining whether or not a behaviour is a valid manifestation of a risk

factor then the reliability with which risk-assessment measures are scored may be decreased, increasing the risk of item drift (Webster et al., 2002).

The overlap between risk factors and behavioural manifestations, such that the behaviours identified as examples of manifest dynamic risk factors and pro-social equivalents were themselves risk factors in the risk-assessment tools (e.g. 'substance use'), was an unexpected finding. In addition, the same behavioural manifestations were identified within several risk factors. For example, 'not accepting responsibility for behaviour' was identified as relevant for the following risk factors: 'problems with treatment'; 'substance use'; 'psychopathic personality disorder'; 'attitudes that support or condone sexual violence'; 'problems with employment'; 'extreme minimisation or denial'; 'non-sexual criminality'; 'offence planning'; and 'problems with self-awareness'. These findings raise questions about the most appropriate way of scoring a risk-assessment item; that is, how to determine which risk factors are relevant when a specific behaviour is observed in custody (e.g. if drug use is observed, is it best captured under the risk factor 'substance use' or does it also infer 'poor coping'?). Finally, this finding may highlight the strong inter-relationship between dynamic risk factors (Thornton, 2002), such that any behaviour may relate to various risk factors (e.g. using drugs in prison may be a consequence of problems related to child abuse, and problems with stress or coping).

Limitations and Future Research Directions

Not all risk factors related to sexual offending are directly observable in a custodial environment. Therefore, it can be difficult to determine whether these risk factors are active. For instance, offence-supportive attitudes, sexual preoccupation, and sexual deviance are all predictive of sexual offending (Mann et al., 2010) but tend to be related to thoughts

or more covert behaviour. This research highlights the potential difficulties associated with determining whether risk related to sexual self-regulation is current (e.g. sexual deviance and sexual compulsivity). Behaviour related to this risk area is generally reliant on offender self-report. Although participants in the current study appeared to have some understanding of behavioural manifestations of this risk area, the more overt manifestations suggested are likely to occur only in individuals who have significant problems with sexual self-regulation (e.g. excessive masturbation). The participants in the current study were all psychologists. In this role, they may only see offenders in clinical contexts rather than the general custodial environment. This limitation may have influenced the participants' ability to generate potential observed behaviour.

Prior research has demonstrated that behavioural manifestations of dynamic risk factors and the opposing pro-social behaviour can be reliably identified in incarcerated violent offenders through the use of a rating guide that identifies relevant behaviour from the VRS (Mooney & Daffern, 2013). Similarly, it is important in a custodial environment for inexperienced staff to have some understanding of the common behavioural manifestations of risk-related and positive behaviours. These behaviours could be monitored to assist risk assessment and change measurement. Although risk may not be eliminated, it can be reduced through interventions, including psychological treatment. The observations that participants provided in the current research can be used as examples of behaviours indicative of ongoing risk or pro-social change in important risk-related domains. In addition, custodial staff who may not be familiar with the relevant risk factors specific to sexual offenders could benefit from access to information about these manifestations to help their monitoring and management of offenders. Intended future research involves a trial of a checklist of potential risk-related behaviours in a

treatment programme for sexual offenders, based on the results of this study. Weekly completion of this checklist by both offenders and custodial staff could provide information about changes in the frequency of various risk-related behaviours over the course of treatment. Both clinicians and custodial staff could use such a checklist in the monitoring of incarcerated sexual offenders and for conducting risk assessments. A longer-term project could validate the checklist through a recidivism study that could evaluate whether changes in these behaviours (reductions in dynamic risk factors/OPBs and increases in the frequency of positive behaviour/PABs) over the course of treatment are related to changes in recidivism rates.

Conclusion

There are several practical and theoretical implications for these findings and two important conclusions. The first conclusion is that participants were able to identify behavioural manifestations that they considered relevant for each risk factor. They provided important information about the behaviours they expect to observe in incarcerated sexual offenders. The next step in furthering this field of research is to test the relationship between these behaviours and sexual offending following release into the community. One potential method may be to conduct a prospective evaluation to determine whether the most frequently observed behavioural manifestations are related to recidivism. That is, the behaviours identified in the current study have face validity, but further research is required to demonstrate their predictive validity.

The second conclusion is in relation to the differential ease with which behavioural manifestations were identified for different risk factors. This finding has implications for risk assessment. The finding that some behavioural manifestations identified were vague, or alternatively, more applicable to another risk factor might suggest that, at times, clinicians

can improperly identify a behavioural manifestation as an indicator of a particular risk factor when it may in fact be an indicator of a different risk factor. This issue is relevant for the scoring of risk assessments; it might lead to an inflated estimation of risk if one behaviour is considered indicative of a range of risk factors.

Given the complexity of the risk-assessment process, it might be beneficial to present information regarding common manifestations of risk factors to less experienced clinicians and other staff (e.g. custodial staff) to assist in the identification of relevant behavioural manifestations for each offender. They would then be better equipped to understand the relevance of specific behaviours to dynamic risk items in structured risk-assessment measures for each offender and use this information to contribute to progress monitoring and risk assessment. The use of structured risk-assessment tools may become more reliable and item drift may decrease. Additionally, it has been highlighted that clinicians need to accurately determine the risk areas that are active when particular behaviours are observed, through the process of functional analysis (e.g. whether aggressive communication is related to coping or interpersonal difficulties); this may help to determine whether the behaviour is relevant to a particular risk factor. Another important implication of the current research is related to the finding that the participants had more difficulty identifying pro-social behaviour than problematic behaviour. Clinicians should be encouraged to monitor and discuss the positive behaviour observed in offenders in order to enhance the risk-assessment and treatment process.

References

- Andrews, D. A., Bonta, J., & Wormith, J. (2006). The recent past and near future of risk and/or need assessment. *Crime and Delinquency*, 52, 7–27.
- Andrews, D. A., Bonta, J., & Wormith, J. (2011). Risk-Need-Responsivity (RNR) Model: Does adding the Good Lives Model contribute to effective crime prevention? *Criminal Justice and Behavior*, 38(7), 735–755.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77–101.
- Braun, V., & Clarke, V. (2013). *Successful qualitative research: a practical guide for beginners*. London: Sage.
- Daffern, M. (2010). A structured cognitive behavioural approach to the assessment and treatment of violent offenders using offence paralleling behaviour. In M. Daffern, L. Jones & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 105–120). Chichester: John Wiley and Sons Ltd.
- Daffern, M., Jones, L., Howells, K., Shine, J., Mikton, C., & Tunbridge, V. (2007). Refining the definition of offence paralleling behaviour. *Criminal Behaviour and Mental Health*, 17, 265–273.
- de Vries Robbé, M., de Vogel, V., & de Spa, E. (2011). Protective factors for violence risk in forensic psychiatric patients: a retrospective validation study of the SAPROF. *International Journal of Forensic Mental Health*, 10, 178–186. doi:10.1080/14999013.2011.600232
- Gordon, A., & Wong, S. (2009). *The offence analogue and offence reduction behaviour rating guide*. Unpublished user manual.
- Gordon, A., & Wong, S. (2010). Offence analogue behaviours as indicators of criminogenic need and treatment progress in custodial settings. In M. Daffern, L. Jones, & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 171–183). Chichester: John Wiley and Sons Ltd.
- Hanson, R., & Harris, A. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13, 105–122.
- Hart, S., Kropp, P., Laws, D., Klaver, J., Logan, C., & Watt, K. (2003). *The Risk for Sexual Violence Protocol (RSVP)*. Burnaby, BC: Simon Fraser University, Mental Health, Law, and Policy Institute.
- Jones, L. (1997). Developing models for managing treatment integrity and efficacy in a prison based TC: The Max Glatt Centre. In E. Cullen, L. Jones, & R. Woodward (Eds.), *Therapeutic communities for offenders* (pp. 121–157). Chichester: Wiley.
- Jones, L. (2010). History of the offence paralleling behaviour construct and related concepts. In M. Daffern, L. Jones, & J. Shine (Eds.), *Offence paralleling behaviour: A case*

- formulation approach to offender assessment and intervention* (pp. 3–23). Chichester: John Wiley and Sons Ltd.
- Liamputtong, P. (2012). *Qualitative research methods* (Fourth ed.). Australia: Oxford University Press Australia.
- Mann, R., Hanson, R., & Thornton, D. (2010). Assessing risk for sexual recidivism: some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, 22, 191–217.
- Mann, R., Thornton, D., Wakama, S., Dyson, M., & Atkinson, D. (2010). Applying the concept of offence paralleling behaviour to sex offender assessment in secure settings. In M. Daffern, L. Jones, & J. Shine (Eds.), *Offence paralleling behaviour: A case formulation approach to offender assessment and intervention* (pp. 121–135). Chichester: John Wiley and Sons Ltd.
- Markovic, M. (2006). Analyzing qualitative data: health care experiences of women with gynecological cancer. *Field Methods*, 18, 413–429.
- McDougall, C., Clarke, D., & Woodward, R. (1995). Application of operational psychology to assessment of inmates. *Psychology, Crime and Law*, 2, 85–99.
- Miller, H. (2006). A dynamic assessment of offender risk, needs, and strengths in a sample of pre-release general offenders. *Behavioral Sciences & the Law*, 24, 767–782.
- Mooney, J., & Daffern, M. (2013). The offence analogue and offence reduction behaviour rating guide as a supplement to violence risk assessment in incarcerated offenders. *International Journal of Forensic Mental Health*, 12, 255–264. doi:10.1080/14999013.2013.867421
- Olver, M., & Wong, S. (2011). A comparison of static and dynamic assessment of sexual offender risk and need in a treatment context. *Criminal Justice and Behavior*, 38, 113–126.
- Olver, M., Wong, S., Nicholaichuk, T., & Gordon, A. (2007). The validity and reliability of the Violence Risk Scale–Sexual Offender version: assessing sex offender risk and evaluating therapeutic change. *Psychological Assessment*, 19, 318–329.
- Pope, C., Ziebland, S., & Mays, N. (2000). Analyzing qualitative data. *British Medical Journal*, 320, 114–116.
- Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior*, 24, 595–605.
- Sheldrick, C. (1999). Practitioner review: the assessment and management of risk in adolescents. *Journal of Child Psychology and Psychiatry*, 40, 507–518.
- Simourd, D., & Hoge, R. (2000). Criminal psychopathy: a risk-and-need perspective. *Criminal Justice and Behavior*, 27, 256–272.
- Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment*, 14(2), 139–153.
- Webster, C., Müller-Isberner, R., & Fransson, G. (2002). Violence risk assessment: Using structured clinical guides professionally. *The International Journal of Forensic Mental Health*, 1, 185–193.
- Wong, S., Olver, M., Nicholaichuk, T., & Gordon, A. (2003). The Violence Risk Scale: Sexual Offender Version (VRS-SO). Regional Psychiatric Centre and University of Saskatchewan, Saskatoon, Saskatchewan, Canada.

Appendix B: Psychologists' Survey

1. Are you currently working with offenders who have been convicted of a sexual offence? If so, how many years have you worked with them?
2. If not, have you worked with them in the past? How many years did you work with them?
3. What are some risk factors you have observed sexual offenders to exhibit in custody? Please specify whether or not it was while in treatment.
4. Are you familiar with the term "Offence Paralleling Behaviour"? **Next screen: provide definition (e.g., "OPB is a behavioural sequence incorporating overt behaviours (that may be muted by environment factors), appraisals, expectations, beliefs, affects, goals and behavioural scripts, all of which may be influenced by the patient's mental disorder, that is functionally similar to behavioural sequences involved in previous criminal acts"; see Daffern et al., 2007).**
5. What are some examples of indicators that might suggest an offender in custody, for a conviction of sexual offending, is engaging in OPB?

Research has provided evidence to suggest that if individuals who have been convicted of a sexual offence have certain risk factors, they are at greater risk of reoffending than those who do not have these risk factors. Due to the relevance of assessing risk of reoffending, it is important to evaluate these individuals' behaviour in custody. However, risk related behaviours will likely be muted in custody and, therefore, we are trying to identify behaviours that professionals believe are the common manifestations of risk factors in custody.

The following list contains risk factors from two risk assessment tools, the Violence Risk Scale: Sexual Offender version (VRS:SO) and the Risk for Sexual Violence Protocol (RSVP). For each of the risk factors, you will be asked to consider the behaviours you might expect to see in an individual if that risk factor is still operative. In addition, as an individual makes changes, unhelpful behaviours would be expected to decrease and be replaced by more prosocial and adaptive behaviours. Therefore, you will also be asked to consider the behaviours you might expect to see in an individual that suggests the risk factor is no longer important.

- 1) Chronicity of sexual violence (i.e., persistence and frequency)
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:
- 2) Diversity of sexual violence (i.e., multiple types of sexual violence, varying in nature and victim selection)
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

- 3) Escalation of sexual violence (i.e., become progressively more frequent, serious, or diverse over time)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

- 4) Physical coercion in sexual violence (i.e., acts that are intended to cause harm or fear of physical harm in the victim)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

- 5) Psychological coercion in sexual violence (i.e., acts that involve threatened loss or promised gain of status, privilege, favour, or affection)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

- 6) Extreme minimisation or denial of sexual violence (i.e., failure to admit to or accept responsibility for acts of sexual violence and the consequences of those acts)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

- 7) Attitudes that support or condone sexual violence (i.e., beliefs and values that directly or indirectly encourage or excuse coercive sex, sex with minors, and other sexual violence)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

- 8) Problems with self-awareness (i.e., lack of appreciation for the factors and processes that place the person at risk of sexual violence)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

- 9) Problems with stress or coping (i.e., extent to which the person's psychosocial adjustment is unstable or susceptible to external events and occurrences)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

- 10) Problems resulting from child abuse (i.e., serious problems in psychosocial adjustment that are the result of abuse experiences in childhood or adolescence, including childhood victimisation, sexual abuse, physical abuse, neglect)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

- 11) Sexual deviance (i.e., stable pattern of deviant sexual arousal)

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

12) Psychopathic personality disorder
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

13) Major mental illness
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

14) Problems with substance use
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

15) Violent or suicidal ideation
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

16) Problems with intimate relationships
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

17) Problems with non-intimate relationships
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

18) Problems with employment
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

19) Non-sexual criminality
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

20) Problems with planning
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

21) Problems with treatment
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

22) Problems with supervision
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

23) Sexual compulsivity
Expected behaviours if risk factor is operative:
Adaptive behaviours expected to replace the risk factor:

24) Offence planning

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

25) Interpersonal aggression

Expected behaviours if risk factor is operative:

Adaptive behaviours expected to replace the risk factor:

Please list any other examples of behavioural manifestations that have not been covered in the previous questions:

If you have any comments about the value of monitoring Offence Paralleling Behaviour in custody or have any additional comments, please provide them here:

Appendix C: Behavioural Checklist

Please indicate how many times you engaged in the following behaviours over the *past week*, in the blank column next to the scale. If you cannot recall exactly how many times, please provide an estimate by placing a cross on the rating scales. If you engaged in the behaviour over 50 times, please write the number in the blank column (e.g. 100+).

	0	50
Not accepting responsibility for offending behaviour (e.g. anger at discussing it)		
Difficulty engaging in treatment (e.g. refuses to discuss offending)		
Accepting responsibility for offending/unhelpful behaviour (e.g. discusses behaviour without defensiveness, takes responsibility for mistakes)		
Engaging with treatment (e.g. accepts feedback, considers function of behaviour)		
Blaming victims of sexual offences or justifying behaviour		
Supporting others' unhelpful attitudes about sexual offending		
Challenging unhelpful comments about sexual offending		
Victim awareness (understanding the harm of offending)		
Not accepting own role in decisions (e.g. minimises own contribution to difficulties)		
Difficulty identifying when triggers for unhelpful behaviour arise		
Accepting own role in decisions (e.g. understands consequences, reflects on behaviour)		
Managing unhelpful behaviour (e.g. removes self, discusses plans to minimise risk)		
Fluctuating emotions (e.g. irritability, overreacts to triggers)		
Withdrawal / isolation		
Talking about problems with supports		
Good problem solving (e.g. thinks before acting, manages emotions, healthy routine)		
Problems forming appropriate attachments (e.g. needy, keeps emotional distance)		
Fear or mistrust (e.g. defensive, withdrawing, hypervigilant)		
Appropriate attachments (e.g. appropriate boundaries, trusts others and talks openly)		
Self awareness of impact of childhood trauma (e.g. discusses emotional reactions)		
Collecting inappropriate images (e.g. children, violent sex)		
Inappropriate sexual arousal (e.g. TV shows, arousal to talking about offending)		
Open about sexual thoughts/behaviour (e.g. talks about problems managing them)		
Managing sexual thoughts (e.g. replaces deviant thoughts, leaves unhelpful situations)		
Manipulating others (e.g. plays others against each other, wants special treatment)		
Superficial (e.g. not applying learned information to self, disconnection from emotion)		
Perspective taking (e.g. thinking of others' needs)		

Interest in meaningful relationships (e.g. mutual interactions)		
Extreme mood states (e.g. low or elevated)		
Auditory or visual hallucinations (e.g. talks to self, speaks in a non-sensical manner)		
Medication compliance		
Stable mental states (e.g. stable behaviour; appropriate responding)		
Positive urinalysis		
Drug seeking behaviour (e.g. medication seeking from the Clinic)		
Negative urinalysis		
Engaging in AOD intervention (e.g. Getting SMART, medical assistance)		
Threats of violence or self harm (e.g. talks about wanting to)		
Emotional instability (e.g. irritability, sudden change in presentation)		
	0	50
Reduction in threats of self harm or violence to others		
Seeking help to address self harm (e.g. uses distress tolerance strategies)		
Communication difficulties (e.g. aggressive phone calls to partner)		
Difficulty relating to others (e.g. superficial, discomfort with closeness)		
Taking responsibility for relationships (e.g. maintains contact, repairs relationships)		
Stable relationships (e.g. ongoing positive interactions, invested in relationships)		
Withdrawal or isolation (e.g. avoids supports)		
Difficulties maintaining friendships (e.g. spends time alone, difficult interactions)		
Interacting with others appropriately (e.g. respectful, socialising)		
Meaningful friendships/contact with supports (e.g. with inmates/staff)		
Avoiding domestic duties		
No goals (lack of motivation)		
Regular completion of domestic duties without prompting		
Seeking training		
Warnings or charges		
Pro-criminal attitudes (e.g. accepts returning to custody, jokes about criminality)		
No warnings or charges (abide by the rules)		
Future planning (e.g. plans for future without offending)		
Difficulties planning for release (e.g. no plans / unrealistic plans)		
Difficulty with problem solving (e.g. does not think about consequences)		
Release planning (e.g. realistic plans, talks to supports about future)		
Good problem solving (e.g. identifies thoughts, thinks about consequences)		
Disengaging from treatment (e.g. does not complete task work)		
Not thinking about what his needs are in treatment (e.g. passive participation)		

Appendix D: Ethics Approval



Department of Justice & Regulation

Justice Human Research Ethics Committee

Information Integrity & Access
Melbourne Victoria 3000
GPO Box 123A
Melbourne Victoria 3001
Telephone: (03) 8684 1514
DX210077

8 December 2015

Reference: CF/15/18371

Professor Michael Daffern
Swinburne University of Technology
505 Hoddle Street
Clifton Hill VIC 3068

Re: Validating the Satisfactory Behaviour Rating Scale (SBRS)

Dear Professor Daffern

The Department of Justice and Regulation Human Research Ethics Committee (JHREC) considered your response to the issues raised in relation to the project *Validating the Satisfactory Behaviour Rating Scale (SBRS)* at its meeting on Thursday 19 November and has now granted **full approval** for the duration of the investigation. The Department of Justice and Regulation reference number for this project is CF/15/18371. Please note the following requirements:

- To confirm JHREC approval sign the Undertaking form attached and provide both an electronic and hardcopy version within ten business days.
- The JHREC is to be notified immediately of any matter that arises that may affect the conduct or continuation of the approved project.
- You are required to provide an Annual Report every 12 months (if applicable) and to provide a completion report at the end of the project (see the Department of Justice and Regulation Website for the forms).
- Note that for long term/ongoing projects approval is only granted for three years, after which time a completion report is to be submitted. The project must be renewed with a new application before the initial three year period has expired.
- The Department of Justice and Regulation would also appreciate receiving copies of any relevant publications, papers, theses, conferences presentations or audiovisual materials that result from this research.
- All future correspondence regarding this project must be sent electronically to ethics@justice.vic.gov.au and include the reference number and the project title. Hard copies of signed documents or original correspondence are to be sent to The Secretary, Justice Human Research Ethics Committee, Level 24, 121 Exhibition Street, Melbourne, VIC 3000.

If you have any queries regarding this application you are welcome to contact the Secretary on (03) 8684 1514 or email: ethics@justice.vic.gov.au.

Yours sincerely

Ashleigh Bell, Secretary
Department of Justice & Regulation Human Research Ethics Committee



Department of Justice & Regulation Human Research Ethics Committee

UNDERTAKING

Project Title: Validating the Satisfactory Behaviour Rating Scale (SBRS)

Reference No. CF/15/18371

I acknowledge that I have read the conditions outlined in the current guidelines of the Department of Justice and Regulation Human Research Ethics Committee (JHREC), and undertake to abide by them.

Reporting requirements:

- *RE: Amendments: I will ensure that an Amendment Request Form is submitted to the JHREC if amendments to the project are required (e.g. staff changes, extension of completion date and adjustments to aims/methodology).*
- *RE: Amendments: If my JHREC application included a Department of Justice and Regulation (DJR) letter of support, I will advise the DJR contact officer of proposed amendments before an amendment request is submitted to the JHREC.*
- *RE: Annual Reports: I will ensure that annual reports are provided if my project extends 12 months in duration.*
- *RE: Completion Reports: I will ensure that a completion report is provided at the conclusion of the research.*
- *RE: Long term/Ongoing Projects: I acknowledge that if my project is an ongoing/ long-term project I need to provide a completion report at the end of every three-year period and renew by submitting a new JHREC application before the end of the three-year period. I further acknowledge that if I fail to renew the project before the three-year period expires, the previous JHREC approval will cease to have effect on expiry of the three year period.*

Name of Principal Researcher: _____

Signed (Principal Researcher): _____

Date: _____



Human Ethics Certificate of Approval

This is to certify that the project below was considered by the Chair of the Monash University Human Research Ethics Committee. The Chair was satisfied that the proposal meets the requirements of the *National Statement on Ethical Conduct in Human Research* and has granted approval.

Project Number: CF13/2736 - 2013001469

Project Title: An analysis of offence paralleling behaviour in a custody-based sex offender treatment program

Chief Investigator: Assoc Prof Michael Daffern

Approved: From: 1 October 2013

To: 1 October 2018

Terms of approval - *Failure to comply with the terms below is in breach of your approval and the Australian Code for the Responsible Conduct of Research.*

1. Approval is only valid whilst you hold a position at Monash University and approval at the primary HREC is current.
2. **Future correspondence:** Please quote the project number and project title above in any further correspondence.
3. **Final report:** A Final Report should be provided at the conclusion of the project. MUHREC should be notified if the project is discontinued before the expected date of completion.
4. **Retention and storage of data:** The Chief Investigator is responsible for the storage and retention of original data pertaining to a project for a minimum period of five years.

Professor Nip Thomson
Chair, MUHREC

cc: Assoc Prof Richard Kemp, Ms Tamara Sweller



Corrective Services NSW Attorney General & Justice

20 Lee Street
SYDNEY NSW 2000
GPO Box 31 SYDNEY NSW 2001 | DX 22
Tel 02 8346 1333 | Fax 8346 1010
www.correctiveservices.nsw.gov.au

D13/ 538 250

23 August 2013

Ms Tamara Sweller
Psychologist
Metropolitan Special Programs Centre
Long Bay Correctional Complex
Locked Bag 20, Anzac Parade
Matraville NSW 2036

Dear Ms Sweller

I refer to your research application entitled "*An analysis of offence paralleling behaviour in a custody-based sex offender treatment program: Can this behaviour be used to predict risk of reoffending*". The project will seek to investigate whether an inmate's behaviour changes in custody as a result of the treatment they receive. Individuals commit criminal offences for a variety of reasons, their behaviour preceding, during and after the offence seen as an attempt to achieve a goal. Specifically, the behaviour of inmates will be analysed to identify offence paralleling behaviour (OPB) and the pro-social alternate behaviours (PAB) that replace OPB as a consequence of successful treatment.

I am pleased to inform you that conditional approval has been given for your research project. The conditions of approval are that you comply with the 'Terms and Conditions of Research Approval' [Attachment A]

I wish you every success in your endeavours.

Yours sincerely


PETER SEVERIN
COMMISSIONER

To: Prof Michael Daffern/Ms Tamara Sweller

Dear Michael and Tamara

SHR Project 2014/068 An analysis of offence paralleling behaviour in a custody-based sex offender treatment program

Prof Michael Daffern, FHAD; Ms Tamara Sweller et al

Approved Duration: 01/04/2014 to 31/12/2018 [Adjusted]

I refer to your application for Swinburne ethics clearance for the above supervised student project transferred from Monash University to Swinburne.

Documentation pertaining to the request was contained in emails, with attachments, sent 27 February and 3 March 2014 (T Sweller to K Wilkins) and 17 March 2014 (M Daffern to K Wilkins, including clarification re project status). This documentation was put to a delegate of Swinburne's Human Research Ethics Committee (SUHREC) for consideration significantly on the basis of the prior ethical review documentation and clearances issued. In this regard, information re the original clearances issued by Monash University (CF13/2736 – [2013001469](#)) and Corrective Services NSW (CSNSW) (D13/538250 of 23 August 2013) and subsequent clearance of the transferred project by CSNSW (undated letter issued by 14 February 2014) was noted. Also of significance, in this instance, sensitive/personal/health information (as per applicable Privacy and health records legislation) is being secured within the student researcher's workplace within NSW.

I am pleased to advise that, as submitted to date, Swinburne ethics clearance has been given prospectively for the project to continue in line with standard on-going ethics clearance conditions outlined below. SUHREC can be expected to ratify the clearance in due course.

Special Note:

CSNSW may need to be apprised of the Swinburne ethics clearance here issued and note/endorse the revised consent instruments now to be used and note that previous participants would accordingly be apprised of the project's changed circumstances. Responsibility for previous clearances issued remain with Monash University and CSNSW respectively (as applicable) and as regards researcher compliance with the clearances.

- All human research activity undertaken under Swinburne auspices must conform to Swinburne and external regulatory standards, including the current *National Statement on Ethical Conduct in Human Research* and with respect to secure data use, retention and disposal.
- The named Swinburne Chief Investigator/Supervisor remains responsible for any personnel appointed to or associated with the project being made aware of ethics clearance conditions, including research and consent procedures or instruments

approved. Any change in chief investigator/supervisor requires timely notification and SUHREC endorsement.

- The above project has been approved as submitted for ethical review by or on behalf of SUHREC. Amendments to approved procedures or instruments ordinarily require prior ethical appraisal/ clearance. SUHREC must be notified immediately or as soon as possible thereafter of (a) any serious or unexpected adverse effects on participants and any redress measures; (b) proposed changes in protocols; and (c) unforeseen events which might affect continued ethical acceptability of the project.
- At a minimum, an annual report on the progress of the project is required as well as at the conclusion (or abandonment) of the project. (Reports and requests made to CSNSW and their HREC also being submitted to Swinburne Research for processing/endorsement may suffice.)
- A duly authorised external or internal audit of the project may be undertaken at any time.

Please contact the Research Ethics Office if you have any queries about Swinburne on-going ethics clearance, citing the project number. A copy of this clearance email should be retained as part of project record-keeping.

Best wishes for the transferred project.

Yours sincerely

Keith

Keith Wilkins
Secretary, SUHREC & Research Ethics Officer
Swinburne Research (H68)
Swinburne University of Technology
P O Box 218
HAWTHORN VIC 3122
Tel [+61 3 9214 5218](tel:+61392145218)
Fax [+61 3 9214 5267](tel:+61392145267)

Appendix E: Participant Information Sheet and Consent Form

INFORMATION SHEET FOR PARTICIPANTS



Project: An analysis of offence paralleling behaviour in a custody-based sex offender treatment program

This information sheet is for you to keep.

My name is Tamara Sweller. I am a PhD student at Swinburne University of Technology, and I am conducting a research project with Professor Michael Daffern, from the Centre for Forensic Behavioural Science and Legal Studies, Swinburne University of Technology, as well as Associate Professor Richard Kemp, from the University of New South Wales. This means that I will be writing a PhD thesis. The aim is to publish the results of this project.

YOU ARE INVITED TO TAKE PART IN THIS STUDY. PLEASE READ THIS INFORMATION SHEET IN FULL BEFORE MAKING A DECISION.

This project seeks participants who have been convicted of a sexual offence and are currently in a treatment program within NSW prisons. The aim of the project is to study behaviours around the time of participants' offending and similarities/differences with their behaviours in prison.

Step 1: The researcher will look at official documents from our computer system (e.g., Judge's Sentencing Remarks, Psychological reports) and will interview you to get more information about what your life was like around the time of offending and about your offence. After this information is collected, the researcher will write up an outline of the relevant factors involved in your offending behaviour. Based on this information, the researcher will predict the kinds of behaviours that you might engage in while going through treatment.

Step 2: Each week, a custodial officer in CUBIT will complete a form with a list of behaviours that are common in offenders in treatment. The custodial officer will fill out this form to say whether they saw you behaving in certain ways over the past week. You will also be asked to fill out this form each week so that you also have input in saying which behaviours you think you engaged in during that week.

Step 3: The researcher will record the changes in these observed behaviours over the time you are in treatment, to see what impact treatment has on the way you think and behave.

Step 4: Several years after completing this study, research might be done to determine whether behaviours seen by offenders in treatment can be used to predict the likelihood of reoffending after release. This research would be done by accessing your Police records. You will not be asked to do anything as part of this step.

If you consent to participate, you will not be asked to do anything that will cause discomfort or risk. Talking about your offending might be uncomfortable, but the interview will be similar to other interviews you have done before. The project will begin from the time you start treatment until you finish treatment, but if you consent to participate, you are free to withdraw consent and discontinue participation in the project at any time. There will be no penalty or prejudice of any kind if you do not participate in the project. When you are asked questions in the initial interview, you do not have to answer any questions you do not feel comfortable answering.

The interview will take about 30 minutes and the weekly form will take between five and 10 minutes each week.

The information you give to the researcher during participation in the project is confidential. It will not be linked to your treatment. However, the researcher may have to inform authorities if you talk about any offence that you have not been charged for before.

If you choose to consent to participate in the project, permission will be given to the researcher to get information from Corrective Services NSW about your offences and your behaviour. Your individual information will not be made public. No identifiable information will be sent from NSW to Victoria. The information collected during this research will be kept by the researchers for seven years, or for five years after the results have been published, whichever is later. If you want to gain access to the final results of the research, you may do so.

If you have any questions about the project (e.g., about procedures), you can contact the researcher, Tamara Sweller, on 9289 2434 at CUBIT, MSPC 2.

Thank you for your time.

Tamara Sweller

CONSENT FORM FOR PARTICIPANTS



Project: An analysis of offence paralleling behaviour in a custody-based sex offender treatment program

NOTE: This consent form will remain with the Swinburne University of Technology researcher for their records

Student Researcher: Tamara Sweller

Supervisor: Professor Michael Daffern

Co-Supervisor: A/Professor Richard Kemp

I.....agree to participate in the Swinburne University of Technology research project specified above. I have had the project explained to me and I have read the Information Sheet, which I can keep for my records.

The study has been explained to me.

I understand that the researcher may talk to me about my offences and my behaviour.

The researcher may take written notes on what I say.

I may have to talk to the researcher for about 30 minutes, but I can stop sooner if I want to. I do not have to answer any questions if I do not want to.

Staff members in CUBIT will observe my behaviour and will give this information to the researcher.

I will be asked to report on my behaviour in CUBIT.

I know that no-one will mind if I decide that I do not want to take part in the study and I can pull out of the study at any time.

The researcher has agreed not to tell anyone my name or any other personal details.

It is OK for the researcher to look in my Corrective Services NSW files.

It is OK for researchers to access my Police record as a follow up study.

I have read or have had read to me, this consent form.

I understand that if I discuss an offence for which I have not been charged or

convicted the researcher will be obliged to report it to the authorities (i.e., custodial management, NSW Police etc).

If I have any questions about the study I can contact the researcher, Tamara Sweller, on 9289 2434.

PARTICIPANT (signature).....

DATE.....

WITNESS (signature).....

DATE.....