

Swansea University E-Theses

Evaluation On Probation: Developing Gender-Responsive And Jurisdictionally Appropriate Evaluation Systems For Offender Services In The British Isles

Kerr, Bridget A.

How to cite:

Kerr, Bridget A. (2018) *Evaluation On Probation: Developing Gender-Responsive And Jurisdictionally Appropriate Evaluation Systems For Offender Services In The British Isles*. Doctoral thesis, Swansea University.
<http://cronfa.swan.ac.uk/Record/cronfa49211>

Use policy:

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence: copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder. Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

Please link to the metadata record in the Swansea University repository, Cronfa (link given in the citation reference above.)

<http://www.swansea.ac.uk/library/researchsupport/ris-support/>

Evaluation on Probation:
Developing Gender-Responsive and Jurisdictionally
Appropriate Evaluation Systems for Offender
Services in the British Isles

Bridget Kerr

Submitted to Swansea University in fulfilment of the requirements
for the Degree of Doctor of Philosophy

Swansea University

2018

THESIS SUMMARY

Candidate's Surname / Family Name	KERR
Candidate's Forenames	BRIDGET ALICE
Candidate for the Degree of	PhD
Full title of thesis	Evaluation on probation: Developing gender-responsive and jurisdictionally appropriate evaluation systems for offender services in the British Isles.

Summary:

This thesis is an account of the application of a structured, evidence-based Canadian instrument for the evaluation of offender intervention services (the Correctional Program Assessment Inventory-2010), and explores the limits of technology transfer between jurisdictional contexts. As well as analysis of evaluation results in two services in the British Isles, this has entailed investigation of the response made by criminal justice agencies and practitioners and how the instrument may need to be modified to operate meaningfully in diverse contexts.

During this study, inter-jurisdictional application of the CPAI-2010 exposed turbulent political environments, which had far-reaching implications for the potential of evaluation. This study demonstrates that, to sustain evaluation as a resource for change and development, it must respond to and reflect the political context of agencies, integrate all relevant approaches (eg, Risk-Need-Responsivity, desistance-focused contributions, gender-responsive work, etc) and engage all key social actors (service-users, practitioners, managers, etc) in a process of reflection. Its findings indicate that adopting a motivational language, placing a service on a continuum and timing evaluation at an appropriate frequency are all factors that can reflect and ameliorate organisational difficulties. Furthermore, this study highlights that evaluation of organisational context is key to meaningful engagement with the core issues of a service, indicating the inadequacy of evaluation systems focusing solely on practice and/or outcomes.

Recidivism rates may be a particularly inadequate outcome measure for women's projects, which have their emphasis on the provision of holistic and diversionary approaches, aiming to keep women (who, as a service-user group, present a low risk of re-offending in any case) out of the disproportionately punitive prison system.

The innovation of this study is the development of an evidence-based evaluation approach that meets the needs of services, specifically women's projects, in the British Isles: the Swansea Service Evaluation Inventory – Women's Projects (SSEI-W).

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed (candidate)

Date

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed (candidate)

Date

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed (candidate)

Date

CONTENTS

Thesis Summary	p. i
Declarations and Statements	p. ii
Contents	p. iii
Acknowledgements	p. iv
List of Tables and Figures	p. v
List of Abbreviations	p. vi
Key to Interview Participants	p. viii
Introduction	p. 1
Chapter 1 From So Many Words to the Last Word?	p. 10
Chapter 2 From Research to the Real World Part I: the Implementation Gap	p. 42
Chapter 3 Methodology	p. 70
Chapter 4 Findings from Phase One: CPAI-2010 Evaluation Results	p. 107
Chapter 5 Findings from Phase Two: Themes from Semi-Structured Interviews	p. 144
Chapter 6 Amendments, Adaptations and Developments: the SSEI-W	p. 164
Chapter 7 From Research to the Real World Part II: Bricks for the Bridge	p. 181
Chapter 8 Conclusion: the Study in Context	p. 197
Index to Appendices	p. 218
Bibliography	p. I

ACKNOWLEDGEMENTS



Thank you, Peter, for sticking with me through all the missed deadlines. I'm so grateful for having had the opportunity to work under your supervision.

Thanks to everyone who made the study possible, most especially Paul, Yvette and Paula in Canada, Pamela and Sue at Swansea University and the management, practitioners and service-users of the participant services in Wales and Jersey.

Thanks to Tim. You really had my back.

For Alimum.

TABLES AND FIGURES

Tables

Table 1.1	Mean Predictive Validity Estimates for Gender Informed Risk/Need Factors by Gender (Andrews & Bonta, 2010, p. 68)	p. 26
Table 1.2	The Risk-Need-Responsivity (RNR) Model of Effective Correctional Assessment and Crime Prevention Services (Andrews & Bonta, 2010, pp. 46-47)	p. 38
Table 3.1	The Coding Process in Inductive Analysis (adapted from Creswell, 2002, Figure 9.4, p. 266)	p. 84
Table 3.2	Study Participants	p. 106
Table 4.1	Core Correctional Practices (CCPs)	p. 130
Table 4.2	Domain G Scoring of Relationship and Structuring Skills	p. 132
Table 6.1	CPAI-2010 and Potential SSEI-W Categories	p. 177
Table 6.2	CPAI-2010 and SSEI-W Subsections	p. 178
Table 6.3	CPAI-2010 and SSEI-W Scoring Distribution	p. 178
Table 7.1	CPAI-2010 Language Issues	p. 191

Figures

Fig. 4.1	CPAI-2010 Evaluation Scores by Domain	p. 108
Fig. 7.1	Perceptions of Evaluation Dichotomy: “Reflection” vs “Inspection”	p. 184

LIST OF ABBREVIATIONS

ADAPT	Adapt Domestic Abuse Prevention Training
AI	Appreciative Inquiry
BASS	Building a Safer Society strategy (Jersey)
BSC	British Society of Criminology
CCP	Core Correctional Practice
CJS	Criminal Justice System
CLMM	Conceptual Level Matching Model
CPAI	Correctional Program Assessment Inventory
CPAP	California Program Assessment Process
CPC	Correctional Program Checklist
CPC-CSA	Correctional Program Checklist – Community Supervision
CPC-DC	Correctional Program Checklist – Drug Court
CPC-GA	Correctional Program Checklist – Group Assessment
CR	Confidence Rating (on the CPAI)
CRC	Community Rehabilitation Company
CRP	Crime Reduction Programme
CREDOS	Collaboration of Researchers for the Effective Development of Offender Supervision
CRP	Crime Reduction Programme
CSAP	Correctional Services Accreditation Panel
DSPD	Dangerous and Severe Personality Disorder programme
EPICS	Effective Practices in Community Supervision
ESRC	Economic and Social Research Council
ETE	Education, Training and Employment
ETS	Enhanced Thinking Skills programme
GLM	Good Lives Model
GPCSL	General Personality and Cognitive Social Learning theory
GRCPAI	Gender-Responsive Correctional Program Assessment Inventory
GRPA	Gender-Responsive Program Assessment tool
GT	Grounded Theory
HMIP	Her Majesty’s Inspectorate of Probation
HMPPS	Her Majesty’s Prison and Probation Service
IOM	Integrated Offender Management
ISP	Intensive Supervision Program
I-Level	Interpersonal maturity level
JPACS	Jersey Probation and After-Care Service
JS3	Jersey Supervision Skills Study
LS/CMI	Level of Service/Case Management Inventory
LSI-R	Level of Service Inventory – Revised
LSI-R-SV	Level of Service Inventory – Revised – Screening Version
MI	Motivational Interviewing
NAPO	National Association of Probation Officers
NPS	National Probation Service
NOMS	National Offender Management Service
OASys	Offender Assessment System

OGRS	Offender Group Reconviction Scale
OINTOC	Offending Is Not The Only Choice programme
OPDP	Offender Personality Disorder Pathway
ORA	Offender Rehabilitation Act
PbR	Payment by Results
PCM	Paragraph Completion Method
PCS	Proactive Community Supervision
PIC-R	Personal, Interpersonal and Community-Reinforcement perspective
PRO	Patient Reported Outcome
RNR	Risk-Need-Responsivity principles/model
SEEDS	Skills for Effective Engagement, Development and Supervision
SFS	Salient Factor Score
SMART	Self-Management And Rational Thinking programme
SSEI-W	Swansea Service Evaluation Inventory – Women’s Projects
SSET	Swansea Service Evaluation Team
STARR	Staff Training Aimed at Reducing Re-arrest
STICS	Strategic Training Initiative in Community Supervision
TF	Think First programme
TIER	Trauma Informed Effective Reinforcement System
TR	Transforming Rehabilitation
WPT	Wales Probation Trust
YJ	Youth Justice

KEY TO INTERVIEW PARTICIPANTS

Service	Partner Agency	Role	Identifier	Code
JPACS (J)		Chief Probation Officer (CPO)		JCPO
		Assistant Chief Probation Officer (ACPO)		JACPO
		Team Leader (TL)		JTL
		Probation Officer (PO)	Numbered 1-5	JPO1, etc
		Research & Information Officer (RO)		JRO
		Practitioner (P)	Numbered 1-2	JP1, etc
		Magistrate (MAG)		JMAG
		Service-User (SU)	Numbered 1-6	JSU1, etc
Women's Project (W)		Project Manager (PM)		WPM
		Research Officer (RO)		WRO
		Team Manager (TM)		WTM
		Probation Officer (PO)	Numbered 1-5	WPO1, etc
		Service User (SU)	Numbered 1-3	WSU1, etc
	Partner Agency 1 (PA1)	Manager (M)		WPA1M
		Practitioner (P)		WPA1P
	Partner Agency 2 (PA2)	Manager (M)		WPA2M

INTRODUCTION

It is a fact of life that researchers can produce as much high-quality research as they wish, but, unless there are those who can span the ravine between research and practice, the books and journals are only so many words on paper. (Leschied et al, 2001, p. xvii).

This thesis is an account of an attempt to “span the ravine between research and practice” in the British Isles using an evidence-based Canadian instrument for the evaluation of offender intervention services, the Correctional Program Assessment Inventory – 2010 (CPAI-2010).

The issue of policy transfer – the way and extent to which criminal justice policy has been influenced by international developments, particularly those of the United States – is of considerable interest in Britain, where “there is a very clear perception that the United States has been either the direct source of, or at least the inspiration for, a number of the policy developments” (Jones & Newburn, 2007, p. 1), including the ‘what works’ evidence base upon which the CPAI is built (ibid, p. 3). Central to the current study is the interrelated concept of *technology* transfer. At its most basic level, this concerns “transfer from the “experts” to the practitioners” (Gendreau, 1996b, pp. 150-151) and has the deceptively simple aim of “getting the necessary information into the hands of those who need it” (ibid, p. 154). Gendreau goes on to argue that successful transfer in fact requires supportive administration, congruent values, sustainable initiatives and, since practitioners tend not to read journals, an action-oriented approach from academics involving workshops, non-academic conference presentations, media coverage and engagement with professional, private-sector and political organisations. (ibid, pp. 154-155). Dolowitz and Marsh articulate the potential for failed policy transfer in three distinct forms: uninformed transfer (ie, the recipient culture does not have the full picture of the policy and how it operates at source); incomplete transfer (where partial transfer of the policy results in the loss of components integral to its success); and inappropriate transfer (where adequate consideration is not given to social, economic, political and cultural context) (Dolowitz & Marsh, 2000, p.17). This latter point is at the crux of this study, which positions itself at the forefront of efforts to

explore ‘evaluation responsivity’, a term coined by Ugwudike (2016) to describe “the importance of designing research-based evaluation tools that accommodate variations in local service delivery contexts”, in other words, getting the necessary *relevant* and *appropriate* information into the hands of those who need it *in a way that they can use it*.

This thesis seeks to identify what constitutes *responsive* evaluation practice by investigating the CPAI evaluation results for two probation services in the British Isles (one gender-neutral and one women’s project) and drawing upon the evaluation experience of participants. It flags up the potential for improved practices and the unification of purportedly oppositional approaches across diverse contexts where due attention is given to responsivity in the development of evaluation instruments and approaches.

In terms of the social, economic, political and cultural contexts of the evaluated services, it is significant that this study was undertaken at a time when Payment by Results (PbR) was being heralded by the government as the “rehabilitation revolution” in England and Wales, purportedly prioritising the reduction of recidivism. In the months leading up to data collection, Probation Trusts were dissolved and replaced by a single public sector National Probation Service (NPS) (responsible for the management of high-risk offenders) and 21 Community Rehabilitation Companies (CRCs) from the private, voluntary and social sectors (managing low- to medium-risk offenders). The government insistence on a “relentless focus on rehabilitation” not only involved agencies becoming “market providers” but, according to the Ministry of Justice, was to “incentivise innovation, paying providers by results for delivering reductions in reoffending” (Ministry of Justice, 2013a, pp. 3-6).

The “Transforming Rehabilitation” (TR) agenda continues to be disputed by critics. The National Association of Probation Officers’ (NAPO) early arguments that the service would be fragmented by an approach that prioritises the private sector, with its profit-driven ethos resulting in staff cuts that would “severely compromise public protection” (<http://www.napo.org.uk>), appear to have been borne out:

Payment is triggered by task completion rather than anything more meaningful. Additional financial rewards are far off and dependent on reoffending rates that are not altogether within the CRC's gift. CRC total workloads (and therefore income) are less than anticipated when contracts were signed. As CRCs continue to develop and adjust their operating models accordingly, CRCs are hard-pressed and are generally giving priority to work that is rewarded with more immediate and more substantial payment. (HM Inspectorate of Probation, 2016a, p.3).

This study is not concerned with exploring the conflict between the profit-driven ideology of the private sector and that of the public sector, whose practices are rooted in notions of fairness and legitimacy. Nevertheless, it ultimately reinforces the importance of acknowledging the unique pressures of the political environment when developing evaluation for agencies that, aside from commercial concerns, must be evidence-based in order to deliver effective services in the open (or *any*) market.

Central to reforms is The Offender Rehabilitation Act (ORA) 2014, which came into full force on 1st February 2015 and extends statutory supervision to include an unprecedented 'through the gate' community resettlement service for all offenders serving less than 12 months in custody¹. This especially affects women offenders, whose prison population has dramatically increased in recent years², who are proportionately more likely to serve short sentences for non-violent offences (Ministry of Justice, 2017b; Prison Reform Trust, 2017a), and for whom the extension of supervision represents increased possibilities of conflict with the criminal justice system for failing to comply (in addition to the ostensible benefits of increased opportunities for support) (Annison et al, 2015, p. 34). The 2017 *Bromley Briefings Prison Factfile* points out:

- Most women entering prison under sentence (83%) have committed a non-violent offence.
- The number of women recalled to custody whilst under supervision after their release has doubled since the end of 2014 – just before mandatory supervision was introduced for people serving sentences of 12 months or less on release.

¹ The provision of *voluntary* after-care was a key responsibility of probation services prior to the introduction of compulsory supervision for prisoners sentenced to twelve months or more (as a provision of the Criminal Justice Act 1991) alongside a contemporaneous shift away from welfare-oriented services and towards public protection and crime prevention (Maguire et al, 2000).

² The women's prison population in England and Wales more than doubled between 1995 and 2010, from 1,979 to 4,236 (Table A1.2, Ministry of Justice, 2014). McIvor notes that the average number of women in prison increased by 173% between 1992 and 2002, compared with an increase of 50% in the average male population at the same time (2007, pp. 7-8). More recently the number of women in prison in England and Wales has declined slightly, with 3,893 women in prison on 1st June 2018 (Ministry of Justice, 2018b).

- 48% of women are reconvicted within one year of leaving prison. This rises to 61% for sentences of less than 12 months and to 78% for women who have served more than 11 previous custodial sentences. (Prison Reform Trust, 2017b, p. 34).

It is clear to see why issues specific to women across custodial and community services are the focus of a growing body of literature and campaigns. Such is the impact of this work that a statutory duty to identify and take into consideration the needs of female offenders in the commission of supervision and rehabilitation services has been embedded into Section 10 of the ORA 2014, though, as Annison et al pointed out in 2015, “this recognition has not extended to the development of clearly differentiated services, commissioning or ring-fenced funding” (2015, p. 33). The current *Female Offender Strategy* (Ministry of Justice, 2018c) indicates progress with its acknowledgement that “short custodial sentences do not deliver the best results for female offenders”, many of whom “could be more successfully supported in the community where reoffending outcomes are better” (p. 6), and with its commitment to investing in community provision for women as well as a pilot for residential women’s centres. However, the intention to develop more gender-informed probation services may yet be undermined in the context of TR arrangements, where women’s centres have faced reduced interest and funding (and, in some cases, closure); where there have been widespread cuts to social services that are needed to address issues related to women’s offending (mental health, substance misuse, victimisation, etc); and where there remains a poorly defined plan for operationalisation or funding (Gelsthorpe & Padfield, 2018).

Due to the challenges arising from TR, the context for probation is set to change once again. In July 2018, Justice Secretary David Gauke announced a number of drastic revisions, including bringing forward the end date for CRC contracts from 2022 to 2020, aligning CRC areas with those of the NPS (creating 10 probation regions in England), and the launch of a new procurement exercise with an emphasis on the involvement of the voluntary sector in the delivery of services (Gauke, 2018). In Wales, where a motion for the political devolution of criminal justice was backed by the National Assembly in March 2018, a more uncompromising approach has been adopted, with TR and its attendant CRC being consigned to history. In June 2018 the Office of the Secretary of State for Wales announced the decision to return the supervision of all

offenders to a single National Probation Service to “better reflect the needs of communities in Wales... strengthen the supervision of offenders and increase confidence in community sentences” (Office of the Secretary of State for Wales, 2018).

A reliable evidence-base is crucial in building a foundation to support such systemic shifts and the need for it has been recognised in a number of initiatives. The *Strengthening Probation, Building Confidence* consultation was launched by the Justice Secretary in July 2018 with the stated intention of using “input...to introduce changes that strengthen our probation system and, in turn, help to break the cycle of reoffending” (Ministry of Justice, 2018d, p. 4). In the same month, former Chief Inspector of Probation, Lord Ramsbotham, launched a consultation for the Labour Party on the feasibility of returning probation to the public sector (Grierson, 2018). Meanwhile, a consultation had been announced in February of that year by the Commission on Justice in Wales, calling for evidence regarding five work streams, one of which being criminal justice, including policing, probation and prisons (Commission on Justice in Wales, 2018). Regardless of whether the decisions of policy makers and commissioners are being made at a systemic, organisational or practice level, evidence of *effectiveness* must be the driver for change.

The Risk-Need-Responsivity (RNR) model of offender rehabilitation is unparalleled in the criminal justice field in terms of its empirical validity, proposing and demonstrating principles for effective practice through a vast body of theory and research. Women, though, have too often been “correctional afterthoughts” (Ross & Fabiano, 1986) in the literature, not to mention policy and practice, due to their relatively small numbers in the offending population³, and because of some uncritical approaches which assume that what works for men is simply transferrable to women (Gelsthorpe & Hedderman, 2012). Furthermore, no matter how great a strategy may be, it will always fail if it is not implemented properly. Failings in implementation of the RNR model have led to deficits in service delivery for both male and female offenders and, in many cases, professional resistance to the RNR framework, leaving it vulnerable to criticism from

³ Even after exponential increases in the intervening twenty years or so, by 2011 they still only represented 25.4% of arrests in North America and 15.8% of arrests / 23.9% of convictions in England and Wales (Snyder & Mulako-Wangota, 2015; Ministry of Justice, 2012; Palmer et al, 2015), and accounted for around 15% of the probation caseload and less than 5% of the prison population in England and Wales in 2017 (Prison Reform Trust, 2017b, p. 34)

other academics in the field, especially proponents of alternative models. The focus for RNR advocates has therefore necessarily shifted from offender assessment and treatment to implementation, which, in spite of being a fundamental component of effective rehabilitative practices, remains relatively neglected in research, in policy and in practice.

In the vanguard of attempts to address this discrepancy is The Correctional Program Assessment Inventory (CPAI) (Gendreau et al, 2010). The CPAI enables a detailed evaluation of the extent to which interventions adhere to principles developed within the RNR framework, empowering agencies to articulate the work they do in terms of effective practice, identify their strengths and weaknesses, and develop strategies to improve services. Where implementation has been a weakness in the roll-out of promising innovations in offender management in Britain in the past, a crucial strength of the CPAI is that it addresses the quality of implementation, including management issues and organisational culture. Bonta et al's 2013 use of the CPAI hypothetically to evaluate a proposed pilot roll-out illustrates its full potential as an evaluability assessment tool.

The current study emanates from the work of the Swansea Service Evaluation Team (SSET), whose impetus is the current lack of an evidence-based methodology with which to measure quality of implementation in the British Isles and the resulting paucity of empirical evidence showing whether the principles of effective practice are being incorporated into rehabilitative work. This study's pilot of the most recent version of the CPAI in a women's project under TR in Wales and a gender-neutral probation service in the Bailiwick of Jersey provides a unique opportunity to investigate issues of implementation at a jurisdictional, as well as agency and practice levels, since Jersey is autonomous from the legal systems of England and Wales, and so has been independent of strategic and political developments in probation under the National Offender Management Service⁴, including TR.

⁴ NOMS, which was replaced by a new executive agency of the Ministry of Justice, Her Majesty's Prison and Probation Service (HMPPS), on 1st April 2017

As Leschied et al (2001) assert, implementation is key and no less so regarding the implementation of an evaluation tool. With the need for evaluation established and the correlation between positive CPAI scores and reduced recidivism now well evidenced in the US, the case for using the CPAI appears cut and dry. Less clear, however, is the place for the gender-neutral, Canadian CPAI in light of the uniquely British context, where there is a new statutory focus on addressing the needs of female offenders, a continuing argument that the TR agenda will have a particular, negative impact on women (Birkett, 2017; Gelsthorpe & Hedderman, 2012; Gomm, 2013), and where restructuring of probation services is currently planned for Wales. With the concept of evaluation responsivity at its core, this thesis aims to redress the balance and place context, culture and women front and centre when asking: *Can the CPAI be usefully developed to address the need for evaluation in the British Isles?*

Chapter One explores ‘what works’ for women as part of the powerful case in support of RNR adherence. It takes into account the meta-analytical literature largely generated in Canada and the US, the seminal work of Canadian behavioural psychologists in delineating and testing principles of effective practice and the enrichment of the empirical literature by the sometimes oppositional contributions of feminist theorists. The chapter proposes the possibility of a gender-responsive *and* RNR framework for intervention.

Chapter Two explores issues relating to implementation. It discusses more closely what the principles mean in practice (particularly the under-articulated Responsivity principle) and how misunderstanding and misapplication of the model has led to criticisms from service-users, practitioners and policy makers as well as academics proposing alternative models. The chapter argues that these criticisms are best addressed through a proper explication of the model and an integrative approach that draws both on the RNR and wider evidence base, especially the contributions of desistance theorists. Chapter Two presents the CPAI as a tool with which issues of implementation can be addressed in the real world and sets out the need for this study with a view to establishing the importance of contextual transposition and translation to women’s services and to the jurisdictions of the British Isles.

Chapter Three details the methodologies employed to pilot the CPAI in women's services in Wales and in Jersey Probation and Aftercare Service (JPACS) and gather data relating to its usefulness and potential for development in these contexts. The sample is introduced and described, ethics considerations are made under the guidance of the British Society of Criminology's (BSC) *Statement of Ethics* and the Economic and Social Research Council's (ESRC) *Framework for Research Ethics*, and the methodological limitations of the study are explored.

Chapter Four explores the findings of the CPAI-2010 evaluations of a women's project in Wales and JPACS's community supervision service by domain to discuss the key issues that were identified, contextualise these by reference to the literature relating to effective practice and explore the sample agencies' responses regarding the usefulness of findings and recommendations generated by the evaluation.

Chapter Five analyses participant perspectives of evaluation, examining similarities and differences across practitioners and management, different agencies, gender-responsive and gender-neutral services and different jurisdictions to reach conclusions about the experience, usefulness and limitations of CPAI evaluation.

Chapter Six explores the potential for use of the CPAI in different jurisdictional contexts, adaptations and modifications that might be made to enable its use in the British Isles, and the need to develop alternative, bespoke instruments for specific contexts, including the development of a gender-responsive evaluation instrument, the Swansea Service Evaluation Inventory – Women's Projects (SSEI-W).

Chapter Seven systematically addresses the research questions presented at the end of Chapter Two with a view to determining how knowledge about effective practice is best packaged and applied when attempting to improve practice in agencies.

Chapter Eight, the conclusion, focuses on the study's original contribution to the field by contextualising its findings in the existing literature and policy context and

evaluating its methodology and limitations in order to identify directions for future work.

CHAPTER ONE

From So Many Words to the Last Word?

Introduction

This chapter sets out the case for Risk-Need-Responsivity (RNR) adherence in work with women offenders. A brief survey of historical endeavours in the field illuminates the key issues in the debate as to what, if anything, comprises effective treatment for offending behaviour, a debate comprehensively referred to as “what works?”, for which concerns include whether interventions work, how interventions work and how to measure effectiveness. The emergent meta-analytical methodologies are explored, along with the resulting vast empirical base for rehabilitative work with offenders. The evidence for the fundamental principles of Risk, Need and Responsivity will be discussed in detail, with close focus on developments in both RNR and feminist criminologies regarding ‘what works’ for women offenders. The chapter concludes with Polaschek’s argument that the RNR model is not sufficiently comprehensive to facilitate effective implementation, reminding us that regardless of the breadth and validity of the evidence base, without the mechanisms to translate theory into practice, RNR may remain “only so many words on paper” (Leschied et al, 2001, p. xvii).

Does Anything Work?

It is a surprising and perhaps even shocking fact that our present day society is engaged in many activities which have no more support in terms of reliable evidence than the incantations of medicine men and the potions of witches. (Wilkins, 1969, p. 9).

Wilkins made his observation of rehabilitative practices five years before American sociologist Martinson asked the question “what works?” and, following a review of 231 intervention programme evaluation studies conducted between 1945 and 1967, famously answered that there is “very little reason to hope that we have...found a sure

way of reducing recidivism through rehabilitation” (Martinson, 1974, p. 49)⁵.

Martinson’s caveat that “It is just possible that some of our treatment programs *are* working to some extent, but that our research is so bad that it is incapable of telling” (1974, p. 49; [Martinson’s italics]) was largely overlooked and his paper nevertheless paved the way for a ‘nothing works’ policy response to crime on both sides of the Atlantic that proposed deterrence as an alternative to rehabilitation and therefore increased punitive control as practice.⁶ However, a number of dissenting reviews were undertaken that indicated positive outcomes for some programmes, including those that had been used to evidence that ‘nothing works’. This led to Martinson himself recanting his earlier position, acknowledging the errors in ‘nothing works’ reviews and reporting that the positive evidence from more recent work and re-analysis of his own data was “too overwhelming to ignore” (Martinson, 1979, p. 252).

Cullen and Gendreau (2001) point out the failings of Martinson’s contemporaries to scrutinise his findings scientifically. Had they done so, Cullen and Gendreau maintain, they would not have accepted the notion that ‘nothing works’ since, amongst other problems, of the 231 studies reviewed only 138 used some measure of recidivism as a criterion, of which fewer than 75 were valid rehabilitative interventions as opposed to ‘sanctions’ such as imprisonment. He also failed to take account of behavioural programmes in his categorisations in spite of the available evidence base at that time (Cullen & Gendreau, 2001, p. 322).

A significant criticism levelled against Martinson’s 1974 article is that positive evidence was ignored in the review. In fact, 40% to 60% of the studies reviewed by Martinson reported reduced recidivism for some offenders (Andrews & Bonta, 2010, p. 70), and women’s services were amongst those discounted. For example, Martinson refers to one

⁵ Martinson’s 1974 article “What Works? Questions and Answers About Prison Reform” pre-empted Lipton et al’s later and longer report, commissioned by NY State. This tempered Martinson’s pessimistic conclusions, finding instead that “the field of corrections has not *as yet* found satisfactory ways to reduce recidivism *by significant amounts*” (Lipton, Martinson & Wilks, 1975, p. 627; [My italics]).

⁶ In the United Kingdom objectives changed for probation. Treatment was relegated from its dominant role whilst the emphasis was placed on the use of alternatives to custody such as community service and day training centres (Bottoms & McWilliams, 1979; McGuire, 2004, p. 136; Pease et al, 1977; Smith, 1982; Vanstone, 2000, p. 173). Raynor (2012, pp. 933-934) explores this shift in detail, noting that in this new era of probation work, community service pilots were evaluated by the Home Office on the basis of the feasibility of implementation, court take-up and numbers diverted from custody. “In other words, the community service research agenda was about effects on systems rather than people, and a departure from the ‘treatment’ agenda”, whereas day training centres, which were based on a more rehabilitative ethos, were neglected in government research (Raynor, 2012, p. 933).

study of adolescent female inmates, stating that “these programs seem to work” under specific conditions, notably when practitioners are selected according to their relationship skills, their “empathy” and “non-possessive warmth” (Martinson, 1974, p. 32). Palmer’s findings from Martinson’s own dataset led him to conclude that some intervention programmes were successful while others failed, and to propose a new line of enquiry around “which methods work best for *which* types of offenders, and under *what* conditions or in what types of setting?” (Palmer, 1975, p. 150; [Palmer’s italics]). It was under this new line of enquiry that principles of effective intervention began to emerge from studies and narrative reviews such as the tellingly titled 1979 paper by prominent Canadian scholars Gendreau and Ross, *Bibliotherapy for Cynics*. As psychologists from a learning theory background (and so entirely at odds with the ‘nothing works’ concept (Gendreau & Ross, 1979, pp. 465-466)), they assessed 95 studies published between 1973 and 1975 and found a substantial number of interventions to be successful at reducing recidivism. Amongst these the programmes they identified as especially effective were behaviourally orientated, targeted criminogenic needs and were individualised to the offender participants (Gendreau & Ross, 1979), this latter point being somewhat under-emphasised in the ensuing literature, with groups including women offenders being largely disregarded in research focused on the predominant offender population, ie, white males of a lower socio-economic status.

These early narrative reviews, then, had distinct benefits in that the reviewer could begin to derive meaning from diverse research studies undertaken in a variety of settings by adding weight to individual studies and focusing on the richness of information therein. Conversely, as can be seen through the oppositional findings of Martinson and Palmer from the same data, narrative reviews were prone to subjectivity. The weaknesses of narrative reviews meant that, whilst they enabled the genesis of principles of effective practice, the results could not be reported with certainty since:

...traditional (narrative) reviews...tend to suffer from (a) selective inclusion of studies, (b) differential subjective weighting of studies in the interpretation of a set of findings, (c) misleading interpretation of study findings, (d) failure to examine other study characteristics as potential explanations for consistent results across studies, and (e) failure to examine the effect of moderator variables in relationship to the outcome variable. (Redondo et al, 1999, p. 252).

The development of the meta-analysis methodology in criminological studies enabled researchers to undertake statistical reviews of large bodies of research and to identify effective components of rehabilitation programmes with a greater degree of objectivity and validity than previously afforded by narrative reviews (Andrews & Bonta, 2010, p. 32).

Meta-Analysis

The introduction of meta-analysis as a strategy for reviewing research studies enabled researchers to measure ‘effect size’ – a quantitative estimate of the importance of the results – across any number of conceptually comparable studies with the same and/or different dependent variables. Meta-analytic computations take into account the effect size of every reviewed study, regardless of whether the primary study found treatment outcomes for a particular intervention to be statistically significant. This approach both reduces subjectivity and increases the statistical power of the exercise. The end result of the meta-analysis is the mean effect size – a “precise point estimate of the relationship of the treatment on the outcome measure across all studies” (Cullen & Gendreau, 2000, p. 134). This is most usually reported as Pearson’s r , which is used to compute a correlation between two continuous variables.⁷ Use of the Binomial Effect Size Display (Rosenthal & Rubin, 1982) enables effect size to be directly interpreted as the improvement percentage of treatment groups comparative to the controls (McGuire, 2004, p. 137).⁸

The results produced by meta-analyses are by no means definitive. The quality of a meta-analysis is dependent on the quality (and quantity) of the primary studies reviewed, and the quality of coding scheme used by the researcher. The conflation of differing outcome measures, the combination of studies that vary in quality and the lack of control for researcher error at either the primary study or meta-analytic level have all been cited as problematic areas (Lipsey & Wilson, 2001; Logan & Gaes, 1993; Pratt,

⁷ Though the phi coefficient is used where variables are dichotomous, and results can be translated into other effect size indexes such as the standardised mean difference or odds ratio.

⁸ This is computed from a base rate of 50%, so a correlation of $r=0.20$ would be reported as a 40% recidivism rate for the treatment group compared with 60% in the control group, a 20 percentage-point difference (Andrews & Bonta, 2010, pp. 14-15).

2002; see Lowenkamp, 2004, p. 14), but there are a number of ways in which confidence in meta-analysis results can be strengthened. Multivariate analysis can be used to ascertain whether the reported effect size has been influenced by the research design as well as to discern any moderating factors, such as offender risk level, that impact on treatment effectiveness. Statistical calculations like Orwin's Fail Safe N (a variant of Rosenthal's earlier formula) and Duval and Tweedie's Trim and Fill assess the potential for publication bias to have influenced the results. Also, replication of meta-analyses can be undertaken by other scholars to assess the coding decisions or sample of studies used and test the robustness of the original findings (Cullen & Gendreau, 2000, p. 135-136).

Nevertheless, Mair points out that, even where criticisms regarding the statistical validity of meta-analysis are answered, the lack of practical utility diminishes its "privileged" status since the organisation and context of the evaluated interventions are not addressed (2004, pp. 16-17). Lack of utility notwithstanding, Redondo et al were optimistic about the possibilities for meta-analysis, noting that:

Applied to the programme evaluation context, meta-analytic techniques achieve three main objectives: (a) to obtain a global index of programme effectiveness; (b) to determine the homogeneity of the results related to the global index; and (c) if homogeneity is not met, to search for study characteristics that may explain the variability in effect size. (Redondo et al, 1999, p. 252).

Unfortunately, women remain largely invisible in the majority of meta-analyses, which either examine men's programming alone or fail to provide a statistical breakdown by gender. Dowden and Andrews's 1999 meta-analysis (1999a) is a notable exception, with its focus on studies of programmes delivered exclusively or predominantly to women, as is Gobeil et al's 2016 meta-analytic review of women's services. Smith et al's 2009 meta-analysis is also significant in its examination of the predictive validity of an RNR assessment tool (the Level of Service Inventory – Revised (LSI-R)) for women offenders. Alongside the generalised (male) meta-analytical evidence base, as well as a comparatively small but growing number of reviews and studies specific to women's services, these contribute to an emergent literature that, finally, recognises and serves women.

Principles

North American treatment advocates have been predominant influences in formulating principles of effective practice through the interpretation of meta-analytic findings. Like the multi-disciplinary Personal, Interpersonal and Community-Reinforcement perspective (PIC-R) (see below) in which they have their roots, these principles are presented as gender-neutral, where gender exists alongside culture, class and ethnic origin as background contextual constants which “cannot account for variation in individual conduct within particular social arrangements” (Andrews & Bonta, 2010, p. 137). Since the evidence for these principles derives largely from adult male offender groups, their use in the treatment of women has been controversial and criticised (Bloom et al 2005; Hannah-Moffat, 2004; Sorbello et al 2002), and it is argued that gender-responsive principles and strategies must be grounded in women-centred perspectives and not the male-derived RNR or ‘what works’ evidence base. However, women-centred work in the field, such as Smith et al’s 2009 meta-analysis, which finds the RNR assessment model applicable to women offenders, and Blanchette and Brown’s 2006 study, which envisages the possibility of integrating a more nuanced understanding of gender into RNR assessment models, appears to have paved the way for a more collaborative response. Indeed, Hubbard and Matthews state that it would now be “tantamount to knowledge destruction” for either perspective to fail to take the other into account when developing correctional programming for girls and women (2008, p. 251), and the prominent critical feminist (and “persistent critic of...RNR” (Andrews & Bonta, 2010, p. 514)), Hannah-Moffat, herself now appears to endorse a gender *and* RNR-informed approach:

The absence of empirical evidence to support the effectiveness of alternative feminist approaches to risk assessment, treatment, and programming places managerially minded correctional agencies in a quandary as to how gender “ought to matter.” Gender is officially acknowledged as important, but an understanding of how to use gender to redesign correctional assessment, as well as how to organize and deliver correctional treatment, remains limited...research that is explicitly attentive to the literatures on gender (and race) as well as the expansive research literature on risk and correctional treatment can meaningfully advance theory, research, and policy. (Hannah-Moffat, 2009, p. 216).

It is in this spirit of collaboration and mutuality that the evidence relating to the RNR model is presented here. It is also worth noting that a great majority of primary studies

incorporated in meta-analyses derive from North America, though European countries are represented. Where possible this thesis will be attentive to the literature pertaining to Europe, and especially the British Isles, as well as that relating to gender-responsivity, since issues of jurisdictional context are also of central importance to the study.

Human Service

Broadly, the catalogue of meta-analyses indicates an average .10 effect size (a 10 percentage-point reduction in recidivism) (Lösel, 1995; 2012), with considerable variation in results across types of intervention studied. Suffice to say, with the effect being averaged across interventions ranging from behavioural to punitive approaches, effectiveness is diverse where punishment frequently produces null or negative effect sizes (ie, increases recidivism). McGuire notes that “if such effects were excluded from the overall calculation, the average for the remaining treatments would be higher than 0.10” and that the pattern of differences in effect sizes has been “much more informative than the grand mean” of .10 (McGuire, 2004, p. 146).

RNR proponents have adopted the term ‘the human service principle’ to describe the need to prioritise human, clinical and social services over conventional notions of criminal justice, where the judiciary seeks to use due process and the legal system to punish, deter, extract restoration and vengeance. They seek instead to address risk/need considerations, evidenced through meta-analyses that consistently find that solely punitive sanctions and/or increased control have negligible effects on reducing recidivism (Andrews & Bonta, 2010, pp. 46-47 & pp. 365-369).

Meta-analyses by Lipsey (1992), Lipsey & Wilson (1998), Andrews et al (1990a), Dowden & Andrews (1999a, 1999b, 2000), and Redondo, et al (1999) expose the fundamental inability of the “typical legal and judicial principles” (Andrews & Bonta 2010 p. 47) to address reoffending. Indeed, at the 41st Annual Meeting of the American Society of Criminology in 1989, Lipsey reported his findings that programmes based on deterrence theory (shock incarceration and “Scared Straight”) increased participants’ reoffending by a sobering 24%. Published in full in 1992, his meta-analysis reviewing

443 studies of young offender programmes suggested that in addition to an increase of recidivism where punishment or deterrence were the basis, the human service principle was evident in improved recidivism rates amongst participants of programmes with some service provision (Lipsey, 1992). In 1998 and using a subset of this original study, Lipsey and Wilson found that programmes that did not deliver services either had no effect or increased recidivism amongst serious juvenile offenders, whilst human service programmes reduced recidivism by 10-40% (Lipsey & Wilson, 1998). Similarly, Andrews et al (1990a) found that “non-service criminal sanctioning” was associated with a 7 percentage-point increase in reoffending and a series of meta-analytical reviews undertaken by Dowden and Andrews (1999a, 1999b and 2000) interrogated the growing databank of studies (at Carleton University, Ottawa) to determine the relevance of the human service principle for programmes engaging with female offenders, young offenders, and violent offenders. These consistently reported a substantially greater effect size in terms of reduced recidivism in programmes delivering services (.18 for female offenders, .13 for juvenile offenders and .12 where violent offending was used as the outcome measure) than in those delivering only sanctions (.01, -.02 and -.01 respectively). Using the Binomial Effect Size Display, Dowden & Andrews (1999a) report that women offenders in human service programmes had a 41% recidivism rate, compared to a 59% recidivism rate amongst women subjected to criminal sanctions alone. This finding is supported by the gender-responsive literature, which finds punitive and restrictive sanctions all too often inappropriately used, exacerbating the low-risk behaviours of vulnerable women (Covington & Bloom, 2003; Hannah-Moffat & Shaw, 2003; Holtfreter & Morash, 2003). The argument for women’s services in the community is at the very centre of the gender-responsive perspective. In the UK, Baroness Corston’s 2007 report into women in the criminal justice system has been strongly influential in its criticisms of the disproportionate punishment of women offenders through incarceration and illustrates the need to deliver holistic, women-centred community services⁹.

⁹ Malloch et al (2014) note that both “holistic” and “community” are highly contested concepts and that lack of definition results in superficial CJS responses. The authors distinguish between “community” as the basis for a transformative agenda that engages with issues of power and privilege (which they argue underpin processes of women’s criminalisation) and “community-based” interventions that take a multi-agency approach to address diverse needs. It is important to note that, whilst their argument is that the structural contexts of women’s offending can only be properly addressed through engagement with (and transformation of) the “deep structures” of society (the law, the CJS, the welfare system etc), their analysis of a community-based women’s service in Scotland finds that “changes are evident in the lives of women who have had the opportunity to access ‘holistic’, community-based interventions” (Malloch et al, 2014, p. 406).

In addition to the North American meta-analyses, Redondo et al's 1999 meta-analysis of European studies indicates that the human service principle is relevant in Europe and the British Isles. Their research took into account Thornton's 1987 British study of a penal/dissuasive intervention programme which revealed a slight increase in the recidivism rates of its young offender participants (an effect size of -.006), whereas they found varying reductions in recidivism (effect sizes ranging from .08 to .231) for programmes that delivered services.

Further to establishing the broad principle of human service, Andrews et al's 1990 meta-analysis of 80 programme evaluation studies was influential in drawing out an assortment of factors that could improve recidivism rates. Andrews et al coded programmes according to a hypothesis derived from clinical psychology (notably PIC-R) and the results of previous narrative reviews – namely, that “appropriate” intervention programmes are delivered to high-risk offenders, target criminogenic needs and use cognitive-behavioural modes of treatment matched with client need and learning style (Andrews et al, 1990a, p. 369). The effect size across all programmes was .10 (a 10 percentage-point reduction in recidivism) with considerable heterogeneity in effects – the effect size for “appropriate” programmes was .30, the effect size for “unspecified” programmes (those that could not be categorised due to lack of information) was considerably less at .13 and the effect size for “inappropriate” programmes (inconsistent with the hypothesised principles) was -.06, indicating a 6 percentage-point increase in recidivism where offender participants were engaged with “inappropriate” programmes, unfounded on the principles of Risk, Need and Responsivity (Andrews et al, 1990a).

Criticisms of Andrews et al's 1990 meta-analysis, notably from Lab & Whitehead (1990) and Logan & Gaes (1993), were, arguably, based on a misunderstanding of the way in which hypotheses derived from a subset of studies were tested against the main set. Their arguments that the findings were therefore tautological are answered by McGuire, who notes that the method “is an accepted confirmatory principle in research” and that “the criticisms also miss the point that there were independent theoretical grounds for predicting the pattern that was observed” (McGuire, 2004, p. 145).

In any case, these criticisms were subsequently dealt with by the findings of a large body of research that has consistently supported the findings of Andrews et al that programmes adhering to RNR principles reduce recidivism and has developed upon these principles to build a more comprehensive picture of ‘what works’ (see McGuire, pp. 140-141). Andrews et al’s original research has been extended twice and the databank now contains information on 374 effect sizes across 225 studies. The most recent findings (described in Bonta & Andrews, 2017, pp. 228-236) report an overall mean effect size of .08, a mild reduction in recidivism in treatment groups that nevertheless counters the ‘nothing works’ argument. Moreover, analysis of the results by the extent to which the studied programmes adhered to the hypothesised principles of effective practice confirms previous findings and reveals the extent of the correlation between programme integrity and treatment outcomes. Where treatment was inconsistent with each RNR principle the effect size indicates a 2 percentage-point increase in recidivism amongst the treatment group (-.02). Where one principle was adhered to the effect size is .02 (a 2 percentage-point decrease in recidivism for the treatment group). Where the treatment conforms to two principles the results are markedly more positive with an effect size of .18. Where a programme conforms to all three RNR principles the effect size is greatest, at .26. Expressed using the Binomial Effect Size Display, it can be reported that, where treatment adhered to all three RNR principles, the recidivism rate for the treatment group was 37% compared to a 63% recidivism rate in the control group – a clear positive effect (Bonta & Andrews, 2017, pp. 228-236).

The Risk Principle

Core to the Risk principle is the matching of offender risk to level of service, which comprises two elements: the acceptance that the probability of future offending can be estimated through assessment; and a concerted effort to deliver interventions based on that assessment. Essentially, the principle is concerned with delivering intensive and extensive services to high-risk offenders whilst minimal or no intervention is delivered to those at low risk of offending. Lowenkamp et al comprehensively tested the Risk principle in their 2006 study of 97 residential and community-based programmes of

intervention in Ohio State. In community-based settings, recidivism was reduced by 9 percentage-points where higher-risk offenders were targeted for intensive and extensive services, that is, over 66% or more of the participant sample were assessed as higher risk, more services were delivered to that group and they were subject to longer or equal length of programme compared to lower-risk offenders. This compares to a 10 percentage-point increase in recidivism where the programmes did not take the Risk principle into consideration. In residential programmes, effectiveness of adherence to the Risk principle equated to an 18 percentage-point reduction in recidivism while non-adherence resulted in a 1 percentage-point increase (Lowenkamp et al, 2006). Using a subset from this original sample to focus exclusively on female offenders, Lovins et al (2007) found that recidivism was considerably reduced in higher-risk women engaged in intensive residential treatment, who were 14 percentage-points less likely to be rearrested than those subject to the lesser interventions of parole or post-release control. Conversely, the recidivism of low-risk women increased by 13 percentage-points when exposed to residential services, leading the authors to conclude that low-risk women exposed to intensive treatment were three times more likely to be rearrested than those who were not.

In Britain, a 2004 retrospective analysis by Hollin et al of the effects on reconviction of offending behaviour programmes delivered in England and Wales found that appropriate allocation according to the Risk principle “has the optimum effect in terms of impact”. They found that low-risk offenders had a low rate of reconviction irrespective of participation in a programme, whilst there was an 11 percentage-point reduction in recidivism for programme-completers of an “appropriate” risk classification. Hollin et al controlled for gender in their analyses but noted that, with women making up only 9.6% of the experimental group and 18.5% of the comparison group, the study was “heavily weighted towards males” (p. 5).

Accounting for the Risk principle in meta-analyses has been challenging. In their 1990 research, Andrews et al found that effect size correlated with adherence to the Risk principle, though they were unable to calculate risk in several of the studies included in their analysis and their definition was questioned (Lowenkamp, 2006, pp. 18-19). It is,

of course, difficult to account for how risk is coded in the different primary studies included in a meta-analysis and Lipsey (2002) introduced an aggregate-sample approach, which involved coding each study according to the proportion of the sample group that had a criminal history. Lipsey's findings were negligible but nevertheless larger effect sizes were evidenced where the sample comprised of mainly "high-risk" offenders. Lipsey & Wilson (1998) went on to analyse data relating to studies on serious juvenile delinquents and found that community programmes delivering interventions to a participant group entirely comprised of higher-risk young offenders (those with a prior record) were more effective than those delivered to *mostly* higher-risk offenders. Most recently, Lipsey (2009) carried out a meta-analysis of 548 studies of juvenile delinquent programmes undertaken between 1958 and 2002. He found that, on average, interventions applied to "high-risk" young offenders produced larger reductions in recidivism than those applied to "low-risk" young offenders, though noted that 87% of the analysed studies were all or mostly male and only 4% were studies of all-female programmes.

Lipsey's aggregate sample approach has also been applied to the Carleton database, where the data continues to reflect the efficacy of intervention programmes targeted at higher-risk offenders as opposed to lower-risk offenders. In their 1998 re-analysis, Andrews and Bonta found a mean effect size of .11 for programmes delivered to higher-risk cases, while it was .02 for those engaging lower-risk offenders. Dowden and Andrews's series of meta-analyses found that programmes delivered to lower-risk female offenders increased recidivism (average effect size -.04), while those delivered to higher-risk female offenders had an average effect size of .19 (Dowden & Andrews, 1999a); programmes delivered to higher-risk young offenders were comparably more effective than those delivered to lower-risk young offenders (an average effect size of .13 compared to .03) (Dowden & Andrews, 1999b); and where violent offending was used as an outcome measure Dowden and Andrews found that adherence to the Risk principle was associated with a slight, but statistically insignificant, increase in effectiveness (average effect size 0.9 against .04). In their meta-analytic review of the databank now comprising 374 effect sizes, Andrews and Dowden reported that the Risk principle accounted for a 3 percentage-point reduction in recidivism where offenders

were engaged in non-adherent programmes, compared with a 10 percentage-point reduction where “high-risk” offenders were targeted (Andrews and Dowden, 2006). Unsurprisingly, the meta-analytic database comprised comparatively few studies of exclusively or predominantly female programmes (27.3%), however it was this sample that provided most support to the Risk principle. Programmes targeting “high-risk” women delivered a 30 percentage-point decrease in recidivism, compared to an 18 percentage-point increase where “low-risk” women were targeted.

This aggregate sample approach is flawed, in that it pre-supposes that all offenders with one or more previous convictions are “high risk”. Reminiscent of Mair’s caution regarding the practical utility of meta-analytical findings, it is certainly the case that coding for risk in this way is not transferrable to practice – ie, it would not be appropriate to subject offenders with one or more previous convictions to lengthier and more intensive interventions *purely on that basis*. Evidence indicates that other factors (such as an offender’s employment status and peer group) must be taken into account in effective risk assessment (see the *Need Principle* section below).

Whilst they cannot achieve the same broad scope as meta-analyses, individual studies such as Hollin et al (2004) and Lowenkamp et al (2006) are able to use more sophisticated methods to code for risk. Each accounted for powerful predictors of reconviction such as offender’s age and offending history using actuarial assessment scores – the Revised Offender Group Reconviction Scale (OGRS2) and a modified version of Hoffman’s 1994 Salient Factor Score (SFS) (which also accounts for unemployment and substance misuse) respectively. However, the use of actuarial tools to assess and classify women’s risk is controversial because they were developed using a characteristically male evidence base. According to some commentators, this may constitute discrimination, over-classifying women and/or leading to their needs being neglected. (Bloom & Covington, 2000; Funk, 1999; Hannah-Moffat & Shaw, 2001; Holtfreter & Cupp, 2007; Reisig et al, 2006; Van Voorhis & Presser, 2001). Despite these concerns, a number of studies and meta-analyses have demonstrated that risk classification instruments developed within the RNR framework and originally validated for use with male offenders (the Level of Service Inventory – Revised (LSI-R)

and its gender and culturally-informed update, the Level of Service/Case Management Inventory (LS/CMI)) are powerful predictors of recidivism for females in Canada, the US and Britain (Andrews et al, 2012; Palmer & Hollin, 2007; Rettinger & Andrews, 2010; Smith et al, 2009; Van Voorhis et al, 2010; Vose et al, 2009).

One of the most promising lines of enquiry that has emerged from tensions and developing recognitions between proponents of the Risk principle and advocates of gender-specificity is the potential to enhance the validity of gender-neutral assessment tools through increased attention to gender-specific factors and/or gender-informed weighting of neutral factors. Smith et al (2009) note that incorporating gender-specific factors into gender-neutral assessment “improved the prediction of recidivism meaningfully” (p. 202), and Andrews et al’s 2012 meta-analysis found that, as previously theorised by feminist thinkers (eg, Bloom et al, 2002), substance misuse “was more strongly related to the recidivism of female offenders” than males. These issues highlight that the different features of the RNR model do not operate as isolated factors. Consideration of risk assessment and categorisation must take into account matters of criminogenic need and Andrews et al conclude their analysis by calling for gender-informed positions regarding risk and need to be enhanced yet further through consideration of the Normative and Specific Responsivity principles of RNR (2012, p. 130).

It is also important to add that the Risk principle is limited in respect of offenders presenting the highest risk of reoffending, eg, psychopathic offenders, who have generally been viewed as incurable due to “biological deficit and/or early childhood experiences so severe that they are beyond hope” (Andrews & Bonta, 2010, p. 215). Lösel suggests that “an inverted U-shaped relationship between risk level and effect size may be realistic” (Lösel, 2012, p. 1001), given the difficulties in reducing recidivism amongst this group, though Andrews and Bonta (2010) note that “there is insufficient evidence to say whether treatment does or does not make a difference”. Moreover, the principles of effective practice have only recently been applied to psychopaths, who have historically been subjected to “poorly conceived interventions and milieu therapies” (Andrews & Bonta, 2010, pp. 215-216). It may yet prove to be the case that

the risk of psychopathic offenders can be successfully addressed through RNR evidence-based interventions such as Wong and Hare's psychopathy treatment programme in Canada (2005) and the Chromis programme in England and Wales (accredited in 2004), both of which have tentatively reported emerging positive outcomes (Olver et al, 2013; Tew & Atkinson, 2013). It is important to note, however, that attempts to treat personality-disordered offenders in England and Wales¹⁰ have in the past resulted in negligible outcomes and, according to O'Loughlin (2013), have been co-opted so as to contain or "warehouse" offenders on the basis that they haven't responded to treatment.

Clearly, some of the meta-analyses give resounding results in support of the Risk principle,¹¹ even, in some cases, indicating that recidivism is increased where low-risk offenders are targeted, but results are varied and some of the reviews mentioned here find little effect in delineating offenders in this way. Nevertheless, the most recent work and the bulk of previous endeavours lend empirical gravitas to the principle of risk and, notwithstanding the debates around gendered assessment and classification, it is certainly the case that protecting low-risk women from the harm of engagement with the Criminal Justice System (CJS) is a preoccupation of the RNR and gender-responsive literature alike.

The Need Principle

Research by Andrews, Bonta, Gendreau and others has identified eight significant risk factors associated with criminal behaviour. Previous accounts of these have distinguished between factors which prediction studies (mostly with general offenders) found to be "major predictor variables and indeed the major causal variables in the criminal behavior of individuals" (Andrews & Bonta, 2010, p. 55) – history of criminal behaviour; antisocial personality pattern; procriminal attitudes and antisocial associates – and factors found to be moderately associated with risk – deficits in family/marital circumstances; school/work; leisure/recreation and substance misuse (Andrews &

¹⁰ Under the aegis of the Dangerous and Severe Personality Disorder (DSPD) programme (most recently restructured as the Offender Personality Disorder Pathway(OPDP))

¹¹ Eg, Lowenkamp et al's 2005 review of drug courts found risk-principle adherent programmes to be doubly effective when applying Lipsey's aggregate method of coding for risk

Bonta, 2010, pp. 58-60). However, Andrews et al's 2012 research found the effect of gender on the validity of substance abuse as a predictive criterion to be highly statistically significant (a mean effect size of .46 for female offenders, compared with .17 for males), leading the authors to question whether it should be re-categorised as a major rather than moderate causal/predictor variable for women's offending in a gender-informed RNR model. In fact, demarcation between the "Big Four" and "Moderate Four" has not been borne out across numerous recent studies of different service-user groups, including general offenders (Olver et al, 2014), young offenders (Grieger & Hosser, 2014); mentally-disordered offenders (Bonta et al, 2014), racial minorities (Gutierrez et al, 2013) and drug offenders (Wooditch et al, 2014). This has resulted in Bonta and Andrews removing this distinction in their 2017 explication of the General Personality and Cognitive Social Learning (GPCSL) theory of criminal conduct, though they go on to argue that "nevertheless, the Central Eight risk/need factors are consistently observed across the meta-analytic reviews" (Bonta & Andrews, 2017, p. 44). Effective risk/need assessment tends, therefore, to be shaped around these factors, though some minor risk/need factors (such as mental disorder and seriousness of current offence) may also be taken into account. Amongst these factors, dynamic aspects or circumstances (in contrast with static factors such as an offender's age or criminal history) are referred to as criminogenic needs, that is, areas of an offender's life in which change is both *possible and needed* in order to reduce risk of re-offending. The Need principle describes the increased effectiveness of interventions that target criminogenic needs for treatment.

There is consensus across RNR and gender-responsive perspectives that women should be engaged with interventions that treat needs related to their offending behaviour, though the relevance of criminogenic needs derived from a male-centric evidence base is contentious (Blanchette & Brown, 2006, p. 84). In fact, several of the needs identified through women-centred investigation are already well accounted for in the RNR literature, eg, substance misuse, mental health issues and experience of abusive relationships (Bloom et al, 2002; Corston, 2007; Earle et. al, 2014; Hollin & Palmer, 2006; Koons et al, 1997). However, the extent and nature of the relationship between women offenders and so-called gender-neutral needs is widely disputed. It is certainly

the case that some needs are salient for women and not for men, though causal evidence is extremely limited. For example, in Britain, a Ministry of Justice Social Exclusion Task Force study found that 60% of women offenders presented with mental health issues compared with 36% of men (2009, p. 12) and, in a 2017 HM Inspectorate of Prisons survey, 65% of women, compared with 42% of men, self-reported mental health issues (p. 57), though the RNR literature suggests that mental health is only minimally predictive of recidivism. Andrews & Bonta (2010) examined eight meta-analyses for evidence of gender-specificity in areas generally identified by critics as being of particular importance to women – age, ethnicity, social class, emotional distress, abuse history, poverty and housing – and found that each factor was, in fact, minimally to mildly criminogenic for both males and females:

Table 1.1 Mean Predictive Validity Estimates for Gender Informed Risk/Need Factors by Gender: Overall Mean r was Averaged Over Mean Estimates Found in up to Eight Meta-Analyses (Andrews & Bonta, 2010, p. 68)

	Female	Male	A Gender-Neutral Factor?
Being Younger	.06	.15	Yes, but Male Salient
Being Non-White	.07	.06	Yes
Lower-Class Origins	.06	.07	Yes
Emotional Distress	.11	.12	Yes
Abuse History	.13	.06	Yes, but Female Salient
Poverty	.19	.16	Yes
Housing	.16	.16	Yes

Some critics propose additional needs that they identify as uniquely pertinent to women's pathways to offending, such as low self-esteem, victimisation and issues to do with sexuality and spirituality (Bloom, 1999; Bloom & Covington, 2001; Covington, 2000; Covington & Bloom, 1999; Koons et al, 1997, Van Voorhis & Presser, 2001). Whilst Andrews and Bonta acknowledge that addressing these issues may be very important in work with women, they refer to the lack of empirical evidence showing them to be causally linked to female offending, ie, criminogenic (2010, pp. 507-508). Referring to the general offender population, Cullen and Gendreau note that "the

research suggests that many factors thought to cause crime, such as low self-esteem¹², are unrelated or only weakly related to recidivism” (2000, p. 145), though it is possible, if not probable, that non-criminogenic, women-specific needs may become precursory to criminogenic need through interplay with gender-neutral and other needs (Hollin & Palmer, 2006, p. 191). Additionally, Andrews et al (2012, p. 117) note that “there is nothing in the evidence of similar risk factors that requires that women and men should offend in exactly the same way, from the same motivation or for the same reasons”. Far from being anti-socially motivated, women are often found to have offended due to “economic pressures, or... in the context of oppressive and coercive relationships with criminal men” (p. 117), whilst Langan and Pelissier (2001) found that men were more likely to report “hedonistic” substance misuse and women were more likely to report drug use “to alleviate physical or emotional pain” (p. 299; Hollin & Palmer, 2006, p. 185). Andrews et al (2012) argue that, rather than impacting on the extent to which criminogenic needs are predictive of offending, recognition of the differences and inequalities in gendered life-chances, opportunities and motivations provides the basis for empirical understanding of the way in which criminogenic needs are distributed and acquired (p. 117). Andrews and Bonta observe that programmes focusing on non-criminogenic needs are unlikely to impact on the future recidivism rates of participants, although addressing non-criminogenic needs in order to motivate and/or engage offenders in criminogenic work is both “important and valued”, as is the targeting of non-criminogenic needs on humanitarian grounds, but this “may not necessarily reduce recidivism” (Andrews & Bonta, 2010, p. 49).

Andrews and Dowden’s meta-analyses of the Carleton databank have provided empirical support for adherence to the Need principle. The most recent data available indicate a 1 percentage-point increase in recidivism amongst participants of

¹² However, it is important to differentiate between self-esteem, most commonly identified as Rosenberg’s “[expression of] an attitude of approval or disapproval toward oneself” (1965, p. 5) and self-efficacy, which Bandura defines as “people’s beliefs about their capabilities to produce effects” (1994, p.71). These concepts are often conflated, though evidence suggests that, unlike self-esteem, a lack of self-efficacy *is* in fact criminogenic (Andrews & Bonta, 2010, p. 58; Maruna, 2001, pp. 76-77).

programmes that fail to target predominantly criminogenic needs whilst adherence to the Need principle results in a 19 percentage-point reduction in recidivism (Andrews & Dowden, 2006; Bonta & Andrews, 2017, p. 230). Their previous meta-analysis focusing on female offenders (1999a) found a 4 percentage-point decrease in recidivism rates for non-adherent programmes and a 26 percentage-point decrease where criminogenic needs were targeted. Their meta-analysis of young offenders (1999b) also provides support for the Need principle, with a 1 percentage-point increase in recidivism evidenced amongst non-adherent programmes and a 22 percentage-point decrease for adherent programmes. It reviewed the impact of targeting individual criminogenic needs (those most frequently targeted amongst the programmes incorporated in the primary studies) as well as individual non-criminogenic needs (again, those most regularly represented amongst the sample). They found that “criminogenic needs are the key when developing correctional treatment programs”, since effect sizes indicating reductions in recidivism from 4 to 36 percentage-points were found for programmes targeting criminogenic needs, whereas non-criminogenic needs proved to be illegitimate targets, with results ranging from null effect to an 18 percentage-point increase in recidivism (Dowden & Andrews, 1999b). Notably, Lipsey did not address criminogenic needs in his 2009 meta-analysis. Andrews and Bonta (2010) question whether this may be due to difficulties in coding, since not all programmes incorporated in studies are explicit about targets for change. However, they point out that they have dealt with this in their own analyses by attributing the code of 0 or ‘not present’ to programmes with unspecified targets and lament that:

It is something of a problem that after 20 years of research on the RNR principles, the RNR team remains the only group of meta-analytic investigators attending to intermediate targets in a systematic manner. If we are off base in our coding, we would appreciate evidence regarding how the crucial issue of intermediate targets might be better explored. It is too important an issue for the primary investigators to not be specific about it and for meta-analysts to be paying so little attention to it. (Andrews & Bonta, 2010, p. 378).

Even so, the meta-analytic evidence presented by RNR advocates provides a convincing case for the Need principle. Gendreau et al’s 2002 report adds further weight through examination of the density of criminogenic needs targeted. They found a 31 percentage-point reduction in recidivism amongst participants of programmes targeting 4 to 6 more criminogenic needs, whilst programmes that targeted 1 to 3 more non-criminogenic

needs resulted in a slight increase in recidivism (Gendreau et al, 2002a). Nevertheless, the targeting of non-criminogenic needs is valued, which enables the RNR model to encompass approaches considered essential for women by paying heed to their lived experiences, personal and social environment and pathways to crime (Bloom, 2000; Chesney-Lind, 2000; Covington, 2000).

The Responsivity Principle

The idea of Responsivity comes from the acknowledgement that offenders need to be successfully engaged in programmes for them to be effective. RNR advocates propose two elements to the principle, which aims to describe the need for interventions to be delivered in such a way that the best available strategies (for reducing risk of reoffending) are employed at a level meaningful to the offender, taking account of individual attributes such as learning style and ability.

General Responsivity

The General Responsivity principle emerges from the Canadian RNR proponents' background in behavioural psychology. It recognises cognitive-behavioural and cognitive-social learning strategies as the most successful means to change behaviour. These approaches rely on techniques including skilled application of modelling, problem-solving, skill-building and cognitive restructuring. Several reviews have reinforced the notion that effective intervention programmes are behavioural in nature, and this has been proved across different populations of offenders including young offenders (Lipsey, 2009), female offenders (Dowden & Andrews, 1999a), violent offenders (Dowden & Andrews, 2000) and sex offenders (Gallagher et al, 2000). Dowden and Andrews (1999a) found that social-learning or cognitive-behavioural programmes incorporating modelling, role-playing, reinforcement and graduated practice as core correctional practices reduced the recidivism of their predominantly female programme participants (average effect size .27) while non-adherent programmes had significantly less impact (.08).

Andrews et al (1990a) and Lipsey & Wilson (1998) rated cognitive-behavioural programmes in particular to be “top tier” amongst interventions with regard to reducing recidivism (Lipsey, Landenburger & Wilson, 2007). Meta-analysis of the Carleton databank reveals a 23 percentage-point reduction in recidivism associated with behavioural, social learning and/or cognitive-behavioural strategies, compared to a 4 percentage-point reduction in recidivism for programmes that fail to adhere to the General Responsivity principle (Bonta & Andrews, 2017, p. 230). Lipsey (2009) found that cognitive-behavioural programmes were most successful at reducing the recidivism of young offender participants and behavioural programmes came second (though any difference in effect size between behavioural and some “non-adherent” programmes (group counselling and mentoring) was statistically insignificant). Redondo et al’s 1999 European meta-analysis evidences the General Responsivity principle’s significance in Britain and Europe, where findings indicate that “the most effective programmes are based on behavioural and cognitive-behavioural theoretical models” (p. 273).

Blanchette and Brown (2006) address some of the criticisms against using cognitive behaviourism with women, including arguments by Kendall (2002; 2004) and Hannah-Moffat & Shaw (2000) that empirically derived principles are intrinsically problematic due to the inability of meta-analyses to incorporate qualitative studies, some of which reach conclusions that contradict quantitative findings. Hannah-Moffat and Shaw note, for example, that qualitative research has found that women express a preference for individual, client-centred counselling, an approach that the meta-analytical literature finds ineffective (2000, p. 9). However, dismissing the extensive and persuasive meta-analytic evidence base for this reason amounts to knowledge destruction (Andrews & Bonta, 2010; Andrews & Wormith, 1989; Gendreau et al, 2002b). It seems entirely sensible to incorporate both quantitative and qualitative research in developing empirically-informed strategies for women, as Blanchette and Brown suggest (2006, p. 131).

Blanchette and Brown find Kendall’s (2002, 2004) anti-scientific approach, which also discounts the value of white, middle-class male psychologists’ contributions to the development of treatment strategies for culturally diverse women, to be of philosophical

importance. However, they re-affirm their “belief in the self-correcting nature of science” (2006, p. 131) as increasing numbers of feminist advocates and scientists contribute to the evidence base.

Kendall’s assertion that cognitive behaviourism dehumanises women by focusing on “otherness” (2002) is simply not demonstrable in findings from the growing research around women’s participation in cognitive-behavioural interventions (Blanchette & Brown, 2006, pp. 131-132).

Blanchette and Brown give most attention to responding to the argument that women are ill-served by treatment that targets individual change and not their ecology (personal relationships and social context). A number of critics cite structural factors (and not deficient thinking patterns) as defining a woman’s pathway into crime and argue that a more holistic approach must be taken with women (Covington & Bloom, 2003; Hannah-Moffat, 2004; Kendall, 2002, 2004; Pollack, 2005; Shaw & Hannah-Moffat, 2004). However, Blanchette and Brown respond in a number of ways (2006, pp. 127-128): they cite programmes that have taken a cognitive-behavioural *and* holistic approach, treating both the offender and other individuals in their immediate environment; they look to the RNR model, which includes advocacy brokerage and interagency communication to treat the environment *as well as* the individual; and they cite “renowned *feminist foremothers*” (p. 131) Worell and Remer as incorporating cognitive-behavioural techniques into feminist therapy, identifying it as “a comfortable choice” that “meet[s] the criteria for a feminist format” (Worell & Remer, 2003, p. 103); furthermore, they identify cognitive-behavioural work as an important tool in the kit required to empower women to overcome trauma, combat poverty and prevent future victimisation (Blanchette and Brown, 2006, pp. 129-130).

The appropriateness of the cognitive-behavioural therapeutic approach for women has nevertheless been drawn into question by some research studies that find no discernible difference in the reoffending rates of women engaged in cognitive-behavioural programmes. In England and Wales, research found positive outcomes for male offenders who completed the Enhanced Thinking Skills (ETS) and Think First (TF)

programmes (Hollin et al, 2008; McGuire et al, 2008; Palmer et al, 2007), though not for women (Cann, 2006; Palmer et al, 2015). In Palmer et al's 2015 analysis, they found no statistically significant difference in the reconviction rates between women who completed the programmes and the comparison group, though they noted an extremely low rate of completion, and increased recidivism rates amongst non-completers compared to the comparison group. Palmer et al theorise that a failure to target women-salient needs and respond to women-specific barriers to engagement (childcare and/or cultural issues, etc) may have contributed to the high number of non-completers and lack of treatment effect, concluding that their findings "offer support to the argument that interventions for women offenders need to be gender-responsive, although in this study it was not clear whether these programs 'failed' women in terms of their criminogenic needs, responsivity, or both" (p. 357). Given that some research has shown cognitive-behavioural programmes to be effective with women offenders (Spiropoulos et al, 2005; Gobeil et al, 2016), it is clear that additional attention needs to be paid to the Risk and Need principles regarding effective targeting, and to the Specific Responsivity principle to ensure that women are able to engage with treatment effectively.

Specific Responsivity

Specific Responsivity is arguably the least refined of the RNR principles and Andrews and Bonta (2010) acknowledge that it has not yet been the subject of thorough meta-analytical review. However, the idea that a programme's success will depend on the individual participant's ability to interact with it, and that this is likely to be affected by their cultural, social, cognitive and biological attributes, is widely accepted. The findings so far (though "scattered") are generally supportive of the principle (Andrews & Bonta, 2010, p. 46). Put simply, Specific Responsivity involves matching treatment to the individual offender and therefore necessitates assessment of, amongst others, learning style, conceptual level, diversity issues, psychopathy, depression, anxiety and motivation as well as the ability to adapt treatment best to encourage and enable participant engagement. Andrews et al's 1990 review sample provides some evidence that programmes delivering differential treatment according to the motivation, maturity,

empathy and conceptual level of participants has a positive impact in reducing recidivism (Andrews & Bonta, 2010, p. 51). Prochaska and DiClemente's work developing the psychotherapeutic transtheoretical model of change in the field of addictions acknowledged that clients' success would vary according to the stage of change that they had reached. This in turn informed the growth of "motivational interviewing" (MI) techniques, ie, methods practitioners can employ in order to increase participant motivation, which is influential in the delivery of offender intervention programmes (Miller & Rollnick, 2013; Prochaska & DiClemente, 1982).

Crucially, Andrews and Bonta identify two "fundamental dimensions of interpersonal exchanges" that underpin responsivity practice: The relationship principle and the structuring principle. An effective relationship between offender and practitioner (ie, within which "interpersonal influence by antecedent and consequent processes is greatest") is "characterized by open, warm, enthusiastic, and nonblaming communication, and by collaboration, mutual respect, liking and interest" (2010, p. 381). This has particular relevance to work with female offenders, where Pollack (1986) and Bloom et al (2003) found a prevalence of negative attitudes towards women amongst criminal justice professionals, while relational theory identifies meaningful connection with others as especially important to women, who will respond most positively to relationships characterised by empathy, empowerment and mutuality (Blanchette & Brown, 2006; Miller, 1986).

Establishing a positive relationship creates a therapeutic environment within which meaningful work can be undertaken to structure changed behaviours. The structuring principle refers to the use of "effective authority practices, antirriminal modeling, differential approval and disapproval, problem-solving, skill building, advocacy, brokerage, the structuring aspects of motivational interviewing, and cognitive restructuring" to influence thinking and behaviour (Andrews & Bonta, 2010, p. 381).

A growing body of international research has demonstrated that practitioner skills are associated with positive outcomes including reductions in reoffending (eg, Chadwick et al, 2015; Trotter, 2013; Ugwudike et al, 2018) and a study of probation practitioners in

the Bailiwick of Jersey demonstrates the importance of utilising both relationship and structuring skills in work with offenders. Having used a skills checklist to score practitioners during interview sessions with offenders, the authors found that higher skills ratings were significantly associated with lower reconvictions. The two-year reconviction rate for offenders interviewed by staff with below-median skill ratings (who tended to score low particularly on structuring skills) was 58%, while the use of a wider range of skills was associated with a 26% reconviction rate (Raynor et al, 2014).

It follows, then, that adherence to the RNR model not only involves the application of valid risk and need assessment tools, the use of screening tools to identify responsivity issues and the development of cognitive-behavioural interventions, but practitioners must also be trained, skilled and supervised to a high level. The Responsivity principle demands that a myriad of permutations and possibilities are allowed for in its applications. The targeting of non-criminogenic needs for humanitarian or engagement reasons, for example, is a responsivity issue.

A “gender-responsive framework” must address those needs that prevail especially amongst women offenders whilst enabling positive engagement with interventions. In 2006 Blanchette and Brown attempted to re-conceptualise the Responsivity principle with women-specific factors in mind. Treatment targets consistently represented in the literature as either impeding or facilitating women’s engagement with interventions are childcare and antenatal services; protection from abusive partners; physical and mental health care; safe and affordable housing; access to reliable transportation; and access to staff after hours (Ashley et al, 2003; Bloom et al, 2003; Richie, 2001; Wellisch, Anglin & Prendergast, 1993). Though there is no evidence to support its direct relationship with reduced recidivism (Dowden, 2005), Blanchette and Brown (2006, p. 120) highlight evidence that suggests empowerment is a valid responsivity factor, in that it helps to develop competencies and enables women to achieve independence (Austin et al, 1992; Blanchette & Eldjupovic-Guzina, 1998).

Blanchette and Brown (p. 121) also note differences in the learning and communication styles of women, who are more likely than males to be intrinsically motivated

(Severiens & ten Dam, 1994) and also to require enhanced occupational self-efficacy (Bandura, 1997), whilst the male propensity to talk and interrupt more than women in mixed-group settings has been explored as an impediment to women's treatment (Kennedy, 2004). Studies of substance abuse programmes have found that women perform better in single-gender versus mixed-gender formats (Ashley et al, 2003; Dahlgren & Willander, 1989; Lex, 1995).

Baroness Corston's 2007 report has been influential in the UK in demanding "a distinct, radically different, visibly-led, strategic, proportionate, holistic, woman-centred, integrated approach" to addressing female offending (Home Office, 2007, p. 79). In their research for the Fawcett Society, Gelsthorpe et al (2007) devised a nine-point list outlining desirable constituents of such an approach:

1. Be women-only to foster safety and a sense of community and to enable staff to develop expertise in work with women;
2. Integrate offenders with non-offenders so as to normalize women offenders' experiences and facilitate a supportive environment for learning;
3. Foster women's empowerment so they gain sufficient self-esteem to directly engage in problem-solving themselves, and feel motivated to seek appropriate employment;
4. Utilize ways of working with women which draw on what is known about their effective learning styles;
5. Take a holistic and practical stance to helping women to address social problems which may be linked to their offending;
6. Facilitate links with mainstream agencies, especially health, debt advice and counseling;
7. Have the capacity and flexibility to allow women to return to the centre or programme for 'top up' of continued support and development where required;
8. Ensure that women have a supportive milieu or mentor to whom they can turn when they have completed any offending-related programmes, since personal support is likely to be as important as any direct input addressing offending behaviour;
9. Provide women with practical help with transport and childcare so that they can maintain their involvement in the centre or programme.

(Gelsthorpe et al, 2007, p. 54).

Bourgon and Bonta (2014) highlight the difficulties of measuring gender-informed programmes' adherence to the Responsivity principle, which is complicated further because of the use of recidivism rates as a measure, an outcome that they argue is more relevant to the Need principle. They use Messina et al's 2010 study of a gender-responsive substance abuse programme as an example. The study found that women subject to the gender-responsive treatment programme (GRT) had better outcomes than those subject to a standard therapeutic community programme. However, Bourgon and Bonta (2014) observe that, whilst both programmes targeted substance misuse, the GRT

targeted more criminogenic needs (eg, family, peers and attitudes) than the standard programme (in addition to women-specific, non-criminogenic needs, such as memories of trauma and victimisation). Findings for the GRT also suggest greater adherence to the Risk principle through a treatment dosage effect whilst the use of cognitive-behavioural interventions in the GRT relates to the General Responsivity principle and the use of therapeutic approaches aimed at enhancing participants' learning relates to Specific Responsivity. As Bourgon and Bonta conclude, "dissecting the independent influence of the RNR principles and in particular the responsivity principle is a challenge".

Gobeil et al's 2016 meta-analysis of 37 studies of interventions for women offenders (incorporating 38 effect sizes) identified extensive, unexplained variability in findings and found no difference in outcomes for gender-neutral versus gender-informed approaches until only higher-quality studies (assessed using the Maryland Scientific Methods Scale) were included in their analysis, these 18 studies demonstrating a greater effect size for gender-responsive interventions (1.68, compared with 1.19 for gender-neutral interventions). Unable to explain this, Gobeil et al note their inability "to control for participant characteristics that could have helped to determine whether women in the sample studies had gender-specific risk or need factors, or gendered pathways to crime" (2016, p. 316). They go on to hypothesise that "this may have contributed to the heterogeneity in our overall results for gender-informed interventions", pointing to recent investigations which have found that gender-responsive approaches may be of increased effectiveness only where service-users have followed gendered pathways into the criminal justice system (Day et al, 2014; Saxena et al, 2014). They surmise that "findings related to gender-informed interventions may have been muted [in the meta-analysis] because this approach may not be optimal for *all* justice-involved women" (Gobeil et al, 2016, p. 316; [My italics]).

Whilst it is acknowledged that quantitative measurement of Specific Responsivity is hard to account for in meta-analyses and has therefore been largely neglected in the research (with notable exceptions such as Gobeil et al's 2016 review seeking to build upon the work of primary studies), services cannot afford to neglect it in practice. The

literature regarding women is emphatic in its insistence on programmes that are gender-responsive and the 2015 Ministry of Justice Data Lab re-offending analysis of women's centres throughout England provides empirical support for their effectiveness (a reduction of between 1 and 9 percentage points compared with a matched group), whilst studies such as Raynor et al's (2014) begin to demonstrate the impact of specific responsive practices on recidivism rates. Their use of scoring of observed practice has much to offer alongside qualitative research in providing an evidence base for responsive practice that can guide successful implementation of the model across different groups, individuals and structures.

The RNR Model

Empirical evidence for the Risk, Need and Responsivity principles has been presented here separately but it is important to recall Bourgon and Bonta's (2014) observation that dissecting each principle's individual influence is challenging, and that Andrews et al (2012) called for risk assessment to be enhanced through greater attention to the need and responsivity principles. In practice, the principles reinforce each other and must be implemented in unity. Engaging high-risk offenders in interventions that neglect their criminogenic needs will not address that risk. Work targeting criminogenic needs will be less effective where offenders are not engaged with services at an appropriate level. Without responding to the individual, there is simply no possibility for effective engagement with either risk assessment or interventions.

Furthermore, whilst the principles of Risk, Need and Responsivity are central to the Canadian academics' conceptualisation of rehabilitation, the RNR model incorporates additional "clinical" as well as "overarching" and "organizational" principles to position them at the heart of a unified approach. The full RNR model illustrated by Andrews & Bonta below (2010, pp. 46-47) accommodates findings from the body of empirical research generated through 'what works' endeavours as well as its theoretical roots in the GPCSL approach and, more specifically, the PIC-R perspective. This "broad cognitive social learning perspective on human conduct" incorporates elements of

Table 1.2 The Risk-Need-Responsivity (RNR) Model of Effective Correctional Assessment and Crime Prevention Services (Andrews & Bonta, 2010, pp. 46-47)

Overarching Principles	
1.	Respect for the Person and the Normative Context: Services are delivered with respect for the person, including respect for personal autonomy, being humane, ethical, just, legal, decent, and being otherwise normative. Some norms may vary within the agencies or the particular settings within which services are delivered. For example, agencies working with young offenders may be expected to show exceptional attention to education issues and to child protection. Mental health agencies may attend to issues of personal well-being. Some agencies working with female offenders may place a premium on attending to trauma and/or to parenting concerns.
2.	Psychological Theory: Base programs on an empirically solid psychological theory (a general personality and cognitive social learning approach is recommended).
3.	General Enhancement of Crime Prevention Services: The reduction of criminal victimization may be viewed as a legitimate objective of service agencies, including agencies within and outside of justice and corrections.
Core RNR Principles and Key Clinical Issues	
4.	Introduce Human Service: Introduce human service into the justice context. Do not rely on the sanction to bring about reduced offending. Do not rely on deterrence, restoration or other principles of justice.
5.	Risk: Match intensity of service with risk level of cases. Work with moderate and higher risk cases. Generally, avoid creating interactions of low-risk cases with higher-risk cases.
6.	Need: Target criminogenic needs predominantly. Move criminogenic needs in the direction of becoming strengths.
7.	General Responsivity: Employ behavioral, social learning and cognitive behavioral influence and skill building strategies.
8.	Specific Responsivity: Adapt the style and mode of service according to the setting of service and to relevant characteristics of individual offenders, such as their strengths, motivations, preferences, personality, age, gender, ethnicity, cultural identifications, and other factors. The evidence in regard to specific responsivity is generally favorable but very scattered, and it has yet to be subjected to a comprehensive meta-analysis. Some examples of specific responsivity considerations follow: <ul style="list-style-type: none"> a) When working with the weakly motivated: Build on strengths; reduce personal and situational barriers to full participation in treatment; establish high-quality relationships; deliver early and often on matters of personal interest and start where the person “is at.” b) Attend to the evidence in regard to age-, gender-, and culturally responsive services. c) Attend to the evidence in regard to differential treatment according to interpersonal maturity, interpersonal anxiety, cognitive skill levels, and the responsivity aspects of psychopathy. d) Consider the targeting of noncriminogenic needs for purposes of enhancing motivation, the reduction of distracting factors, and for reasons having to do with humanitarian and entitlement issues.
9.	Breadth (or Multimodal): Target a number of criminogenic needs relative to noncriminogenic needs.
10.	Strength: Assess strengths to enhance prediction and specific responsivity effects.
11.	Structured Assessment: <ul style="list-style-type: none"> a) Assessments of Strengths and Risk-Need-Specific Responsivity Factors: Employ structured and validated assessment instruments. b) Integrated Assessment and Intervention: Every intervention and contact should be informed by the assessments.
12.	Professional Discretion: Deviate from recommendations only for very specific reasons. For example, functional analysis may suggest that emotional distress is a risk/need factor for <i>this</i> person.
Organizational Principles: Settings, Staffing, and Management	
13.	Community-based: Community-based services are preferred but the principles of RNR also apply within residential and institutional settings.
14.	Core Correctional Staff Practices: Effectiveness of interventions is enhanced when delivered by therapists and staff with <i>high-quality relationship skills</i> in combination with <i>high-quality structuring skills</i> . Quality relationships are characterized as respectful, caring, enthusiastic, collaborative, and valuing of personal autonomy. Structuring practices include prosocial modeling, effective reinforcement and disapproval, skill building, problem-solving, effective use of authority, advocacy/brokerage, cognitive restructuring, and motivational interviewing. Motivational interviewing skills include both relationship and structuring aspects of effective practice.
15.	Management: Promote the selection, training and clinical supervision of staff according to RNR and introduce monitoring, feedback, and adjustment systems. Build systems and cultures supportive of effective practice and continuity of care. Some additional specific indicators of integrity include having program manuals available, monitoring of service process and intermediate changes, adequate dosage, and involving researchers in the design and delivery of service.

Sources: Andrews, 1995, 2001; Andrews et al, 1990b; Andrews & Bonta, 1994, 2006; Andrews et al, 1990a; Bonta & Andrews, 2007; Gendreau, 1996.

motivational and control theories and recognises social bonding in order to explain the mechanisms of criminal behaviour (Andrews & Bonta, 2010, p 141).

With regard to gender-specificity, there are two features of the model to which we must pay particular attention. Firstly, an overarching “respect for the person and the normative context”, where the authors place a duty on the agency to attend to issues such as gender, pointing out that a “premium” may be placed upon, for example, interventions aimed at addressing trauma and/or parenting concerns in women’s services. Secondly, the need to “attend to the evidence in regard to...gender-responsive services” is a feature of Specific Responsivity, though adherence must, of course, impact on the assessment of risk and targeting of need as well as the mode, style, influence strategies and service practices.

Introducing Integrity

Gendreau’s persuasive work of 1996 has had the effect of professionalising the RNR model by incorporating additional principles relating to the duration and intensity of delivery, the expectation of appropriate staff skills and management supervision, the use of booster sessions and multi-disciplinary practice embodied within a solid referral culture (Gendreau, 1996a). As early as his 1979 review with Ross, Gendreau was concerned with the issue of “therapeutic integrity”. They found that, amongst other things, many programmes had no theory base, bore no relevance to risk/need factors relating to offending, were conducted by untrained staff and periods of intervention were often inadequate. In short, the understanding that programme and treatment integrity are crucial became of the essence to Gendreau’s work and he and Andrews went on to refine the “Eight Principles of Effective Correctional Intervention” as the basis of a tool with which to evaluate programme integrity in real-world applications (Latessa, et al, 2002, p. 45). This tool, the Correctional Program Assessment Inventory (CPAI), has since been used extensively in the United States and Canada, where studies have found that the programmes with highest CPAI scores had greater measured effect in terms of reducing re-offending (Gray, 1997; Holsinger, 1999; Nesovic, 2003; Lowenkamp, 2004). A full description of the current CPAI-2010 and its validation

follows in Chapter Two, which will focus more closely on the broader issues of programme and treatment integrity, and evaluation, with particular emphasis on the current context in Britain. Suffice to say that the language of “evidence-based” rehabilitative practice has become common parlance amongst the criminal justice agencies delivering interventions on both sides of the Atlantic. However, it is as recent as 2002 that Latessa et al coined the term “correctional quackery” to describe a criminal justice culture where agencies, policy makers and practitioners continue to ignore or eschew empirical knowledge, maintaining Wilkins’s earlier diagnosis of “common sense” practice (Latessa et al, 2002).

Where there is a failure to implement principles in practice, ie, programme and treatment integrity are lacking, it is unsurprising that RNR has been the subject of criticism from academics, practitioners and offenders alike, since experience of its application and outcomes in the real world is varied and often negative, potentially causing the failure-resistance-failure cycle that produces competing models and conflicts in the field.

Not the ‘Last Word’

In concluding this introductory chapter, it may be useful to consider Polaschek’s critical appraisal of the RNR model (2012), in which she argues that deficits in implementation are the result of unrealistic expectations where the model is “not fully developed” as a multifactorial rehabilitation theory. Pointing out that “theory evaluation should be conducted with reference to the intended type or level of theory proposed” (p. 5), she locates the RNR model within Ward & Hudson’s (1998) meta-theoretical framework in which rehabilitation theories (originally informing sex offender treatment) are ascribed levels according to their depth of abstraction and breadth of comprehensiveness.

In Ward and Hudson’s (1998) meta-theoretical framework, theories in the top tier – Level I – are global and multi-factorial. They leave unspecified important details about the inner workings of the phenomenon they seek to explain. Level II theories instead deal with a single aetiological factor: specifying mechanisms, and describing how the factor interacts with other factors. Finally, Level III theories are local theories of the offence process itself. (Polaschek, 2012, p. 5).

In Polaschek's analysis, the RNR model does not meet the requirements of a Level I theory, chiefly on the grounds that it lacks sufficient Level II mechanisms to enable the design and implementation of effective programmes across all offender groups and settings. Her argument is that it provides no more than a Level I "theoretical framework", which is not yet developed enough to span theory and practice. Whilst this thesis will take into account the RNR model's "limited explanatory depth" (p. 8), it is written from a perspective that, like Polaschek, recognises its "unifying power and external consistency", "empirical validity" and "fertility and practical utility". Furthermore, Polaschek's central argument that RNR is not the "last word" on offender rehabilitation is interpreted here as a strength, since the model functions as a broad framework of incomparably empirically validated parameters within which dynamic, innovative and gender-responsive endeavours may flourish.

Chapter Two will take this analysis further using Polaschek's critique of the model's limitations as the impetus to fully interrogate issues around the misunderstanding, misapplication and potentially failed implementation of RNR. This prepares the ground for the chapter to discuss instruments and approaches that have the potential to bridge such a gap between theory and practice.

CHAPTER TWO

From Research to the Real World Part I - The Implementation Gap

Introduction

In theory there is no difference between theory and practice; in practice there is.
(Anonymous).

This chapter's intention is to explore the issues surrounding implementation in order to examine the role of the Correctional Program Assessment Inventory (CPAI) in addressing the gap between principle and practice, identified by Polaschek (2012) as having unintended and damaging consequences for the Risk-Need-Responsivity (RNR) model. Studies of adherence to the RNR principles reveal that very few programmes are implemented with fidelity and that fidelity is significantly correlated with effectiveness (Gendreau et al, 1999; Lowenkamp et al, 2010; Lowenkamp, Latessa & Smith, 2006). The increasing research base clearly demonstrates that, in the words of Bourgon et al (2010), "the effectiveness of RNR interventions diminishes when 'what works' knowledge is translated from well-controlled research projects to the 'real world' of everyday corrections" (Andrews & Bonta, 2010; Bourgon et al, 2010; Leschied et al, 2001). Of the interventions that Lipsey reviewed in his 1999 study, for example, the 'demonstration' pilot projects were on average twice as effective as their 'practical' counterparts (those subject to routine implementation), leading him to conclude that "rehabilitative programs of a practical 'real world' sort clearly can be effective; the challenge is to design and implement them so that they, in fact, are effective" (Lipsey, 1999, p. 641; Raynor, 2004, p. 318). It was as early as 1979 that Gendreau and Ross asked of correctional interventions:

To what extent do treatment personnel actually adhere to the principles and employ the techniques of the therapy they purport to provide? ... How much is treatment diluted in the correctional environment so that it becomes treatment in name only? (Gendreau & Ross, 1979, p. 467).

Taking Polaschek's 2012 critique of RNR as a starting point, this chapter will consider under-articulated aspects of the model that now require increased emphasis on practitioner skills and characteristics alongside supportive organisational structures and processes. It will be argued that dealing with these issues will go some way towards addressing contemporary conflicts between RNR and alternative models, making progress towards an integrative approach. The CPAI is presented as a tool with which issues of implementation can be addressed in the real world, facilitating the processes by which policy makers, commissioners, services and practitioners can operate with fidelity to the theory. Finally, the mindful application of the CPAI will be considered, with a view to establishing the importance of contextual transposition and translation to women's services and the jurisdictions of the British Isles around which this study revolves.

The Answer Within - Returning to Responsivity

Styles and modes of service are matched to the learning styles and abilities of offenders. A professional offers a type of service that is matched not only to criminogenic need but to those attributes and circumstances of cases that render cases likely to profit from that particular type of service. (Andrews et al, 1990b, p. 20).

Polaschek identifies the difficulties in implementing the RNR model as resulting from its limited "explanatory depth", particularly with regard to the Responsivity principle, originally defined by Andrews et al above. As Polaschek notes, the principle "contains much of what makes the application of the model both humane and effective" and yet it is "theoretically unsophisticated" and a "catch all category", in spite of its central role (2012, p. 8). The characteristics of offenders to which Andrews et al refer have largely been left open to interpretation, along with the way in which these may differ between individuals (Robinson & Crow, 2009, p. 100), and emphasis of research has most frequently been placed on client factors and not on practitioner attributes and skill level (Bourgon & Bonta, 2014).

According to Bourgon and Bonta (2014), the "vagueness of the original conceptualization of responsivity" has contributed to the principle becoming a "poor cousin" of risk and need in terms of research. Whilst the notion of General Responsivity

(ie, the use of cognitive-behavioural techniques) went on to become well-founded empirically (eg, Landenberger & Lipsey, 2005; Lipsey et al, 2007; Pearson et al, 2002; Redondo et al, 1999; Wilson et al, 2005), Specific Responsivity, relating to the need to *individualise* treatment to offender characteristics (Andrews et al, 1990b), has received far less attention, especially with regard to the way in which practitioner attributes and delivery impact outcomes (Bourgon & Bonta, 2014). As a result, McNeill points out that in practice “a complex and important principle [has been reduced] to the generalised induction: ‘Employ social learning and cognitive-behavioural strategies’” (McNeill, 2012, p. 5) and, as such, RNR programmes have been implemented as “one size fits all”. This has attracted considerable criticism due to the perceived failure of the RNR model adequately to consider the specific needs, values and issues of individual offenders (Ward & Maruna, 2007), though it is, in fact, in direct contradiction to the intellectual grounding of the Responsivity principle. As Bourgon and Bonta (2014) note, the psychotherapy literature from which the RNR principles were conceived “was well aware that no one mode of therapy or type of therapist was equally effective with all clients and that interaction of therapist, technique and client needed to be considered” (Clavert et al, 1988; Paul, 1967; Stein & Lambert, 1984).

Acknowledging the important consequences of underdevelopment of the Responsivity principle, Ogloff and Davis (2004) provide a more detailed overview of the factors affecting responsivity. They distinguish between idiographic factors, which are internal to the individual and include intellectual functioning, self-esteem and motivation level; and nomothetic factors, which are external and include staff characteristics, therapeutic relationships, environmental support, programme content and delivery (p. 233). Bourgon and Bonta’s 2014 “reconsideration” of the Responsivity principle sets out the requirements for adherence:

1. Know the client’s attributes that limit and/or facilitate the client’s learning style. These are bio/psycho/social factors. Examples of biological factors are race, age/interpersonal maturity and gender. Psychological factors may include intelligence, personality (e.g., impulsive; interpersonally insensitive), emotions (e.g. anxious), and poor motivation. Examples of social factors are poverty and culture. Some client attributes may be a mix of factors (e.g., a client from a racial minority has biological factors operating and perhaps social factors in the case of minorities living in poverty).

2. Create an optimal environment conducive to learning. Learning in this context is very broad; it is the acquisition of skills. To create such an environment, the first requirement is for the service provider to understand what client characteristics can affect his or her ability to learn. Next, the service provider creates the environment through his or her skills, language, and intervention activities that encourages client engagement in the learning activities and promotes efficient and effective client learning of what is being taught. (Bourgon & Bonta, 2014, p. 5).

Whilst more detailed conceptualisation of the Responsivity principle provides expectations of practice, Polaschek argues that real-world adherence is not achievable without the contributions of (Level II) theories which focus on a single aetiological factor – these are uniquely placed to be able to specify mechanisms, describe how the factor interacts with other factors and so enable the model to span theory and practice (Polaschek, 2012). The motivation level of an offender, for example, is an important idiographic factor affecting her or his ability to learn, and offenders may be uninterested or resistant to change, regardless of their apparent needs (McGuire, 2000, pp. 68-69). The Responsivity principle places responsibility on the service provider to use skills, language and activities to create an environment that encourages client engagement. Polaschek (2012) argues that the RNR model should take more detailed account of the considerable (Level II) developments made in the field, such as Motivational Interviewing (MI) (Miller & Rollnick, 2013). Indeed, Bourgon and Bonta's reconsideration of the Responsivity principle provides an account of how the skills, techniques and strategies of MI align with the "Core Correctional Practices" (CCPs) used by the "change agents" (ie, practitioners) of RNR – CCPs being the relationship and structuring skills outlined in Chapter One of this thesis (pp. 33-34) which are undertaken in the "MI spirit: a collaborative, person centred form of guiding clients" (Bourgon & Bonta, 2014; Miller & Rose, 2009). In England and Wales, where the requirement of an offender's consent to probation was abolished in 1997, Raynor (2014) highlights the importance of a therapeutic alliance between offender and practitioner, within which joint ownership of supervision can nevertheless be negotiated. In their study of mandated treatment, Skeem et al (2007) found that this alliance has a significant association with client resistance, motivation, cooperation and compliance, factors considered as "primary responsivity outcomes" in the RNR literature (Bourgon & Bonta, 2014). Dowden and Andrews's 2004 meta-analysis reviewed the Carleton databank with a new emphasis on staff skills and characteristics and found that "the

elements of CCP significantly enhanced the positive effects of clinically appropriate correctional treatment programs” (p. 212).

Mobilised by meta-analytical findings that community supervision had no impact on recidivism, Bonta et al observed 62 probation officers’ practice in supervision sessions in the Canadian province of Manitoba and found minimal adherence to the Risk, Need or Responsivity principles, citing this as evidence of the need for a “systematic and structured training agenda to help probation officers become more effective agents of change” (Bonta et al, 2008, p. 248). As such, a number of training initiatives undertaken in the US and Canada have focused on promoting the acquisition of effective skills and techniques including CCPs and MI, as well as maintenance of these skills, aiming for ongoing treatment integrity through coaching, supervision and mentoring. Evaluations of Effective Practices in Community Supervision (EPICS: Latessa et al, 2013), Proactive Community Supervision (PCS: Taxman et al, 2006), Staff Training Aimed at Reducing Re-arrest (STARR: C. Robinson et al, 2012) and the Strategic Training Initiative in Community Supervision (STICS: Bonta et al, 2011) have all indicated that practitioners who are highly trained are more effective in reducing recidivism through supervision practices. Preliminary findings from a similar initiative in England and Wales, the Skills for Effective Engagement, Development and Supervision programme (SEEDS, Rex & Hosking, 2013), demonstrate slightly higher levels of compliance among service-users working with trained practitioners (Sorsby et al, 2017). Finally, Chadwick et al’s 2015 meta-analysis, which included studies from Britain (Pearson et al, 2011; Raynor et al, 2014), the US (Latessa et al, 2013; Millson et al, 2010; Taxman, 2008; Lowenkamp et al, 2014), Canada (Bonta et al, 2011), and Australia (Trotter, 1996, 2013), found that practitioner training in CCPs is significantly correlated with reductions in recidivism.

Whilst there is an argument for the integration of theoretical advances in fields such as motivational work (McMurran, 2009), treatment readiness (Day et al, 2010) and compliance (Ugwudike, 2013), it is important that the RNR model remains as a framework of broad parameters within which *all* effective practice operates. It is not conceivable that every aspect of effective practice for each aetiological factor in every individual’s case in every service can be explicitly dealt with in the model. Thus, the

model acknowledges the importance of work undertaken in related fields and places an expectation on services to consider this in their work with offenders, enabling services to access the literature appropriate to their specific client group and to innovate for the benefit of their service-users. In this way, the RNR model organises theory into a framework for practice and facilitates approaches and instruments that absorb new knowledge, at the same time maintaining empirical integrity.

In order to preserve “treatment integrity”, practice must involve application of the theoretical model through professional use of interpersonal (relationship and structuring) skills (McGuire, 2004). The expectation is that services must be professionalised rather than parochial and must actively engage with the wider literature in order to develop and maintain effective practice (Latessa et al, 2002). It is a *sine qua non* that a professionalised service requires organisational and operational expertise.

Doing Without – Organisational Deficits

In their study of an effective Intensive Supervision Program (ISP) in New Jersey, Paparozzi & Gendreau (2005) introduced the measure of “organizational supportiveness”. They applied an adapted version of the Lederman Interview Schedule (1986), an inventory composed of organisational supportiveness indices consistent with elements of successful programmes identified by Petersilia (1990), such as a receptive environment, secure administrators, low staff turnover and sufficient resources. They found that, in addition to service delivery (whereby more treatment is provided to high-risk offenders) and a balanced professional orientation (as opposed to a bias/tendency towards either law enforcement or social casework), organisational supportiveness was an important factor in implementing the programme and reducing recidivism. They concluded that “[it] would be reasonable to speculate that the supportive offices were more diligent about seeing that their [probation officers] provided quality supervision to program participants” (Paparozzi & Gendreau, 2005, p. 462).

The importance of organisational structures is given priority in Taxman’s 2014 consideration of responsivity, in which she hypothesises an additional, third element to the principle – Systemic Responsivity. This amounts to an expectation that the

jurisdiction must provide a range of programmes capable of meeting the requirements of General and Specific Responsivity. If services are to offer effective interventions, she argues, practitioners must be properly resourced. In short, adherence to the Responsivity principle is not possible without due attention to the capacity and organisational structure of services. “Programme integrity” involves managers ensuring that resources are in place, including appropriately selected and trained staff, suitable accommodation and clinical supervision of staff. (McGuire, 2004).

Organisational efforts to ensure treatment integrity across different practice settings, ie, the development of structured and manualised accredited intervention programmes, have themselves fallen foul of systemic deficits in large-scale implementation. In her study of the manualisation of sex-offender treatment, Mann (2009) found that manualised interventions were linked to both effectiveness and the ability to replicate successful treatment *as long as* they were used as a tool to deliver interventions within a constructive therapeutic relationship between practitioner and service-user. In practice, however, they may be used as scripts from which inexperienced and undertrained staff deliver interventions (Mann, 2009; Polaschek, 2012). Representing the Correctional Services Accreditation Panel (CSAP) of England and Wales¹³, Maguire et al (2010) note that a programme’s effectiveness is not guaranteed simply by virtue of its accreditation – indeed, a well-designed programme is practically without value if it is not delivered as intended and in a “mindful and responsive style”. Furthermore:

...effectiveness of individual programmes – however well run by the facilitators in the ‘classroom’ – can be greatly enhanced or reduced by the quality of the environment in which they are delivered, the support accorded to programmes teams by senior management, the quality of offender selection, and the degree of preparation, support and follow up provided to participants. (Maguire et al, 2010, p. 44).

One notable example of the way in which poor implementation of RNR programmes can have far-reaching negative consequences for policy, practice, research and the model itself is the Pathfinder initiative. The development of the Pathfinder programmes in the 1990s was a large-scale government initiative to pilot and evaluate CSAP

¹³ The CSAP was replaced by the Correctional Services Advisory and Accreditation Panel in 2012.

accredited programmes¹⁴ as part of its Crime Reduction Programme (CRP) in England and Wales. Prison-based cognitive-behavioural programmes showed early promise (Friendship et al, 2002), though later evaluations found that the same programmes produced either null or limited effects (Cann et al, 2003; Falshaw et al, 2003), possibly due in part to methodological limitations including researchers' inability to establish a properly matched comparison group (Falshaw et al, 2003; Raynor & Robinson, 2009). Evaluations of community-based programmes aimed at addressing general offending behaviour (Hollin et al, 2004), basic skills (McMahon et al, 2004), enhancing community service through effective practices (Rex et al, 2004) and resettling short-term prisoners (Lewis et al, 2003) were also beset by methodological issues, including the lack of random allocation or appropriately matched comparison groups and concerns over recidivism data quality. Additionally, Raynor & Robinson note that "in all cases the experimental projects did not proceed exactly as planned, with knock on effects on the research designs. The typical result is not confirmation of what works, but rather a case for a better-informed second phase pilot with a further evaluation" (2009, p. 117). Indeed, implementation issues dominated the findings of the first phase of Pathfinder evaluations and, where investment in projects continued after the premature termination of the CRP (after three rather than the planned ten years), results were far more promising. The second phase of the Probation Resettlement Pathfinders, for example, addressed issues in programme design and implementation to make improvements in terms of integrity, participant numbers, data recording, levels of contact and quality of intervention. The second-phase evaluation produced conclusive findings with regards to reduced recidivism (Clancy et al, 2006), a clear endorsement both for the project and for evaluation research (Maguire, 2004). Hollis's 2007 reconviction analysis of offenders under probation in the community attending Offending Behaviour Programmes found that, far from null or limited effects, the re-offending rate for those sentenced to all programmes was statistically significantly lower (-10.3%) than that predicted, and the reduction in re-offending for programme completers was 25.8%.

Maguire (2004) observed that, having initially benefited from heightened political interest (and significant funding), the CRP was undermined by "inherent risks and

¹⁴ Maguire et al (2010) note that where programmes were unable to prove their effectiveness prior to roll-out, accreditation was provisional or granted with a proviso that a proper evaluation should be carried out (p. 48).

tensions that became increasingly prominent as circumstances (and the political climate) changed”:

While initially conceived as research-driven, it was ‘sold’ to politicians as contributing to the government’s challenging crime reduction targets, an aim which progressively took priority over research. It was over-ambitious in scale and raised unrealistic expectations of its outcomes. It suffered from major practical problems caused by unfeasible timescales, slow-moving bureaucratic procedures, and shortages of ‘capacity’. Low commitment to project integrity, cultural resistance among practitioners, and insufficient attention to differences between academics’ and policy makers’ understandings of research, also contributed to its problems. (Maguire, 2004, pp. 213-214).

Unrealistic targets for the Pathfinder programmes were set through negotiation with the Treasury without taking into account offender demographics, pressuring the Probation Service to achieve 30,000 programme completions between 2000 and 2004 and driving a fast-track roll-out by an inexperienced service, which resulted in staff resistance, union opposition and some academic misgivings about accredited programmes (Maguire et al, 2010; Mair, 2000, 2004; Napo, 2001; Raynor, 2003, 2004; Raynor & Robinson, 2009). The rush to roll out programmes also resulted in understaffing, communication problems between staff and agencies, and alienation of practitioners by the top-down management style seen as necessary to progress the roll out quickly (Raynor, 2004; Raynor & Robinson, 2009). Additionally, shortcuts were made. Instead of service-users being referred to programmes based on assessment of their risk and need, they tended to be selected to meet statistical targets. Hollin et al (2004) found that many service-users were selected from outside the offending behaviour programmes’ target group, but that those within it were more likely to do well in terms of reduced reoffending. However, attrition rates across the programmes were so high that the increased recidivism rates of non-completers effectively negated any positive impact on programme-completers in the findings. It seems likely that the high attrition rate was due, at least in part, to the poor matching of offenders to service. (Raynor, 2004; Raynor & Robinson, 2009).

Maguire (2004) expressed his concern that the negative experiences of the CRP may “lead policy makers to adopt crude assumptions associated with the ‘nothing works’ era and consequently, for example, to withdraw investments in the more expensive

interventions...simply on the grounds that they are ‘unproven’ and there are cheaper alternatives” (p. 233), concluding that:

...the ideal of ‘evidence-based policy’ may be more effectively pursued as a quiet iterative process over the longer term, rather than through a risky investment in one high profile and rapidly implemented ‘programme’ which promises more than it can guarantee to deliver. (Maguire, 2004, p. 214).

Van Voorhis and Brown (1996) caution that a roll-out must be well conceived to justify programme evaluation, otherwise results may erroneously lead policy makers and others to write off programmes as ineffective (which then may be used as evidence for alternative approaches such as increased imprisonment or the privatisation of services) and Maruna and LeBel observe that negative experiences of RNR enterprises in England and Wales such as the CRP and Pathfinders have “opened a door” for other approaches to switch focus from ‘what works’ to issues of implementation, delivery and effective practice or “how change works” (2010, p. 69).

Fighting Over Common Ground?

Desistance theorist McNeill’s own rejection of the RNR model came directly as a result of the processes by which it was implemented in England and Wales. The roll-out of RNR programmes was undertaken within the wider context of increased managerialism, resulting in services that were, to all intents and purposes, RNR in name only, lacking both the intention and capability to individualise treatment: “...the significance of the ‘human’ in human services somehow got lost” (McNeill et al, 2010, p. 7). The inadequate articulation of responsivity combines with a failure to conceptualise the model as separate from its application to produce and perpetuate the “contemporary conflict” between the models of RNR and, variously, the Good Lives model (GLM), strengths-emphasising approaches and desistance-based corrections (McNeill, 2012).

Ward and colleagues propose GLM as an alternative to RNR, which they perceive as being too narrowly focused on addressing the risks of service-users. They argue that services should instead be strengths-based, responding to service-users’ “core aspirations and interests” in order to provide them with “the internal and external resources to live rewarding and offence-free lives” (Ward, 2010, p. 41).

Notwithstanding concerns about whether increasing an offender's wellbeing (as opposed to reducing reoffending) is a valid end-goal for rehabilitative efforts, Ogloff and Davis (2004) point out that Ward et al's critiques of RNR (eg, Ward, 2002; Ward & Stewart, 2003; Ward et al, 2012) reflect their narrow construal of the model. A number of the psychologically relevant factors identified by GLM (such as marital relationship, employment skills, etc) are addressed by the Need principle of RNR and the GLM focus on broad human needs and self-determination is addressed through full implementation of the Responsivity principle (Andrews et al, 2011; Ogloff & Davis, 2004; Wormith et al, 2012). Issues around framing interventions as addressing deficits rather than promoting strengths are important but need not constitute a "paradigm conflict". Ogloff and Davis (2004) suggest that dividing responsivity into two levels – responsivity impediments and responsivity enhancements – would go much of the way towards accommodating the GLM perspective within the RNR framework.

McNeill's summary of the major themes emerging from desistance-focused work highlights that, far from contradicting the RNR literature, the desistance perspective in fact enriches understanding of the processes by which the RNR principles can be implemented in practice (2012, pp. 12-13):

1. Since desistance is an inherently individualised and subjective process, approaches to intervention must accommodate and exploit issues of identity and diversity. (Weaver & McNeill, 2010).
2. The development and maintenance not just of motivation but also of hope become key tasks for workers (Farrall & Calverley, 2006).
3. Desistance can only be understood within the context of human relationships; not just relationships between workers and offenders (though these matter a great deal) but also between offenders and those who matter to them (Burnett & McNeill, 2005; McNeill, 2006).
4. Although we tend to focus on offenders' risks and needs, they also have strengths and resources that they can use to overcome obstacles to desistance – both personal strengths and resources and resources in their social networks. Supervision needs to support and develop these capacities (Maruna & LeBel, 2003).
5. Since desistance is about discovering agency, interventions need to encourage and respect self-determination; this means working with offenders not on them (McCulloch, 2005; McNeill, 2006).
6. Interventions should work on 'human capital' (or developing offenders' capacities and skills) (Maguire & Raynor, 2006), but also on 'social capital' (developing relationships and networks that generate opportunities) (Farrall, 2002, 2004; McNeill & Maruna, 2007; McNeill & Whyte, 2007).

As discussed, responsivity is successfully operationalised through attention to individualised issues (such as identity, diversity, motivation, etc) within an effective therapeutic alliance and the desistance-focus on the strengths and resources of service-users is consistent with the capacity-building focus of RNR theory. Andrews and Bonta (2010) identify promising targets for change as including, for example, “promoting family affection/communication” and “promoting identification/ association with antirriminal role models” (p. 504).

Polaschek (2012) sees strengths-based approaches as making an important contribution to consideration of the emphasis of interventions, which must give service-users “reasons to want to...change [and] not just the capacities to do so” (p. 8). A focus on desistance has led, for example, to a greater understanding of how women’s paths to self-efficacy may be impacted upon by identity. Giordano et al’s 2002 research found that women are more likely to identify parenthood as the catalyst for their desistance from crime than are men. Rumgay’s (2004) theory is that the role of “mother” is attractive (by virtue of its socially-approved status) and that, should a woman recognise it as an opportunity to desist from anti-social behaviour, the nature of “motherhood” as a common identity provides her with a “script” by which she can enact and (alongside further skill acquisition) perpetuate this alternative, pro-social role. The desistance perspective endows practitioners with a greater knowledge base from which to respond to those characteristics of an individual that may provide opportunities for change. The literature around core correctional practices identifies the skills with which to respond. For example, effective reinforcement is crucial since “newly acquired identities need validation by those who witness them” (Rumgay, 2004, p. 416).

Porporino (2010) proposes an integrative approach since, without the evidence base of RNR, desistance theory and research “lacks any sort of organised practice framework” (p. 61). McNeill cautions that it is common for practitioners to interpret desistance approaches as meaning that “they ‘just’ need to have good relationships with offenders and to offer a bit of practical help” (2010, pp. 7-8). Conversely, as this chapter has argued, RNR’s incomparable empirical framework

lacks the “rich...descriptive analysis of the forces and influences that can underpin offender change” which desistance-focused efforts provide (Porporino, 2010, p. 61). As McNeill contends, there is no substantial rift between the desistance-based approach and RNR theory, though “...the tensions between narrow interpretations of ‘Evidence-Based Practice’, ...the managerialized implementation of ‘what works’ and a desistance-based perspective are indeed serious and real” (2012, p. 14). To move towards an effective and integrative approach we must turn our attention to implementation – the transfer of knowledge into practice.

Knowledge Transfer and the CPAI

As has been seen, knowledge transfer is not simply a case of disseminating research findings. The processes by which human services can absorb and implement the evidence base must be the focus of efforts to improve practice. To this end, reasons that programmes may not be implemented as designed (and that evaluations may not measure programme effectiveness) have been identified comprehensively by Van Voorhis et al (1995):

- a. The program is not grounded in the knowledge base of the discipline and therefore is not utilizing program strategies that are either empirically or theoretically sound.
- b. The program has chosen a strong program design but is not *operating* according to the design or the clinical dimensions of an intervention (e.g., a social learning program which does not incorporate principles of good role modeling).
- c. There is a lack of specificity in the program's design; we are delivering some global treatment, e.g., counseling, case management, job skills (which can mean different things to different staff), rather than an explicit treatment process that is known to be effective.
- d. Staff members do not understand the intervention. They do not wish to cooperate, or they do not have or do not follow a treatment manual.
- e. The organizational and political climate is too confusing and is not conducive to successful implementation.
- f. Budget cuts create a situation where we are asked to do the impossible-keep the program without the funds.
- g. The "dosage" is inadequate. It may be a good intervention, but the amount of time the client participates in the program is insufficient.
- h. Clients did not or could not attend.

i. The program is too turbulent, undergoing several changes during the evaluation. Evaluators do not know what they tested.

j. Evaluators fail to measure various components of service delivery (e.g., contact hours, attendance, content analysis). The evaluation puts forward assumptions about whether or not the services were delivered.

(Van Voorhis et al, 1995, p. 20).

The detrimental impacts of prematurely conducting outcome evaluations on non-adherent services are itemised by Van Voorhis and Brown, who predicted loss of resources, the collapse of funding chains and the eventual withdrawal of social programmes altogether (1996, pp. 3-4).

In the field of corrections, Gendreau and Andrews turned their attention to implementation in their 1979 review, finding that initiating and maintaining a programme successfully was correlated with how contact with the service was made; knowledge of the service environment; the stability of the organisation; shared values between stakeholders and consultants; active involvement of practitioners and staff; and sustainable funding. At the same time, in response to poor quality evaluations of a number of US federal social and health programmes undertaken in the 1970s, Wholey (1979) developed a model for evaluability assessment. In order to overcome implementation issues (including unclear or unrealistic objectives, differing views of purpose amongst stakeholders, lack of a coherent management structure and illogical programme design) leading to no-effect findings, Wholey proposed preparing programmes for outcome evaluation using an eight-step process summarised by Trevisan and Walser:

(1) Define the program to be evaluated; (2) collect information on the intended program through document review and stakeholder interviews; (3) develop a program model; (4) analyze the extent to which stakeholders have identified goals, objectives, activities, and so forth in measurable terms; (5) collect information on program reality through site visits and document review; (6) synthesize findings to determine the plausibility of program goals; (7) identify options for evaluation and management; and (8) present conclusions and recommendations to management. (Trevisan & Walser, 2015, p. 2).

The model has been developed further in the field of corrections as the RNR principles emerging from the ‘what works’ enterprise have enabled an approach which also takes

into account whether a programme is using effective strategies to target those needs most likely to increase the chances of recidivism. One recent interesting attempt to solve the issue of implementation was Bonta et al's 2013 study where they used the CPAI to assess the likelihood of STICS being rolled out successfully. The CPAI is an evaluability assessment tool that incorporates the principles of effective intervention to measure correctional programme quality. In its most recent version the CPAI-2010 (Gendreau et al, 2010) comprises 143 items in nine domains, eight of which are concerned with whether different aspects of practice adhere to the RNR principles and whether the service attends to programme and treatment integrity through built-in quality assurance measures. Chapter One of this thesis gives a detailed overview of the evidence base for items relating to the RNR principles. CPAI domains and items relating to organisational factors have been formulated based on a wide range of influential findings from a number of fields including the technology transfer literature (eg, Backer et al, 1995), the general management literature, the social psychology of persuasion literature (Gendreau, 1996) and the CPAI developers' own experiences of effective and ineffective programmes, which have been further built upon and refined following Gendreau and Andrews's 1979 review (Gendreau et al, 1999).

Domain A, Programme Demographics, contextualises the evaluation and is descriptive (of elements such as referral sources and general structure and physical set up) rather than assessed. The eight assessed domains are articulated as "the principles of effective correctional treatment" by Gendreau et al in their 2004 summary of key elements (pp. 27-28):

B. Organizational Culture.

The organization has a culture that is receptive to implementing new ideas and has a well-articulated code of ethics. A history of responding to new initiations and coping with problematic issues in a timely manner is evident, as is a proactive orientation to problem solving. Organizational harmony is reflected in low staff turnover, frequent in-service training, and within-house sharing of information.

C. Program Implementation/Maintenance

The implementation of the program is based upon individual level survey data on the need for the service and a thorough review of the relevant treatment literatures. Implementation occurs during a period when the organization does not face contentious issues (e.g., fiscal, staffing levels, stakeholder concerns) that might jeopardize the project.

D. Management/Staff Characteristics

The director of the program has a post-B.A.-level degree in the helping professions and has several years experience [sic]working in offender treatment programs. He/she trains and supervises staff and provides direct service to offenders. The majority of staff involved in direct service delivery have undergraduate degrees in the helping professions and clinical experience working with offenders. Staff are hired on relationship and skill factors that enhance the integrity of the therapeutic relationship; moreover, they have a belief in the value of rehabilitation and confidence in their ability (i.e., self efficacy) to deliver quality service. Staff are reassessed and retrained regularly with respect to their clinical skills.

E. Client Risk/Need Practices

Offenders are assessed on a risk instrument that not only has adequate predictive validities but also contains a wide range of criminogenic need factors. The assessment also takes into account the responsivity of offenders to different styles and modes of service. Risk level is routinely assessed over time (e.g., every three to six months) in order to monitor changes in risk/need levels that will, in turn, have a significant impact on case management strategies.

F. Program Characteristics

The single most important characteristic of effective programs is that they are behaviorally oriented (general responsivity) and target the criminogenic needs of higher risk offenders... The program has a manual that describes the theory and data justifying the program and a curriculum that details the discrete steps to be followed in presenting the manual... Offenders spend at least 40% of their program time in acquiring prosocial skills. The ratio of reinforcers to punishers is 4 to 1 or more, and completion criteria are explicit. Relapse prevention strategy methods are extended to offenders after completion of the initial treatment phase.

G. Core Correctional Practice

Program therapists engage in the following therapeutic practices:

- Anti-criminal modeling;
- Effective reinforcement and disapproval;
- Problem-solving techniques;
- Structured learning procedures for skill building;
- Effective use of authority;
- Cognitive self-change;
- Relationship practices; and
- Motivational interviewing.

H. Inter-Agency Communication

The agency aggressively makes referrals and advocates for offenders (i.e., advocacy brokerage) so that they receive high-quality services in the community.

I. Evaluation

The agency routinely conducts program audits, consumer satisfaction surveys, evaluations of changes in criminogenic need, and follow-ups of recidivism rates. The effectiveness of the program is evaluated by comparing the recidivism rates of the treatment group with a risk-control no-treatment comparison group or a risk-control comparison group that has undergone an alternate or minimal treatment.

Using data gathered through observations of practice, document and file-reads and interviews, the CPAI quantifies a service's adherence to its individual items and domains using scoring and categorisation (Unsatisfactory, Satisfactory or Very Satisfactory according to percentage score – a full description of the scoring protocol for the CPAI-2010 is provided in Chapter Three). In this way, services' progress and comparative performance can be tracked without undertaking (lengthy and costly) outcome evaluations prematurely. There is added value for researchers who have been able to examine the correlation between programme integrity and recidivism, an endeavour that has supported the validity of the CPAI itself with scores showing significant association with programme effectiveness. An abbreviated version of the CPAI was applied by Gray (1997) to 67 evaluation studies of community programmes and a correlation (r) of .41 between CPAI scores and reductions in recidivism was found. Holsinger's 1999 study of nine juvenile correctional facilities in Ohio found CPAI scores to be strongly predictive of post-release recidivism. Using a modified version of the CPAI in her study of 173 programmes across the province of Ontario, Nesovic (2003) also found a significant correlation ($r = .50$) between scores and reductions in recidivism. Lowenkamp (2004) reported similar results (a correlation of $r = .41$) in his study of 38 halfway houses in Ohio. Lowenkamp et al's 2006 study of the same halfway house services again found that scores consistently and significantly correlated with outcomes (correlations ranging from $r = .26$ to .42 depending on the outcome measure). Using the CPAI scoring categories as a measure, Lowenkamp et al's 2010 examination of juvenile community correctional facilities in Ohio found that programmes categorised as "Satisfactory" had an average effect size of $r = .06$, while those rated as "Very Satisfactory" achieved greater reductions in recidivism ($r = .14$).

Van Voorhis and Brown (1996, Table 1) itemise the benefits of evaluability assessments which, as well as assisting goal clarification and resource allocation, etc, facilitate the development of a "learning organization". A full CPAI report is provided to the assessed service, which identifies strengths and weaknesses and makes recommendations for improvements to organisation and practice. The real-world impact of the CPAI on service performance is illustrated by Latessa's account of the State of Oklahoma's response to assessment. In an initial review by researchers from the

University of Cincinnati, only 9% of 29 programmes delivered across the state achieved a Satisfactory CPAI score. The state's response was to embed the CPAI into institutional processes, task programme managers with producing action plans to address deficiencies and implement these prior to reassessment. As a result, 79% of programmes in the state achieved CPAI scores of Satisfactory and above in subsequent assessments (Latessa, 2004, pp. 554-555).

As well as shaping improvements to services, the CPAI has fulfilled a role in the planning stages of programme development. When moving from pilot project to wide-scale implementation of the STICS initiative, Bonta et al (2013) acknowledged the crucial need to focus on maintaining integrity and they planned the roll-out according to the implementation literature through consideration of factors such as organisational and staff readiness. They then administered the CPAI *hypothetically* to assess their plans, ensuring that they had accounted for each of the domains prior to roll-out. This methodology could have significant implications in the private sector, where the CPAI could be used by commissioners (albeit in an incomplete way, ie, without access to an up-and-running programme for observation, interview, file-read, etc) to assess the potential of planned services when weighing up bids for contracts.

Current developments in the privatisation of offender services in England and Wales arguably lead to greater possibilities for the CPAI (though the likelihood of these coming to fruition in the context of the evidence-free, profit-driven reforms of Transforming Rehabilitation (TR) appears slim – see, eg, Raynor, 2018a, 2018b, 2018c). Wright (2014) argues for contract incentives that harness the profit-focus of the private sector to the employment of evidence-based practices and the Ministry of Justice's Payments by Results (PbR) strategy ostensibly involves a proportion of a service provider's payment being determined by the reductions in reoffending they achieve (Bardens & Grimwood, 2013). However, whilst reductions in reoffending are the goal of (RNR) rehabilitative services, recidivism is a problematic measure. Recidivism studies are fraught with difficulties such as the complexities of ensuring a matched sample (since internal and external factors can influence recidivism rates), definitional disagreements regarding recidivism (eg, re-arrest or re-conviction) and the lack of standard, appropriate follow-up periods (Latessa, 2012, pp. 66-67; Wright, 2014,

pp. 180-181). In women's treatment, a further complication is the "holistic" approach advised by the gender-responsive literature, which integrates a number of services for both offenders and non-offenders. In their 2011 analysis of the Together Women (TW) intervention programme in England, Jolliffe et al noted that reoffending rates may be an inadequate outcome measure for such an endeavour:

As a result of data limitations this evaluation was limited to assessing the impact of TW through the measurement of proven re-offending. However, TW offered offenders and non offenders a broad range of services, only some of which appeared to have a clear theoretical basis to support their potential to reduce re-offending. Services offered included mentoring, general counselling (i.e. talking through an issue with a key worker), attending an appointment to another agency, self-esteem courses, support in tackling domestic violence, debt advice, and support in finding training and employment. Service users could also receive support from partner agencies to which they were referred by TW. Given the range of services offered it is questionable whether TW can be appropriately evaluated using only official measures of proven re-offending. (Jolliffe et al, 2011, p. 26).

They further identified that a lack of standardised measurement tools and data monitoring requirements prevented meaningful evaluation. Gelsthorpe & Hedderman (2012) argue this narrow focus on recidivism figures will have extensive, detrimental, implications for women's services, possibly ruining established good practice. In any case, Wright's view is that a lack of standardised recidivism evaluations prevents objective assessments being carried out in the timely manner demanded by the private sector and that specific contract incentives should therefore not be dependent on recidivism measures. He instead proposes linking incentives with performance on the CPAI and suggests that bonuses may be paid for achieving higher scores, the key being that incentives are awarded on the basis of objective assessment of practice, instilling both guidance and accountability into the contract (Wright, 2014, pp. 180-181).

From negotiating contracts to identifying the strengths and weaknesses of services, it can be seen that use of the CPAI may provide the blueprint from which policy makers, commissioners and services themselves are able to build a bridge between theory and practice.

A Case for Change?

It may seem contradictory to promote innovation on the one hand and encourage services to meet a set of fixed criteria on the other. To address potential criticism that the CPAI may discourage experimentation, Van Voorhis and Brown point out the history of popular but ineffective panaceas in corrections and caution against rejection of an empirically sound model of intervention – the CPAI interrogates the validity of services' selection of targets and interventions. Furthermore, they identify the CPAI as an “evolving tool” which has already been subject to several revisions and may be adapted further to accommodate advances in criminological theory, though it is important to note that the evidence for RNR, such as crime correlates, has “shown tremendous stability over time, place and population variabilities” (Van Voorhis & Brown, 1996, pp. 20-21).

In spite of evidence that the CPAI is valid across a variety of different types of programmes, correctional settings and offender types (eg, Nesovic, 2003), several alternative tools have been developed from its template according to perceived additional needs. For example, the California Program Assessment Process (CPAP) includes an evaluation of the strength of existing research around specific programmes (Grattet et al, 2006). Additionally, it seems likely that some areas will have developed their own tools for reasons of cost-efficiency. The CPAI is copyrighted to the developers in all its versions and can only be administered by accredited evaluators, to whom the tool is licensed and who incur costs each time the tool is used. It can nonetheless be said to be cost-effective when viewed in the context of the large amounts of money spent on resourcing and conducting outcome assessments of ineffective programmes (Van Voorhis & Brown, 1996).

Another impetus for modification has been research into the comparative validity of the individual domains and items of the CPAI across different programme settings (Holsinger, 1999; Lowenkamp & Latessa, 2003; Lowenkamp, 2004; Lowenkamp & Latessa, 2005a; Lowenkamp & Latessa, 2005b). The University of Cincinnati's Correctional Program Checklist (CPC), includes only the CPAI items evidenced to be positively correlated with programme outcomes and has been developed further to

include different versions specific to context, including the CPC-Drug Court (CPC-DC), CPC-Group Assessment (CPC-GA), and CPC-Community Supervision (CPC-CSA) (University of Cincinnati). In contrast, a single CPAI tool is available which, just as it is applied across different institutions and interventions aimed at reducing reoffending, is purportedly equally applicable to programmes engaging different offender groups, including women. As Chapter One explains at length, the literature suggests that women's RNR interventions should be delivered within a gender-responsive framework, defined by Bloom and Covington below:

Gender-responsive means creating an environment through site election, staff selection, program development, content, and material that reflects an understanding of the realities of women's lives and addresses the issues of the participants. Gender-responsive approaches are multidimensional and are based on theoretical perspective that acknowledges women's pathways into the criminal justice system. These approaches address social (e.g., poverty, race, class, and gender inequality) and cultural factors, as well as therapeutic interventions. These interventions address issues such as abuse, violence, family relationships, substance abuse, and co-occurring disorders. They provide a strength-based approach to treatment and skill building. The emphasis is on self-efficacy. (Bloom & Covington, 2000, p. 11).

Though literature addressing evaluability assessment in this context is scarce, there is some evidence that the CPAI is both appropriate and useful. In their evaluation of the effectiveness of a gender-responsive, cognitive-behavioural programme for female offenders, Duwe and Clark (2015) found that the programme significantly reduced the risk of re-arrest and reconviction when delivered with fidelity, but that when parts of the curriculum were cut, the length of the programme was shortened and class sizes exceeded recommended numbers, it ceased to reduce reoffending, indicating that integrity is crucial in the delivery of gender-responsive RNR services. In her assessment of women's programmes delivered by the Correctional Service of Canada (CSC) Booth (2012) used the CPAI alongside the Gender-Responsive Program Assessment (GRPA; Covington, 2007) to assess the extent to which women's programming integrated both aspects of the RNR model and principles of gender-responsive programming. With the evaluated programmes scoring "Satisfactory" on the CPAI and "Good" on the GRPA, and respondents to staff and women offender participant questionnaires identifying elements of both effective practice (RNR) and gender-responsive programming as present and important, her findings suggest that implementation of an integrated

approach is desirable, achievable and measurable. Booth's conclusion illustrates the way in which assessment of gender-responsivity can complement the findings of an evaluability assessment by enhancing the focus on areas of distinct relevance to women's programmes (in this case cultural considerations) alongside empirically supported elements (such as organisational and implementation issues):

The results of this study suggest that continued focus on the integration of culturally relevant programming elements is warranted. Moreover, in order for facilitators to be equipped to provide correctional programming to the highest possible standard, operational and implementation issues related to consistent staffing, the provision of ongoing supervision and securing of adequate funding appear to be areas requiring further attention. (Booth, 2012, p. 25).

According to Salisbury's 2015 article, a Gender-Responsive Correctional Program Assessment Inventory (GRCPAI) is currently being developed and piloted by Van Voorhis. Salisbury notes that this has the potential to combine the gender-responsive and RNR evaluability literature to give a holistic, evidence-based evaluation of women's services:

...this gender-informed tool adds the assessment of program dimensions that are pertinent to women offenders, including gender-responsive criminogenic needs (Van Voorhis et al., 2010; gender-informed case management (e.g., *Women Offender Case Management Model*; Millson, Robinson, and Van Dieten, 2010); multidimensional substance abuse programming (used to address the confluence of substance abuse, mental health and trauma); and wraparound services targeting education for sustainable careers, poverty, trauma, healthy relationships, safety, and parenting...Additionally, the assessment also targets program qualities that are pertinent to the treatment of both males and females, including leadership, assessment, case management, staff qualifications and training, cognitive-behavioural programming options, and program resources. (Salisbury, 2015, p. 7).

This chapter has so far presented the case that, (a) a gap between theory and practice exists in the British Isles (as elsewhere), (b) integrity is crucial to the effectiveness of RNR services, including interventions for women, (c) evaluability assessment is a useful method to facilitate knowledge transfer, thus bridging the gap, and (d) the CPAI is a valid and flexible evaluability tool in the field (which may even be used alongside or in combination with assessment of gender-responsivity). Even so, it cannot be assumed that the CPAI, a tool developed and validated in Canada and the US, applies equally in the jurisdictional contexts of the British Isles.

Separated by a Common Language

In the field of medicine psychology, for example, the need for the linguistic validation and cultural adaptation of patient-reported outcome (PRO) instruments is widely accepted and applied. In their study of the cross-cultural adaptation of haemophilia PRO questionnaires (Haemo-QoL, Haem-A-QoL and Hemo-Sat), von Mackensen et al (2013) noted the semantic/conceptual, cultural, idiomatic and grammatical complexities of items that necessitated different English-language “translations” (Canada, SA, UK and US) in addition to different versions in other languages. However, in spite of the unequivocally culture-specific nature of social science in general, translations of social science texts all too often fail accurately to translate concepts so as to be properly understood by the target culture. For example, the English translation of Simone de Beauvoir’s influential study *Le Deuxieme Sexe* from the French uses the word “subjective” colloquially to mean “personal”, rather than, as de Beauvoir intended, an existential term meaning “exercising freedom of choice” (Glazer, 2004; Heim & Tymowski, 2006). Wallerstein cautions that:

A social science text utilizes concepts as the central mode of communication. The concepts are more or less clearly defined and applied by the author. On the one hand, they are shared references of meaning, shared summations of data or classifications of reality. Were they not shared with some others, the text would be gibberish. On the other hand, these concepts are not *universally* shared and are quite often the subject of open and violent conflict. (Wallerstein, 1981, p. 88).

Transposing the language of the CPAI-2010 for use in the British Isles involves more, then, than merely swapping American English terms for their British English equivalent. For instance, “Program” requires more than the addition of “-me” to make the Canadian term understood *as intended* in the jurisdictional contexts of the British Isles. The CPAI’s authors intended “program” to be understood “very broadly to include treatment programming, services, and social controls” (Taxman, 2014, p. 32), indicating the instrument’s wide applicability, whereas, in the criminal justice services of the British Isles, the term “programme” is a limiting concept, most often referring to specific behaviour-change interventions (most often cognitive-behavioural groupwork) that are structured, manualised, and often accredited (see, for example, overviews of

programmes provided by Staffordshire and West Midlands Community Rehabilitation Company (CRC) and Jersey Probation and After Care Service; and NOMS, 2010, p. 11).

Not only are there cultural differences between the country of origin and target jurisdiction, but the conceptual language should ideally be uniformly understood by the different participants in evaluation: the practitioners, management and service-users interviewed using the inventory, and the target audience for the report, which may also include policy makers, the general public, third- and private-sector agencies, etc. Polaschek argues that the detached, technical terminology of the RNR model is both off-putting and prone to mistranslation, noting: “Simpler, clearer and more accessible language serves to reduce misunderstandings, and has other important practical implications: readers are likely to like the model more and view it as easier to use” (Polaschek, 2012, p. 9). Of their work with STICS, Bourgon and Bonta (2014) describe changing the formal language of cognitive-behaviouralism for the purposes of enhancing client engagement, learning and application of the concepts to thinking and behaviour. It stands to reason that meaningful evaluation must successfully translate the formal languages of theory to that of policy, practice and experience in the target culture(s). However, translation must not equate to over-simplifying/undermining the specialist concepts integral to a professionalised service since it is possible that they may only be adequately described using technical language. It is important in the translation of a professional tool such as the CPAI-2010 to discern between concepts that are not understood due to complex terminology and those that are not understood due to a knowledge/skills deficit. Clearly, this distinction should inform development of the instrument - the first may warrant the translation of concepts into a more accessible register, the second the inclusion of a glossary of terms, for example.

A further language consideration is ideological. In line with labelling theory¹⁵, desistance advocates argue that the language of interventions is crucial, possessing the

¹⁵ See, for example, Tannenbaum (1938) who theorises that “tagging” an individual with a negative label contributes to their future delinquency and Link et al (1989; 1999) for an introduction to modified labelling theory, which finds that an individual’s self concept is dramatically impacted upon by (even seemingly minor) societal reaction. However, the attention of desistance theorists departs from labelling theory in its focus on the processes of desistance from crime. As Maruna (2001, p. 20) asserts, labelling theory suggests that the stigma of deviant labelling is “almost irreversible” (Erikson, 1962, p. 311). Desistance theory goes beyond

ability to reinforce negative perceptions of offenders (or “*people* who have offended”), which make it harder for them to give up crime (McNeill & Weaver, 2010, p. 5). The language of both policy and practice must be “future-oriented, optimistic and approach goal focused” to create an environment of positive growth and participation (Ward & Maruna, 2007, p. 127). It is therefore of considerable detriment to the perception and practical utility of the RNR model that it appears to have been developed without due regard to the nuances of the language with which it is articulated. As this chapter has set out to demonstrate, desistance and RNR may not be irreconcilable paradigms, but poor translation could yet confound attempts to integrate them. Use of terminology such as “Correctional Program” may imply that services are “providers of correctional treatment (that belongs to the expert)” and not “as supporters of desistance processes (that belong to the desister)” (McNeill, 2006, p. 46). In other words, because of the language used, and in spite of its underpinning intent of “human service interventions *for the offender*” (Polaschek, 2012, p. 10; Andrews & Bonta, 2010), the RNR model may be perceived to be proposing that interventions are done *to* a service-user in the interests of society rather than *with* a service-user in the interests of achieving shared goals for the benefit of both individual and community (Ward & Maruna, 2007). Harris summarises the dangers of these perceptions, based on his work with incarcerated men in Pennsylvania:

Many people who are currently or were formerly in prison ... hold negative attitudes toward the concept of rehabilitation and correctional treatment programs. In general, the distaste for such programs is linked to a sense that these interventions involve things being “done to” or “prescribed for” passive recipients who are characterized as deficient, ineffectual, misguided, untrustworthy, possibly dangerous, and almost certain to get into trouble again. Although people who have been incarcerated often believe that some staff members or other outside parties and some types of programs can be helpful, their effectiveness stems from the potential they offer for empowering participants rather than trying to compel them to change. Most argue, “No one else can rehabilitate you. You rehabilitate yourself.” (Harris, 2005, p. 318).

Scope for a Bridge

Clearly, a range of factors contribute to the hiatus between principle and practice. Rushed timescales and the need for political expediency impede efforts to deliver interventions as designed. Policy-makers, managers and the different agencies involved

this limitation “to specify the social mechanisms which might operate to ‘return’ the stigmatized secondary deviant to a ‘normal’ and acceptable role in the community” (Trice & Roman, 1970, p. 539).

have multiple, and possibly contradictory, goals that may undermine or obstruct effective practice. Managers and staff may simply not have the resources, understanding or skills to properly implement the principles, which may themselves be under-articulated. These implementation issues combine not only to produce ineffective practice but also resistance and conflict in the field. However, rather than deny the unparalleled evidence base for RNR, this chapter has attempted to survey the distance and proportions of the gap between research and real-world practice in order to determine the scope for a bridge. In this way it reaches similar conclusions to Herzog-Evans's 2018 review of the RNR, desistance and Good Lives models, which calls for an integrative approach, necessitating, amongst other considerations, a renewed focus on organisational context and staff practices.

Herzog-Evans's review goes further still, arguing for inclusion of the legal framework in an integrated treatment model. Her call for "good courts" that work in a "therapeutic alliance" with offenders and "good laws" that sustain rehabilitation through, for example, expunging criminal records to remove the stigmatising and exclusionary effects of criminalisation, extends the responsibilities for change beyond the individual and into a social context that may either support or inhibit rehabilitation (Herzog-Evans, 2018, pp. 112-116). Clearly, the role that probation agencies and practitioners can have is limited in this respect – full operationalisation of Herzog-Evans's model requires policy engagement at a governmental level. However, RNR's responsibilisation of the individual has been a key point of departure for proponents of alternative or integrated approaches (Herzog-Evans, 2018; McNeill, 2012; Williams, 2018).

Departing from responsibilising individuals is especially significant in gender-responsive work, since a number of critics cite women's ecologies (personal relationships and social contexts) as defining their pathways into crime (eg, Covington & Bloom, 2003; Hannah-Moffat, 2004; Kendall, 2002, 2004; Pollack, 2005; Shaw & Hannah-Moffat, 2004) and argue that women's offending can only be properly addressed through engagement with (and transformation of) the "deep structures" of society (Malloch et al, 2014, p. 406; see footnote 10, p. 17). As Blanchette and Brown (2006, pp. 127-128) demonstrated, holistic engagement with individuals in their social

context is not only compatible with an RNR approach but specifically required by the model's inclusion of advocacy and brokerage as core correctional practices (see p. 31). Nevertheless, the responsibility and potential for *change* remains predominantly on and with the individual and this is inconsistent with a conceptualisation of desistance as "a journey towards integration" (Williams, 2018, p. 28). As Williams argues, "[r]eintegration is a two way street. Offenders need to learn new ways but society needs to alter in order to accept them" (Williams, 2018, p. 29). Furthermore, Canton sees this as central to the legitimisation of punishment, since this depends on membership of a moral community that "has responsibilities towards as well as rights against offenders" (Canton, 2018, p. 266) (a philosophy reminiscent of the social welfarist principles upon which probation was founded, but which has been overtaken in recent years by neoliberal responsibilisation (Raynor, 2018d)).

Thus, arguments for an integrated approach (Canton, 2018; Henley, 2018; Herzog-Evans, 2018; Williams, 2018) require the focus for probation to be on '*four forms of offender rehabilitation*' (McNeill, 2012). In addition to the focus on individual-level *psychological* or *personal* change, *legal* or *judicial* rehabilitation is essential, alongside *moral* and *political* rehabilitation, which involves making reparation to the community, and *social* rehabilitation which involves "[r]ebuilding a social ecology of family, friends and contacts who can help to sustain a positive life-style" (Williams, 2018, p. 29).

Whilst the RNR approach acknowledges the importance of social context, it does little to respond to the need for legal, moral or social rehabilitation. In addition to perceptions of these areas as being outside of the reach of probation practice (Herzog-Evans, 2018, p. 116) this is likely, at least in part, because the "effectiveness" of reintegration is not easily measurable (of note, advocacy and brokerage were not significantly correlated with reduced recidivism in Dowden and Andrews's 2004 meta-analysis) and so claims to "evidence-based practice" in these areas are difficult to support empirically. This chapter has argued for an integrative approach based on the commonalities between 'what works', desistance and gender-responsive perspectives. Though there is little to suggest that the perspectives themselves are incompatible, the definition of "evidence" as "what works" (quantitatively) to reduce reoffending may need to be reconsidered to

accommodate the “rich body of theory and research on how and why the rehabilitative process works...known as desistance” (Maruna, 2015, p. 313) in order to develop a truly integrated approach. Without this, RNR endeavours, such as use of the CPAI, will fail adequately to capture promising desistance and gender-responsive practices such as work to re-unify mothers with their children, divert women from criminalisation or engage courts, probation and social services in the provision of therapeutic justice.

For this reason, amongst those discussed throughout the chapter, it is yet to be seen whether the CPAI will be well fitted to women’s services in the British Isles and this needs to be explored with agencies, practitioners and offenders. The hope is that such an endeavour will equip them with the capacity to move forward under the aegis of an unparalleled scientific body of evidence. This study sets out to determine how knowledge about effective practice is best packaged and applied in the process of trying to improve practice in agencies. It asks:

- What processes of evaluation are meaningful to services?
- Are the aims of the CPAI-2010 appropriate to services in the British Isles?
- Are the items of the CPAI-2010 appropriate to services in the British Isles?
- Is the language of the CPAI-2010 appropriate to services in the British Isles?
- How should assessment findings be presented?
- What barriers are there to use of the CPAI-2010 in the British Isles?
- How could these barriers be overcome?
- Can the CPAI-2010 be developed upon to facilitate more responsive evaluation of women’s services in the British Isles?

CHAPTER THREE

Methodology

Introduction

This project is, in essence, a pilot study of the implementation of a pre-existing system of service evaluation, the Correctional Program Assessment Inventory-2010 (CPAI-2010), to establish how useful it is in the context of the British Isles. For the purposes of testing the applicability of the CPAI-2010, this study deals with two distinct types of methodological issues and the initial part of this chapter covers the first of these – the methodology required to use the system. The chapter continues with a description of the study sample and a discussion of the second methodology, which is concerned with assessing the system’s usefulness in the context of services in the British Isles, with a particular focus on women’s services, as well as exploring access and ethics considerations and the methodological limitations of the study.

Phase One: The CPAI-2010 Methodology

Use of the CPAI-2010 is subject to numerous legal requirements and, prior to commencement of this study, Swansea Service Evaluation Team (SSET) signed contracts detailing the terms of use. The CPAI-2010 can only be administered by accredited evaluators and the accreditation process undertaken prior to this study involved both theoretical and practical training with two of the tool’s developers, Paul Gendreau and Yvette Thériault, and completion of a full CPAI-2010 evaluation of a sample programme with the developers’ oversight and guidance. The constrictions of contractual obligations in this instance ensured the robustness of the research methodology when applying the tool *as intended* in this new context.

The nine domains of the CPAI-2010 include 143 items, 133 of which are scored ‘Yes’ (1), ‘No’ (0) or ‘Not Applicable’ (void) according to whether different aspects of practice adhere to the principles of effective practice (Gendreau et al, 2002a):

A. Programme demographics	10 items (none scored)
B. Organisational culture	9 items
C. Programme implementation/maintenance	10 items
D. Management/staff characteristics	18 items
E. Client risk/need practices	13 items
F. Programme characteristics	25 items
G. Core correctional practice	45 items
H. Inter-agency communication	5 items
I. Evaluation	8 items

To enhance reliability and reduce subjective scoring, there is a confidence rating for each scored item that requires the scorer to indicate the reliability of their assessment on a five-point scale according to the evidence available to support the score.

An overall score is calculated by collating the scores in each domain and dividing the sum of these scores by the total number of applicable items. The overall score is assigned to its classification of programme integrity, ranging from Very Satisfactory (70% and above) to Unsatisfactory (below 50%). The average confidence scores for each section are used to ascertain the reliability and validity of the evaluation’s conclusions. The results of the CPAI-2010 evaluation are given in the form of an in-depth report, which includes the scoring summary, overall classification, qualitative evidence and recommendations.

In order to undertake this detailed assessment of services, data is collected using a number of strategies. Data relating to staff and participant demographics (eg, number of qualified staff, participant numbers, etc), funding and referral sources is gathered prior to the on-site evaluation using a staff member survey (Appendix 1) and request for information form (Appendix 2). On-site data collection involves three separate methods:

1. A review of a service's documentation includes scrutiny of open and closed case files, programme manuals, meetings minutes, logs, policy and procedure manuals, training materials, assessments, evaluations of the programme and/or staff, etc. Given that the provision and/or completion of certain documents are requirements of effective practice, it follows that a document review will yield credible evaluation information due to the tangible and contemporaneous nature of the evidence. For example, Item B1 of the CPAI-2010 requires that an agency "clearly document[s] its goals/mandate". A document review should produce definitive evidence in this respect – ie, the evaluator should see a formal document articulating the service's goals and mandate. Documented evidence relating to, for example, risk, need and responsivity assessment practices (including the completed assessments themselves) should also be available in paper or electronic form. However, the file read can provide only very limited information relating to organisational environment and processes, management and staff characteristics, treatment practices and service-user experience.
2. Semi-structured interviews are undertaken with the "Programme Director"¹⁶, supervisors, frontline staff, service-users and other stakeholders. King et al (1987) note that self-report measures are a useful method of data collection when assessing programme implementation, since "It makes sense, of course, to turn to the people who have experienced a program for information about it" (p. 44). However, they caution that data collected using this method "may have credibility problems" due to the intentional bias of interviewees (eg, staff with a vested interest in making the programme look good), the secondhand nature of the evidence (the evaluator reports what interviewees say happened) or the unreliability of recollection/self-awareness (p. 44). Bauwens (2010) refers to "Cohen's (1985) distinction between the practice and the *story*" when recounting her own probation research, which found discrepancies between interview responses and observed practice.

¹⁶ The "Programme Director" is the member of staff responsible for the programme services delivered to offenders. Their remit includes hiring, training and supervising staff as well as development and maintenance of the service. In the British Isles various individuals may be responsible for different aspects of this role.

3. CPAI-2010 evaluators also conduct direct observations of interventions, casual contacts with service-users, interactions between staff, communication between management and staff and interactions between service-users. Bauwens (2010) notes that observational studies in probation are rare, due in part to the perceived sanctity of the relationship between officer and service-user. As Howe recollects, “for a long time it was presumed what did take place was essentially a private, professional matter between the worker and her/his client. What was achieved and experienced was simply a product of that worker’s way of doing things” (1991, p. 202). However, such research as Raynor et al’s 2014 study into *The Impact of Skills in Probation Work* is notable in its use of an observational methodology to assess *how* probation officers work with offenders in supervision sessions. Their findings (that practitioners’ increased use of relationship and structuring skills correlates with reduced recidivism) indicate the value of this kind of research strategy, not least in shaping practitioner training and practice.

All three data collection methods are crucial to the CPAI-2010 methodology. For example, interviews will yield evidence relating to which treatment strategies the practitioners seek to employ, the document review will enable the evaluator to assess whether the service has a manual detailing the “types of treatment provided and the treatment activities” (Item F9), and observation of intervention sessions will provide evidence of how service-users are *actually* engaged in practice. The combination of all three methods will yield evidence relating to “treatment”, the data gained using each method supporting or qualifying the findings of the others. Borrowing terminology from the fields of trigonometry and geometry, this approach is referred to as ‘triangulation¹⁷ by method’ or ‘methodological triangulation’ (eg, Denzin, 1989; Miles & Huberman, 1994). King et al (1987) note that an evaluation of programme implementation “will be more valid and credible if data from several sources converge on the same picture” (p. 45), while Flick (2002) advocates the use of triangulation in research to “systematically extend and complete the possibilities of knowledge production...[increasing] scope,

¹⁷ In the geophysical sciences *triangulation* refers to determining the location of a point using observations from two additional points. In social research, the term has been adopted by theorists (originally Campbell and Fiske, 1959) who argue for the combination of different methods, study groups, local and temporal settings and different theoretical perspectives in dealing with a phenomenon. Triangulation is used to enrich, refute, confirm and/or explain data (Denzin, 1989).

depth and consistency in methodological proceedings” (p. 227). Regarding her methodological approach to research into “the transformation of offender rehabilitation” in Belgium, Bauwens summarises:

...using method triangulation (file analyses, interviews and observations) is a valuable approach that does not merely duplicate data, but also offers complementary insights and understandings, and can reveal some important discrepancies that might have remained uncovered through reliance on interviews alone. (Bauwens, 2010, p. 39).

Denzin’s 1989 systematology includes three further types of triangulation: data triangulation, investigator triangulation and theory triangulation. Data triangulation refers to the use of different data *sources*. The CPAI-2010 evaluation process involves the participation of practitioners, managers, partner agencies and service-users at different times (even retrospectively through file reads) and in different settings. This approach enables evaluators to “purposively and systematically involve persons and study groups, local and temporal settings in the study” (Flick, 2002, p. 226). The CPAI-2010 developers also advocate investigator triangulation, an approach in which more than one evaluator is engaged in analysing the data and scoring the tool, thus minimising biases resulting from the evaluator as a person.

Theory triangulation involves the use of multiple perspectives and hypotheses in the interpretation of a phenomenon and here the methodology of the CPAI-2010 departs from qualitative strategies, which seek to interpret data without preconceptions. The CPAI-2010 provides the theoretical framework within which data collection takes place, an evaluation methodology identified by King et al (1987, p. 21) as unequivocally quantitative, regardless of data collection method. Whilst qualitative data is used in the report to describe the service and provide accounts of its strengths and weaknesses, it is also used (alongside quantitative data) to quantify through scoring whether or not the service is employing evidence-based practices. Interviewees’ responses to Item I2, for example: “How does the program ensure that the manual is followed?” may provide useful and important qualitative data regarding the ways in which management, practitioners and service-users contribute to the processes and evaluation of practice, but it is also used to quantify adherence to the principle that “a mechanism is in place that

consists of regular group observations to ensure fidelity to the program manual” by scoring 1 point for “Yes” or 0 for “No” (CPAI-2010, Item I2).

In spite of its clearly positivist epistemology, the CPAI-2010 takes a pragmatic approach to data collection and benefits from both quantitative and qualitative strategies. Grounded Theory (GT), for example, is a qualitative strategy in which Glaser and Strauss prescribe an approach whereby theory is generated by the coding of emergent data into groups and categories, a concept clearly incompatible with the preconceived theoretical framework of the CPAI-2010. Nevertheless, the GT concept of theoretical saturation is useful to the CPAI-2010 evaluator, for whom neither the data collection period nor a requisite number of interviews and observations is specified.

According to Glaser and Strauss:

The criterion for judging when to stop sampling the different groups pertinent to a category is the category’s *theoretical saturation*. *Saturation* means that no additional data are being found whereby the sociologist can develop properties of the category. As he sees similar instances over and over again, the researcher becomes empirically confident that the category is saturated. He goes out of his way to look for groups that stretch diversity of data as far as possible, just to make certain that saturation is based on the widest possible range of data on the category...The criteria for determining saturation, then, are a combination of the empirical limits of the data, the integration and density of the theory, and the analyst’s theoretical sensitivity. (Glaser & Strauss, 1967, pp. 61-62).

When addressing an item such as G8 “Staff are enthusiastic and express optimism”, for example, the evaluator cannot simply observe one session in which the facilitator demonstrates enthusiasm and optimism. S/he must instead seek out opportunities to observe other sessions, other facilitators, other interventions, etc, until instances are repeated to the point at which the “theoretically sensitive” evaluator can confidently report on practice relating to that item across the service. Whilst confidence scores can be used to indicate the level of evidence available to the evaluator for each item, consistently low confidence scores would negate the value of the evaluation.

The Sample

Sampling for this study was driven by SSET's research agenda to develop evaluation approaches and instruments *with* and *for* services and, as such, the sampling process differed for the different agencies involved in the study.

A Women's Project in Wales

Representatives of Integrated Offender Management (IOM) Cymru and the (now dissolved) Wales Probation Trust (WPT) engaged with CPAI-2010 training alongside members of SSET as well as youth justice (YJ) practitioners. In contrast with SSET's research into YJ services in which YJ practitioners went on to conduct evaluations alongside SSET following training (Ugwudike & Morgan, 2015), the representatives of adult services chose not to continue the accreditation process following training. A participative research methodology whereby research is "carried out with and *by* local people rather than on them" (Cornwall & Jewkes, 1995, p. 1667; [My italics]) was therefore not feasible for this study given that copyright restrictions prevent the CPAI-2010 from being used in any capacity by anyone other than accredited evaluators. As Bergold & Thomas (2012) emphasise "unless people are involved in decisions – and, therefore, research partners, or (co-) researchers – it is not participatory research." However, the participative ethos of enabling "local people to seek their own solutions according to their priorities" (Cornwall & Jewkes, 1995, p. 1668) underlies the initial sampling decision. The women's project that is the focus of this study was identified for evaluation by participants in the training who were "local" to the project as a result of concerns pertinent to its management and operational team. Their concerns were: an increased interest in the treatment of women offenders in light of the Offender Rehabilitation Act (ORA) 2014, the identification of women as a key priority group in the Reducing Reoffending Strategy for 2014-2016 and IOM Cymru's agenda for multi-faceted evaluation of pilot projects to feed into the development of a detailed operating model for women in the Criminal Justice System (CJS) in Wales.

The women's project evaluated for this study was being piloted in an urban Welsh probation area to engage all women from first point of contact with the CJS (usually arrest) and to encompass all women engaged in statutory CJS involvement in the community as well as introducing voluntary engagement through diversion to services as an alternative to charge, during bail and pre-sentence and during resettlement from custody. Additional to this, women could self-refer to the project if they had previously been involved with the CJS or identified themselves as being at risk of committing offences. The prevailing presenting problems of service users reported by the project were deficits in thinking and behaviour, lifestyle and associates, substance misuse, accommodation, mental health issues and victimisation issues¹⁸. The project was delivered in partnership with a number of agencies – The National Probation Service (NPS – responsible for managing high-risk statutory cases), Wales Community Rehabilitation Company (CRC – responsible for low/medium-risk statutory cases), and a number of statutory and third sector partners who deliver services to women offenders and women “at risk” of offending.

Jersey Probation and Aftercare Service community supervision service

Jersey Probation and Aftercare Service (JPACS) functions within a very different jurisdictional context from the women's services in Wales detailed above. Autonomous from the legal systems of England and Wales and France, the States of Jersey legislate for the Channel Island and JPACS is under the governance of a Probation Board consisting of five Jurats (lay elected judges). Independent from strategic and political developments in probation in England and Wales, JPACS has developed long-standing relationships with academics, researchers and professionals in the criminological field, which has shaped the character and development of the Service. Probation Officers, or Delegués, have a number of responsibilities across the Courts and States of Jersey, including the management of Probation Orders (a form of provisional release comparable to the former Probation Orders of England and Wales). This includes the delivery of community supervision for all offenders sentenced to Probation Orders, as

¹⁸ The women's project also reported that the majority of their involvement with service users was related to violent and/or acquisitive offending behaviour, a statement that was repeated in the evaluation report (Kerr, 2014 p. 11). The consequence of uncritical inclusion of data in this way is explored on pp. 166-168.

well as those released from custody. It is this aspect of service that was chosen for evaluation as part of this pilot study in an initial meeting held in Jersey in May 2016 to provide an overview of the CPAI-2010 at the Chief Probation Officer's (CPO) request, following the Service's collaboration with Swansea University researchers in the Jersey Supervision Skills Study (see Raynor et al, 2014).

Co-located with the Magistrates Courts in the urban area of the parish of St Helier (the largest town in Jersey), JPACS was providing community supervision services to 390 adults and 26 youths at the time of this evaluation, 91% of whom were male. The principal presenting problems amongst service-users were alcohol and substance misuse, aggression and emotional instability, relationship issues and, to a lesser extent, financial issues, deficits in employment, negative peer groups and health-related issues. Some elements of service are delivered in partnership with community and third-sector agencies through arrangements such as the Building a Safer Society strategy (BASS). For example, The Adapt Domestic Abuse Prevention Training (ADAPT) group programme is delivered by the Jersey Domestic Violence Forum under BASS.

CPAI-2010 evaluation sampling method

For the purposes of CPAI-2010 evaluation, sampling is undertaken with the intention of providing a 'snapshot' of the evaluated service. As King et al note in their guidance on *How to Assess Program Implementation*:

Usually, of course, you will not be able to collect information from everyone who worked in or was affected by the program – it would take too much effort and time. Instead you will want to sample people within each important role group. Since different groups of program participants may have divergent perceptions, you may want to gather... information, probably on a sample basis, from all key groups...and then compare the information provided by different groups to see if you develop a consistent set of pictures about the program." (1987, p. 44).

The CPAI-2010 interview protocol is that "as a general rule, evaluators should conduct as many interviews as possible" (during the evaluation period) and that "it is permissible to interview a sample of respondents who hold a position, but you should ensure that you interview frontline staff members who have a specialized caseload or

unique job responsibilities” (Gendreau et al, 2010). In this respect, the CPAI-2010 demands that the ‘Program Director’ participates in the evaluation process as a result of several items directly relating to their individual role and responsibilities. Whilst purposive sampling is undertaken, in that the CPAI-2010 defines the parameters for participation (ie, the roles that participants fulfill within the project), the evaluation also involves convenience sampling. Beyond concerns regarding the need for sampling across key groups and the participation of staff members with unique roles, sampling is undertaken according to factors contemporaneous with the evaluation period, such as the availability of practitioners, the willingness of service-users to have interventions observed and/or to engage with interviews themselves, the composition of formal meetings, the type of interventions being delivered, demographics of the service’s environment, etc, and there is a limit to the generalisability (to the project as a whole) of results gained through this approach.

An obvious impediment to generalisability is self-selection bias whereby, for example, practitioners who are resistant to employing evidence-based modalities, or to being evaluated, may choose not to be available for interviews or observation. It may be that practitioners who are more confident about their abilities and adherence to the RNR principles, or are more interested in evaluation, are more likely to participate than those who are not. Additionally, since information about the evaluation and the opportunity to participate in interviews and observation is disseminated to service-users through practitioners, those service-users with more positive views of the project may be encouraged to participate over those with negative views. Equally, service-users and/or practitioners motivated by a desire to air grievances may be more likely to engage with evaluation than those less negatively affected by either the project or their experiences of evaluation. The result of such bias is that the sample may not be representative of the population being studied, or that some finding, or findings, may be exaggerated. For the purposes of CPAI-2010 evaluation, the triangulation of data including data accessed using random (probability) sampling is used to ameliorate these issues. A random sample of active and inactive case files is selected (from a database/dataset of all project service-users, past and present) for review using information system numbers for identification. Data obtained through the file-read may qualify or support data collected

using other methods and the use of random sampling helps to reduce the impact of self-selection bias and increase the validity of the evaluation's findings.

The sampling process for the CPAI-2010 evaluations of the women's project in Wales and JPACS community supervision services involved the identification of key groups (ie, the agencies and individuals involved in the services and their respective roles in the delivery of interventions) through discussion between the SSET evaluator and management teams (ie, the Project Manager and Project Research Officer of the women's project and CPO, Assistant Chief Probation Officer (ACPO) and Probation Team Leader of JPACS), and the distribution of the CPAI-2010 staff member survey (Appendix 1) to all practitioners responsible for delivering services as part of the project (including supervisors/management staff). Interviews with senior staff and the review of documents including case files were arranged individually with the management teams. On-site access for informal observation of services was negotiated across all project premises through the management teams and relevant partner agency managers. All service practitioners were contacted by email to give them the opportunity to participate in interview and observation during the evaluation periods agreed between SSET and the management teams (Appendix 3) and information was also given to front-line practitioners to disseminate to service-users, inviting them to participate (Appendix 4).

CPAI-2010 Evaluation Data Collection

A women's project in Wales

In the women's project, initial orientation meetings regarding the CPAI-2010 were held on 3rd and 28th April 2014. These were used to provide an overview of the CPAI-2010 to the Project Manager and Research Officer. A timescale for the evaluation was agreed at this time, though it was subject to a number of changes and data collection for the first evaluation subsequently spanned the period 23rd June – 12th December 2014. Data collection involved:

- Interviews with the Project Manager, Project Research Officer, Project Team Manager, all full-time front-line NPS and CRC project practitioners (two NPS, three CRC), management staff from two core partner agencies, three service-users and one practitioner from a supporting partner agency;
- Multi-agency practices involving other statutory and third-sector agencies (such as case management meetings and everyday project practice) were reviewed and observed;
- Formal observations were undertaken of two group programme sessions and four offender supervision sessions;
- The document review included project administration documents, project working documents, resource documents for service-users, evaluation documents and 20 active and inactive case files for women offenders managed by WPT, Wales CRC and NPS in the evaluated probation area.

JPACS

An initial orientation meeting was held in Jersey on 10th May 2016 to provide an overview of the CPAI-2010 to JPACS's CPO, ACPO and Probation Team Leader. A timescale for the evaluation was agreed via email following the meeting and on-site data collection took place during the week commencing 12th September 2016, though key interviews with the CPO and Probation Team Leader were undertaken over the telephone on 2nd August and 23rd August 2016 to accommodate conflicting schedules.

On-site data collection involved:

- Interviews with the ACPO, five probation officers, Research and Information Officer, practitioners in substance misuse and restorative justice, six service-users and one magistrate;
- Case management meetings and everyday practice being reviewed and observed;
- Formal observations being carried out of two group programme sessions and five offender supervision sessions;
- The document review of service administration documents, working documents, resource documents for service-users, evaluation and quality assurance documents, training materials, reports and 20 active and inactive case files for offenders (17 male, 3 female) supervised in the community by JPACS.

Phase Two: Feedback - Evaluating Service's Responses to the CPAI

Fieldwork for this study involved two phases: Phase One was project evaluation using the CPAI-2010 (see above); and Phase Two was the collection of feedback regarding participants' experiences of evaluation.

Patton, amongst others, advocates a pragmatic approach to research inquiry which, like this study, aims to “contribute to *useful* evaluation, *practical* problem solving, real-world decision making, action research, policy analysis, and organizational or community development” (2002, p. 145). The social sciences are well known for ontological and epistemological disputes between proponents of different research methods (see, eg, Bryman, 2012), but pragmatism is not committed to one philosophical paradigm. For the utility-oriented pragmatist, reality may exist independently of the mind and meaning, or it may be socially constructed, or it may, most probably, be both. Most important is that “truth is what works at the time” (Creswell, 2014, p. 11). The pragmatic researcher believes that the Platonic tradition of asking questions about reality and the laws of nature has outlived its usefulness (Cherryholmes, 1992). Rather, “they would simply like to change the subject” (Rorty, 1983, p. xiv). Noaks and Wincup succinctly advise criminological researchers to “adopt a pragmatic and theoretically coherent approach to data collection”, in short, “using appropriate methods to answer their research questions” (2004, p. 10).

Since the usefulness of this research to the agencies delivering offender services is paramount to SSET, this study adopted a pragmatic strategy of “matching concrete methods to specific questions...tactically mixing methods as needed and appropriate” (Patton, 2002, p. 69). Data collection for Phase Two, then, was guided by the research questions posed in Chapter Two (p. 69). To ensure the gathering of rich, experiential data, qualitative semi-structured interviews (see Appendix 5 for the interview schedule) were undertaken with relevant individuals who volunteered to offer feedback regarding their experience of the evaluation process and the usefulness of the report. As a result of several issues, including organisational restructuring and staff re-deployment in the women's project in Wales and the self-selecting nature of the sample in both areas, the

sample size for Phase Two was considerably smaller than that of Phase One. In the women's project, second-phase interviews were undertaken with the Project Manager, two NPS and two CRC project practitioners and the manager of one core partner agency. In Jersey, the CPO, ACPO, Probation Team Leader and four probation officers participated in second-phase interviews.

Open-ended questions were used to elicit responses enabling an inductive approach to be taken in analysing the data. Thomas (2003) notes that the purpose of such an approach is to "allow research findings to emerge from the frequent, dominant or significant themes inherent in raw data, without the restraints imposed by structured methodologies", in contrast to deductive approaches, which set out to "test" theory and in which "key themes are often obscured, reframed or left invisible because of the preconceptions in the data collection and data analysis procedures" (p. 2).

A general inductive approach for qualitative data analysis (Bryman & Burgess, 1994; Dey, 1993; Thomas, 2003) provided a straightforward set of procedures with which to analyse systematically the interview transcripts whilst being guided by the specific objectives of the study. This enabled a pragmatic approach to be taken to data analysis, rather than the iterative inductive approach required of, for example, GT methodologies. The interview transcripts were read and re-read to identify and categorise specific segments of information corresponding to particular themes. The upper-level, more general categories were derived from the research questions and the lower-level, more specific categories emerged from the raw data through *in vivo* coding. The category system was revised and refined over several further readings until (i) no new themes emerged (ie, saturation was achieved) and (ii) categories with similar meanings were combined under a superordinate category. In total, seven categories developed that were assessed to be most important given the research objectives, in line with Creswell's model of the coding process in inductive analysis, which considers inductive coding of more than about eight major themes as incomplete:

Table 3.1 The Coding Process in Inductive Analysis (adapted from Creswell, 2002, Figure 9.4, p. 266)

Initial read through text data	Identify specific segments of information	Label the segments of information to create categories	Reduce overlap and redundancy among the categories	Create a model incorporating most important categories
Many pages of text	Many segments of text	30-40 categories	15-20 categories	3-8 categories

Similarities and differences across sub-groups (eg, practitioners vs management, NPS vs CRC, gender-responsive service vs gender-neutral service, etc) and relationships between categories were also explored (see Chapter Five).

Access

Of her research in probation, Rex observed that “gaining access to research participants requires negotiations at every level” (1999, p. 368), a process initiated through the identification of gatekeepers, described by Burgess as “those individuals...that have the power to grant or withhold access to people or situations for the purpose of research” (1984, p. 48). The levels of negotiated access for this study involved engaging with gatekeepers at the top of the organisation, meeting procedural requirements in order to be granted access to services and gaining the cooperation of gatekeepers further down the organisational structure of the services. In terms of access to participant input, there were five levels of negotiations at different research stages:

1. Initial negotiations differed with each service. The participation of the women’s project in Wales was facilitated through the invitation for senior leads from WPT and IOM Cymru to engage with CPAI-2010 training alongside SSET. This enabled a participatory approach in identifying an appropriate project for evaluation and initiated negotiations for access through completion of WPT Practice and Performance Committee consent to research procedures. Noaks and Wincup note that during the ‘getting in’ process, gatekeepers must have a clear and realistic understanding of “the implications of the research for the setting and those who work within it” (2004, p. 57). The participatory approach at this

stage equipped the senior gatekeepers with a detailed ‘inside’ knowledge of evaluation procedures and the purposes of the research project. They then opened up the next level of negotiations through establishing contact between the SSET researcher and Project Manager.

In Jersey, JPACS’s pre-existing collaborative research relationship with SSET engaged senior staff in discussions and correspondence relating to the CPAI-2010 and led to the CPO commissioning SSET to undertake an evaluation of its community supervision services.

2. Initial negotiation with the Women’s Project Manager and JPACS’s senior staff team involved correspondence and meetings to share information regarding the project and the research study and to initiate both the CPAI-2010 and informed consent processes. At this stage, evidence of procedural compliance (ie, provision of a valid DBS certificate for the researcher, the study’s ethical approval and accompanying documents and written consent from the WPT Practice and Performance Committee) was vital to the management teams. Documents relating to the CPAI-2010 evaluation were provided, including a pre-visit letter explaining the evaluation process (Appendix 6), a checklist of materials needed for evaluation (Appendix 7), a request for information form (Appendix 2), and the staff member survey (Appendix 1). This was the level at which broad arrangements such as time-scales and physical settings for the evaluation were agreed.
3. The management teams negotiated access to multi-agency practitioners within the services by facilitating access to stakeholder meetings (through invitation), emailing information relating to the CPAI-2010 and the staff-member survey to all relevant personnel (copying the researcher in and providing researcher contact details for further correspondence – Appendices 1 & 3), and by seeking researcher access to case files through CRC, NPS and JPACS data-management and protection procedures. Responsibility for day-to-day oversight of the evaluation was delegated to administrative members of staff and team leaders

who became contacts for negotiations regarding, for example, room booking and physical access to project premises, practical arrangements for the file read, etc.

4. In the women's project key management, supervisor and practitioner staff members across agencies were identified and engaged in negotiations through their completion of the staff-member surveys, introduction to the research in stakeholder meetings and/or correspondence setting up opportunities to meet and discuss the research project. These initial meetings were used to facilitate informed consent and involved the provision of written information detailing the research study (Appendix 8) and the completion of participant consent forms (Appendix 9). These meetings also provided the opportunity to arrange observations and interviews for the CPAI-2010 evaluation and facilitated the next level of access through the provision of information to be passed on to service-users (Appendix 4).

Key management, supervisor and practitioner staff members involved in the delivery of JPACS's community supervision services were similarly identified through their completion of the staff-member surveys. A timetable was set up for evaluation, with a timescale of one week due to the restrictions of geographical location (ie, the researcher was not able easily to travel to and from Jersey for in-person meetings). Therefore, access was mainly negotiated by email at this stage. A timetable for evaluation was negotiated, with relevant individuals invited to sign up for observation and interview slots. Additional time was allocated prior to each slot to facilitate informed consent processes during the on-site data collection period.

5. Practitioners negotiated researcher access to service-users in the first place by reading information relating to the evaluation (Appendix 4) and inviting them to participate. Negotiations for access were also strengthened through the building of rapport between researcher, practitioners and service-users during informal observation of day-to-day practice. In this way the researcher transitioned from being an 'outsider' to, at least, a known face. Noaks and Wincup note the

importance of gaining “social” as well as “physical” access through “building up rapport with participants and securing their trust” (2004, p. 63).

Negotiations between researcher, management and practitioners were ongoing throughout evaluation and feedback phases of the research. Email correspondence remained active during and between on-site fieldwork to keep practitioners and/or management apprised of the research progress, to update research records regarding organisational developments and to negotiate Phase Two interviews to disseminate research findings/gain feedback. In all, negotiations were facilitated through, (a) researcher understanding of and compliance with the services’ procedural requirements, (b) the building of positive relationships/rapport and the valuing of participant perspectives, (c) the research focus being on developing evaluation processes *for* services, ie, research *in the interests of* its participants (which required the researcher to adopt a reflexive and adaptive approach) and, (d) reciprocal exchange of information in feedback interviews.

Distribution of the CPAI-2010 report is entirely at the discretion of its commissioner(s). JPACS’s CPO disseminated the report (Appendix 10) widely both internally and externally and, as with all its evaluation and quality assurance documents and reports, published it on a public access website, probation.je. In the women’s project, however, the Project Manager disseminated the full report (Appendix 11) to senior leads within IOM Cymru and the National Offender Management Service (NOMS) and the decision was made at that level that the report was not suitable for further distribution, chiefly due to concerns that the project’s Unsatisfactory categorisation might deter potential partner agencies from involvement. Davies and Peters note the importance of “output gatekeepers” who “can suppress findings from being produced and disseminated, or politically motivated pressure can be applied regarding the interpretation of findings” (2014, p. 37). As a result, a sixth level of negotiation in the women’s project concerned the dissemination of findings. It was agreed that a summary report should be prepared in which the scoring and categorisation were omitted, and in which the focus was, instead, on the strengths of the project and recommendations. In this way, the agency received a document which they felt was a positive representation of the project, which they distributed widely, and from which the majority of the evaluation’s findings were

disseminated. Though the findings were, to some extent, censored, research integrity was not compromised. It is SSET's firm belief that evaluation approaches and instruments must be useful to and *used by* services – ascertaining which elements of findings are useful and/or used should shape the development of instruments for future use.

Ethics Considerations

The British Society of Criminology (BSC) advises that “researchers should ensure that research is undertaken to the highest possible methodological standard and the highest quality in order that maximum possible knowledge and benefits accrue to society” (2015, p. 2). In line with the BSC's general principle, this study aims to advance knowledge about criminological issues and to this end the researcher has sought relevant training and opportunities to develop knowledge, skills, attributes and professional integrity. According to the Economic and Social Research Council (ESRC), “the principal ethics consideration should be to ensure the maximum benefit of the research whilst minimizing the risk of actual or potential harm” (2015, p. 2). The study was subject to Swansea University Standard Ethical Approval and WPT Practice and Performance Committee consent to research procedures prior to commencement. Guided by the *BSC Statement of Ethics* (2015) and the *ESRC Framework for Research Ethics (Updated January 2015)*, ethical considerations were paramount to each stage of this study – during planning, fieldwork and the dissemination process – and this will continue to be the case during the archiving, future use, sharing and linking of data.

Pivotal to both ESRC and BSC principles for ethical research is the expectation that “research participants should take part voluntarily, free from any coercion or undue influence” (ESRC, 2015, p. 4). There were two main concerns for this study during the process of planning and implementing this principle: firstly, there is an organisational and professional expectation that staff involved in the delivery of intervention services should participate in the inspection and audit of services and the CPAI-2010 therefore presumes, and operates within, this cultural context. The CPAI-2010 has been developed to be commissioned at a management level and practitioners are generally

instructed to comply with the evaluation. CPAI-2010 guidance is that management should ask staff to “extend [their] cooperation and assistance to the team as they carry out the evaluation” (Gendreau et al, 2010). It was important to recognise this cultural expectation, to differentiate between inspection and evaluation, and, moreover, between evaluation and evaluation *research*, and to address these issues with management and practitioners using the informed-consent processes and ongoing negotiations. Secondly, the study engaged service-users as participants, and informed consent is complicated in the coercive environment of statutory offender services where compliance is rewarded and disengagement punished. In the research context, service-users may feel that they would be advantaged through their participation or disadvantaged if they chose not to participate. The process of gaining informed consent, therefore, needed to be considered with diligence and awareness of these issues.

In practical terms, the BSC advises that the expectation participants must be able to give freely informed consent “implies a responsibility on the part of the researchers to explain as fully as possible, and in terms meaningful to the participants, what the research is about, who is undertaking and financing it, why it is being undertaken, and how any research findings are to be disseminated” (2015, p. 6). The recruitment process for the study involved giving verbal and written information during discussion with managerial staff in stakeholder meetings, dissemination of letters explaining the study to all potential staff participants identified by management (Appendix 3), the provision of pre-evaluation information for service-users (Appendix 4), and a full written explanation of the study given to and discussed with participants prior to obtaining written consent (Appendix 8). Participants were informed at all stages that their participation was voluntary, and the consent form explicitly states: “I understand that my participation is entirely voluntary, I have the right to say no and I can choose not to answer specific questions or to withdraw from the study or stop participating at any time. This will not impact on my work with [the service]” (Appendix 9). In line with BSC guidance that “participants should be able to reject the use of data-gathering devices such as digital recorders”, separate informed consent was sought from all participants for audio recording of both intervention sessions for evaluation and feedback interviews (Appendix 12).

As well as written processes for obtaining and maintaining informed consent, ongoing negotiations with gatekeepers had to be sensitive to the organisational demands of evaluation whilst respecting the constraints of voluntary participation in research. An extract from the fieldwork diary for the CPAI-2010 evaluation of the women's project relates:

16th July 2014... [The Project Manager] asked for the names of practitioners who were not yet engaged, explaining that they were expected to participate in evaluation as part of their professional role. I explained that, as evaluation *research*, their participation must be entirely voluntary and not coerced. We agreed that I would continue to seek the engagement of practitioners through email and encounter at the [service premises] and I didn't provide names, though it may be that the practitioners are under implicit pressure from management to participate and I must continue to emphasise the voluntary nature of participation during recruitment and throughout the study. (Fieldwork diary for the women's project evaluation).

The BSC and ESRC are unequivocal when it comes to the researcher's responsibility and the mitigation of all potential risk and harm (physical, psychological and/or emotional) to participants and researchers through "robust precautions" (ESRC, 2015, p. 4; BSC, 2015). On-site fieldwork took place on the evaluated services' premises and followed the services' risk management procedures (eg, the use of interview rooms monitored by CCTV and/or with panic alarms fitted where there was the possibility of risk to researcher during interview). Practitioner participants were responsible for the management of risk during formal observation of the interventions they were delivering, and the researcher was made aware of agency procedures in this respect. The researcher was subject to Enhanced Disclosure and Barring Service procedures, and the certificate was provided to the agencies prior to commencement of fieldwork.

It is generally accepted that researchers must guarantee some degree of confidentiality for the purposes of developing an environment and relationship where participants can be relaxed, open and honest. Cowburn highlights that, traditionally, the confidential nature of this relationship has been held sacrosanct but questions the privileged role that confidentiality is given in the context of research into offenders. He argues that, where they neglect to act on information indicative of a risk of harm to others, researchers are knowingly colluding and consequently failing to protect the public (Cowburn, 2010, p.

10). This frames the researcher's non-disclosure of certain behaviours, rather than breach of confidentiality, as ethically unacceptable. Disclosure of self-harm is another issue that brings into question the confidential nature of the relationship between participant and researcher. Of his work interviewing prisoners, King writes:

Protecting the vulnerable research participant is one thing but I have never taken the view that confidentiality can be absolute. I *always* tell staff and prisoners that I would not regard as confidential information given to me about planned self-harm or harm to others, for example, or a planned escape, because I always make it clear that I am a citizen and would have my own problems about living with that information. (King, 2000, p. 307).

Whilst the focus of this study was on the service rather than its users or their offending behaviour, the researcher acknowledged that information indicative of risk of harm or vulnerability may be observed or disclosed in interview, and precautions were put in place to address this. In line with BSC requirements that "research participants should be informed about the limits to confidentiality" (2015, p. 6) and, following consultation with service management, all participants were informed that "whilst every effort will be made to respect the confidentiality of information supplied by research participants, action will be taken to manage risk through [the service's] procedures where researchers are made party to information indicating a risk to others or to the participant themselves" (Appendix 8). In practice this would involve the researcher reporting any relevant information at an appropriate level of the service, which would then follow its own risk and vulnerability management procedures, including further assessment, signposting and/or referral to relevant agencies (eg, police, Community Mental Health Team, etc).

These measures were entirely precautionary and did not need to be implemented during the study. To pre-empt inappropriate disclosure and increased risk/vulnerability in the first instance, participants were made aware that interviews are not therapeutic in nature and that the focus of the study is on the interventions programme, not on the participants themselves. However, it was recognised that emotional support may be needed even where official intervention may not be necessary or appropriate. For example, the CPAI-2010 evaluation of the women's project took place at a time of significant organisational restructuring and, faced with the challenges of setting up a pilot in this context, this was a period of extremely low morale amongst some

practitioners. A note from the research diary for this time observes how, at one point, the researcher was in the staff office timetabling observations and the team of practitioners took the opportunity to talk about their feelings of being under-resourced, under-supported and under-valued:

30th November 2014... I saw this as useful data for the evaluation and made some notes while they spoke but was concerned when one practitioner told me “it’s good this, isn’t it? – you’re like a counsellor!” I asked whether there was a need for counselling, but they replied that they needed practical action such as “someone from [the central management team] to come and see what it’s like here. We need computers that work and for someone to see how much work we’re doing and say...you know, you should have less people on your caseload...” They expressed their hopes and expectations that the evaluation report will highlight these issues and make these recommendations and so I spent time talking in detail about the limitations of the evaluation and how, while some of the issues may be covered, it is entirely guided by the evidence relating to effective practice and may not reflect their individual concerns. I feel there is a need, though, to have links to external counselling available for practitioner participants. Stress is repeatedly coming up in interviews and conversations. (Fieldwork diary for the women’s project evaluation).

Whilst the stresses that practitioners expressed were to do with the pressures of their work, rather than the pressures of the evaluation or research, the CPAI-2010 interviews asked them to consider issues such as “Do staff feel they have the tools and knowledge to be able to run the programme effectively?” (Item D18, CPAI-2010) and “How does this organisation respond when problems occur?” (Item B4, CPAI-2010), and this exposed negative feelings which the practitioners felt able to express in the research environment. For this reason, it seemed important to make a clear distinction between the roles of the assiduous, evaluation-focused researcher and that of a benign, client-centred counsellor. Leaflets providing links to local counselling services were available on all service premises.

A further concern was the anonymisation of participants. Identification of individuals can have negative impacts for numerous reasons and in a variety of ways. Should the individual perspectives of practitioner participants in the commissioning or partner agencies be identifiable and attributable, it could impact on their working relationships, the trust afforded to them and other aspects of their professional role. Similarly, an individual service-user’s perspective, if exposed, might affect their relationship with practitioners and/or their relationships with other service-users. In these cases the individual concerned may feel embarrassed, guilty or angry over their personal views or

circumstances being identifiable in reports, articles, etc, over which they have such limited ownership or control. There is an additional concern *vis à vis* service-users, who are also at risk of being identified as an “offender” by individuals outside of the service, which would have the potential of making them extremely vulnerable. As a result of their identity being revealed, however inadvertently, it is conceivable that they could lose employment, relationships, housing, social status or even be made subject to vigilante attacks. Issues of data protection were therefore of the utmost concern in this study:

- With regard to the CPAI-2010 evaluation it was made clear that key practitioners (eg, the Project Manager, CPO, etc) would be easily identifiable in the report from their role descriptions and so were subject to limited anonymity at this stage. However, no individuals were named in the report and, for the majority of the practitioners and for all service-users, characteristics and opinions were generalised and not attributed to any one individual.
- For the women’s project, it was agreed that (with the exception of the CPAI-2010 report, which the commissioning agency owns and disseminates) no intervention programme will be identified by name or specific location in any papers or presentations relating to the study without the permission of the service in question and further consideration given to ethics concerns in that context. JPACS, however, explicitly requested to be identified in all papers and presentations relating to the study in order to benefit from active participation in national and international criminological discourse and research opportunities arising from the study. This reflects the differing concerns of the managerial teams involved in this study. The women’s project was being piloted at a time of intense scrutiny and pressure under the target-driven Transforming Rehabilitation (TR) agenda – a carefully managed profile was considered essential by them to maintain momentum, attract stakeholders and protect services. Conversely, JPACS’s community supervision services were well established and already the focus of several ongoing high-profile research studies into community sentences and their outcomes (Miles et al, 2015), and into skills and strategies in probation supervision (Raynor et al, 2014). For

JPACS, this study would add to its reputation for engagement with academic research, which has included a contribution to the launching of an international research network studying probation practice (CREDOS, the Collaboration of Researchers for the Effective Development of Offender Supervision).

- Where participants have been directly quoted they have given their express permission for this, additional to their consent to participate in the study (Appendix 9), the quotes have been used for illustrative purposes and contain limited identifying features (idiosyncratic turns of phrase, description of circumstances, etc) to minimise the chances of the participant being identified.
- Data relating to the study – completed staff surveys, audio recordings of intervention sessions and interviews, anonymised transcripts of interviews, etc – have been encrypted in a data vault on a password protected laptop. Manual files (eg, hand written notes) were carried by the researcher at all times whilst on agency premises or in transit and kept in a locked filing cabinet at others.
- No personal details (names, dates of birth, area of residence, etc) are, and will not be, included in the CPAI-2010 report, PhD thesis, or any report, presentation or other aspect of the study at any time.
- All case records and documents containing personal details belonging to the evaluated agencies remained on the services' premises, were never in the researcher's sole possession and were subject to the services' data protection procedures, informed by the Data Protection Act 1998 and overseen by service staff, without exception.

Whilst adherence to fixed principles such as those advanced by the BSC and ESRC may be considered an ethically rigorous approach, Armstrong et al (2014) argue for “a more social scientific understanding of rigour where research is not judged by the absence of ethical ambiguities, but by evidence of ethical sensibilities that return us to the heart of the matter – respect for autonomy, beneficence and justice” (p. 217). Contrary to

attempts to put robust pre-emptive measures in place for the purposes of addressing ethical concerns, more ambiguous issues arose as fieldwork progressed. The managerial decision not to disseminate the full CPAI-2010 report of the women's project was, from the research perspective, problematic and raised important ethical concerns. It was essential that evaluation processes under the agency's domain proceeded naturally, without researcher interference. The CPAI-2010 report belongs to the commissioning agency and it would not only be contractually inappropriate for the researcher to disseminate its findings to participants, it would also pre-empt the collection of valuable data regarding lived experiences of the evaluation processes which it is hoped might shape the development of future approaches and instruments. However, this amounted to withholding research findings from participants, which could negatively impact on them and damage trust relationships between participants and researchers. Certainly, there was evidence from interviews undertaken at the time that practitioners were frustrated by the failure to distribute the report. This highlights the importance of the sixth level of negotiations, following which the summary report was prepared and disseminated. Moreover, it was crucial to maintain contact with participants during the period of non-dissemination and at no point to deceive them. A note from the research diary recounts the ways in which the researcher aimed to "respect autonomy, beneficence and justice":

24th March 2015... [The participants] have all expressed anger, frustration or dejection that they haven't seen the report. I'm dealing with this in three ways –

- Data collection: I've explained that it is important for the study to take into account how they're feeling about the whole process, including not having the report, and we've talked about this in interview.
- Debriefing: I've been open and honest about my lack of control over the report and how this fits into the research intentions to develop evaluation tools and processes alongside services and their users – in other words, their experiences of the CPAI-2010 being applied as intended may be negative but we intend for their input to shape better approaches.
- Participatory: I've agreed to consult with the Project Manager to negotiate dissemination of the evaluation findings to participants. (Fieldwork diary for the women's project evaluation).

An overarching ethical concern was the mitigation of intrusion by the researcher, and consideration of whether the introduction of an 'outsider' researcher might compromise the 'safe space' of the women's project was paramount. Due to service-users'

experiences of volatile, coercive and abusive relationships with men, this is women only, and so it was crucial in the first place that the researcher was female. In order to minimise intrusion, it was important to capitalise on the informal observation of day-to-day activity to build respectful, non-judgemental relationships with service-users through warm, friendly but ultimately non-invasive, non-threatening and non-demanding interaction (eg, without the expectation of participation, without note-taking, etc):

3rd December 2014... [A practitioner] didn't turn up for interview today so I spent some time in the waiting room just chatting with service-users, talking about Christmas presents for their children, holidays they'd been on, my job, food they'd liked (but thought they wouldn't), etc. It didn't feel appropriate to take notes for observation but I nevertheless gained useful data and insight regarding their preference for [the project] premises to [the probation office] and the fact that they liked their probation officers here and had felt rushed and ignored or dismissed by their previous officers at [the probation office]. One woman confided about her violent and abusive male relative who could be heard complaining from the street that it was raining and that he was having to wait outside. It struck me how small a safe space [the premises] provides, how vulnerable these individuals are on a day-to-day basis and how privileged I am to have access. My presence and the possibility of compromising what sense of safety they have is an ever-present consideration. (Fieldwork diary for the women's project evaluation).

In Jersey, there was the further consideration of outsider research in a self-governing, culturally-distinct island community. Of Miles's work as an indigenous researcher in Jersey, Miles and Raynor (2014) observe that "'non-local' researchers on Jersey have sometimes struggled to access, understand and interpret local information and elicit useful responses from research participants" (p. 42). They identify the nature of Jersey society as "personality based" (p. 41) and describe a process of gaining trust and reducing the apprehension of research participants by meeting their "vetting" criteria of having insider knowledge and family connections (pp. 41-42). As an 'outsider' in Jersey, the CPAI-2010 researcher was unable to meet these criteria and so took additional time to engage in informal conversation with participants at every opportunity, to show interest in their backgrounds and experiences and to follow up their recommendations for places to visit and things to see and do, deferring to their expertise and showing both willingness and enthusiasm to familiarise herself with Jersey:

15th September 2016... I don't think that it's an exaggeration to say that at least a quarter of each day involves talking to practitioners, staff and clients about Jersey – where best to take

the dogs around the parish of St Martin [where the researcher stayed during the evaluation with her husband and dogs], spots to visit along the coastline, restaurants to try (expensive and cheap) and places of historical interest – our evenings are spent following these recommendations up in order for me to discuss them at JPACS the next day...fitting in is time-consuming but really enjoyable. (Fieldwork diary for the JPACS evaluation).

Methodological Limitations

An obvious methodological limitation of this study is that the sample was relatively small, which raises questions about the generalisability of the findings to, for example, probation areas in England. It is certainly the case that different areas of the British Isles offer different intervention services to different service-user groups and have different ways of working. However, the CPAI-2010 has been developed to evaluate services in any community or residential setting and this study only *tests* this quality (rather than innovates its use) through its implementation in the sample services. Further investigation needs to be undertaken to explore the usefulness of evaluation in different settings – another SSET study is, for example, currently piloting the CPAI-2010 in YJ settings in Wales (Ugwudike & Morgan, 2018) – and, whilst gendered issues are addressed in this study, that is not its predominant concern or innovation, which is to pilot the CPAI-2010 *in the British Isles*. In this respect, what is important is the value of the system in the broader context within which the sample projects operate.

As a single jurisdiction, the way in which probation services operate across England and Wales is bound by the same legal requirements in respect of the way in which offenders are dealt with by the CJS and is subject to the same current policy context of market-building through the introduction of CRCs and Payment by Results (PbR). In this way, the women's project in Wales is representative of the broad statutory context of England and Wales, whilst JPACS provides a useful comparison sample to draw out issues specific to jurisdictional context. However, this study recognises that occupational cultures, defined by Mawby and Worrall as “the values shared by individuals that manifest themselves in the practices of members of that occupation or organization” (2013, p. 5), are at least as important as statutory context when determining the value and appropriateness of evaluation approaches and instruments to services, because:

- They indicate ‘what really matters’ and ‘how things are done around here’
- They provide insight into how practitioners perceive their occupation
- They influence how work is done and how effective it will be
- They influence how new members are introduced into ways of working
- They are a resource for adapting to change and the external context
- They can be a stabilizing force for good
- They can be an obstacle to reform, change and progress
- They come to the fore during turbulent times

(Mawby & Worrall, 2013, p. 7).

In their study of 60 current and former probation workers across urban and rural England, Mawby and Worrall (2013) found that, whilst there were some cultural differences according to location, time, setting, etc, there were significant “family resemblances and common threads that run through and act as cultural locators or indicators against which probation cultures can be identified” (p. 142). In this respect, the sample for the current study, both in Wales and in Jersey, can be seen to be representative of the wider probation population:

1. Mawby and Worrall note that “probation workers are drawn to the job through common values which include a belief in the possibility of change and their own ability to effect it; a faith in both offenders and colleagues...; and an ethos of service or vocationalism” (p. 142). Whilst worker motivation was not a focus of the current study, all practitioners were unequivocally positive in their responses to the CPAI-2010 items D17 “Staff are committed to human service and believe that offenders can change...” and D18 “Staff believe they have the skills to run the program effectively and that the program will be effective...”. As one practitioner in the women’s project put it, “the majority of people here are absolutely passionate about their job...we all want to make a difference” (WPO3).
2. A number of “artefacts” identified by Mawby and Worrall (2013, p. 142) as common to probation cultures were consistently reported by practitioner participants in the current study and observed during evaluation, including: long hours of computer-dominated assessment, reports and records, multi-agency meetings, a female-dominated environment, mutual support amongst colleagues, etc.

3. Mawby and Worrall found that probation workers seek and achieve job satisfaction in a variety of ways, such as “introducing into their work a creativity (or departure from the script) that they believe the organization prohibits (but which the organization is implicitly dependent on)” and that one “core belief” is that of working and building relationships with individual service-users (2013, p. 142). These elements were consistently present in the current study:

“I probably don’t use [the manuals] like you’re supposed to, like scripted, you know, this bit, this bit, then this bit. I pick and choose bits that are going to work with *this* client, with where we’re at at *this* time.” (JPO1).

“...I knew that I’d come to [the project] I wanted to do something that was really creative, really innovative and really outside the box in terms of probation but using my past experiences when I worked in more supportive roles and stuff and building up, like, therapeutic relationships, so I was really keen to do all that....” (WPO2).

4. Mawby and Worrall argue that probation workers find meaning through professionalism, defined by the elements of credentialism, expertise and autonomy/legitimacy (2013, p. 144). As illustrated by the quote from WPO2 above, practitioners in the current study often referred to their previous accomplishments, training and experiences when explaining their place in the service and legitimising their role. Mawby and Worrall point out that the elements of professionalism act as “*boundaries* that distinguish between occupational groups, between an occupational group and lay people/clients and between an occupational group and the market” (2013, p. 144), creating a professional identity. This identity was sometimes expressed by practitioners in the current study as an uncomfortably privileged, but nevertheless real, status:

“It’s all well and good for us as...I don’t like using the word “professionals” but...from us, from the professionals’ side” (WPO1).

Mawby and Worrall (2013) identify a fifth element of probation culture of “*(re) presentation* in the turbulent conditions of the external environment” (p. 145) which relates to “media and public perceptions of probation as a socially tainted occupation, working with undeserving groups” (p. 18) and (perhaps disingenuous) attempts to maintain credibility with, and the confidence of, both offenders and criminal justice

partners (p. 145). Practitioners in the women's project adopted many of the strategies identified by Mawby and Worrall as dealing with uncertainty: "a self-effacing/apologetic/non-aggressive/unthreatening position in relation to public and media criticism; a long-sufferance that redefines success and lives with or rationalizes failure; a sense of solidarity and isolation ('nobody understands us') ameliorated by humour (often dark), blame, arrogance and moaning masochism" (p. 145). Further, the pilot was seen by practitioners as an opportunity to achieve positive representation in a turbulent environment (made more turbulent by the introduction of TR and division of the service), by explanation of and adherence to gender-responsive practice in public launch and PR events; with criminal justice agencies through increased partnership practices; and with the service-users themselves:

"I've seen a massive increase...in compliance. We just build up those... relationships in terms of how much there's more trust there, there's more respect...Especially knowing you've applied to come over to [the project], [service-users] know that and...you've almost built a connection there 'cause it's actually saying – 'd'yer know what? I recognise that women's needs are different to males' hence the reason why I've applied to come over to this pilot scheme and you know if it's successful hopefully it'll be rolled out in other areas'...and I think you build up an instant 'Ah! Respect!'" (WPO2).

Whilst JPACS was itself facing turbulent conditions at the time of the evaluation (with a proposal having been made that the Service should be brought under the jurisdiction of Jersey's Community and Constitutional Affairs Department) the element (of *[re]presentation*) was less readily apparent amongst practitioners who appeared to be somewhat insulated from the external environment, both by their confidence in the management team and by the clarity of their position within the courts. The Service's relationship with the courts was unequivocally identified as positive ("the courts are our friends, definitely, our critical friends" (JACPO)) and magistrates and Jurats were viewed as supportive of the practitioners' aims to care for and assist, rather than simply to control, offenders ("We'll go over [to the courts] and speak with [the magistrates] and agree how best to deal with situations if they happen. They'll listen to us and we might not always agree but we all want the same thing in the end, to help our clients." (JPO5)). The magistrate participant wholeheartedly echoed this sentiment, whilst identifying jurisdictional differences that enabled a "strong, stable line" from the courts for practitioners.

“We’re keen to help, keen to *let them* help...My perception is that probation always find the time to do – I won’t say ‘maternal’ or ‘paternal’ – supervision, a sort of benevolent keeping a watchful eye, a helping hand. Our perception over here of probation in the UK is of political interference. Political machinations get in the way of caring work.” (JMAG).

The impact of these “political machinations” of the States appeared largely absorbed by the management team, with practitioners making references to the proposed takeover as “not great” (JPO1) and “a bit worrying” (JPO3) but not presenting any significant obstacle to their daily work, whilst the need to represent the service positively to protect its position under the jurisdiction of the courts was a (if not *the*) primary concern of the management team in the commissioning of the CPAI evaluation and dissemination of the report.

“I’m waving [the report] at every opportunity – anyone has a pop at us and it’s one of the things I can use...to say...[the service is] working very well under its current stewardship....” (JCPO).

In terms of the “female dominated environment” identified in point 3, it is important to note that the practitioner (as well as the service-user) sample for in-depth qualitative feedback in the women’s project was, in fact, entirely female. Some male staff members from partner agencies attended stakeholder meetings but were not involved in observed service delivery and all the core project CRC and NPS practitioners (including managers) were female, so there were no male participants in the qualitative study. As far as generalisability to the general practitioner population across probation services is concerned, Mawby and Worrall (2013) reported a “feminization of service” represented by female workers making up approximately 68% of the probation workforce in 2013. They concluded at the time that “the organization is increasingly being defined by its female voice”. Whilst this suggests that findings from a women-focused study are of particularly relevant, it is clear that an *entirely* female sample is not representative of probation services *as a whole*. In this respect, the inclusion of male participants in the sample from Jersey was especially important to this study’s generalisability.

It is important to note that the characteristics of occupational culture in probation services in England and Wales may yet change in the new context of TR. This study

took place during a period of significant upheaval and uncertainty. From WPT being dissolved and the movement of staff to roles in the newly created CRC and NPS to the bidding process and acquisition of Wales CRC by Working Links and, finally, the ongoing processes of tendering and partnership, the fact that the sample continued to demonstrate the characteristics of the wider occupational culture of probation services identified and described by Mawby and Worrall (2013) is evidence of the concept's resilience. At the time of this study, staff expressed ambivalent feelings about a radical new approach which they reasoned might revitalise services with new ideas, direction and resources on the one hand, but was perceived to be disrupting practice, dehumanising the service and reducing autonomy on the other:

“We all received [a training package] and came out feeling very positive about it, but then the TR machine took over and the training ceased, wasn't rolled out to everyone and I think we're still waiting to see if that direction will be the way forward again. CRC employees don't know what the organisation is going to bring that buys the company. They may embrace [the training package] or bring in their own methods.” (WTM).

Mawby and Worrall caution that “attempts to dismantle or dilute these cultures by fragmenting probation work and parcelling it out to the lowest bidders, may [loosen] the ‘ties that bind’ probation workers to what was described to us as an ‘honourable profession’” (2013, p. 154). It is, therefore, an inescapable limitation of this study that evaluation and its processes would need to be re-thought if the occupational culture of probation services in the UK becomes markedly different under TR.

There were also practical considerations relating to the CPAI-2010 itself. As discussed, investigator triangulation is used to minimise bias. However, there is only one CPAI-2010 accredited evaluator working in adult services in England and Wales. To overcome this limitation in this study, audio recordings of three of the supervision sessions observed for evaluation of the women's project were shared with CPAI-2010 developers Paul Gendreau and Yvette Thériault, and Paula Smith (who is responsible for overseeing the use of the CPAI-2010 in the US). The final scores for section G were then agreed upon following discussion and clarification between the researcher in Wales and the Canadian/US evaluation team.

Additionally, according to its developers the CPAI-2010 should ideally be administered over a short period of time – preferably one week or less – in order to ensure that the data collected is representative of day-to-day practice and provides a ‘snapshot’ of the service for evaluation. Many factors prevented data collection being completed to these timescales in the evaluation of the women’s project. To some extent, this was due to complexities of the pilot itself – ie, the need to identify and collect data from several different agencies at several different stages whilst the pilot was in its early development and not coherent as a result. As discussed, this evaluation also took place during a period of substantial upheaval, at a time when WPT was dissolved and replaced by the CRC and NPS. Numerous related and unrelated factors including staffing issues, bidding issues (leading to the dissolution of one partner service), delays in the roll-out of the pilot, teething issues, access issues, false starts, communication lapses, etc, all resulted in the lengthy data collection period. Whilst this clearly is not ideal with regard to the integrity of the CPAI-2010, following discussions with the Project Manager, the decision was made to extend the timescale as far as was needed to complete the full evaluation. SSET recognises the need to be responsive to the agency under evaluation and in this way hopes to have provided the opportunity to account for the many changing pressures and achievements peculiar to the evaluated project. However, the extended timescale also had implications for the feedback phase of the study. Many of the sample were no longer available to participate due to factors such as staff redeployment and the completion of Orders (and engagement with the project) by service-user participants. The evaluation of JPACS was undertaken to CPAI-2010 timescales (one week of on-site data collection) but there was no opportunity to engage any service-user participants in Phase Two of the study (following dissemination of the report) in either Wales or Jersey (due to factors such as Order completion and disengagement from service as well as participant availability during fieldwork for Phase Two).

One of the most important contributions of the desistance literature is that “both ex/offenders’ and practitioners’ voices need to be respected and heard” in debates around evidence-based practice (McNeill et al, 2012, p. 35). The inclusion of service-user participants’ contributions from Phase One is therefore extremely important and

practitioners' perspectives from Phase Two on whether/how their service-users access/assimilate the reports provide some insight, though further work would need to be undertaken fully to investigate service-users' perceptions of the value and methods of evaluation and dissemination of findings.

Contrasts of Location in the Research Experience

Whilst the values of practitioners in both the women's project and JPACS appear broadly representative of probation in the British Isles (as conceptualised by Mawby and Worrall, 2013), there were significant differences in the research experience between locations, and these appear to reflect differences in agency culture, management style and political approach.

The organisational response to evaluation in JPACS was systematic and harmonious. Negotiations between the evaluator and management team enabled timetables and processes for data collection to be arranged in advance of fieldwork and these were adhered to throughout the evaluation. The probation team, service-users, partner practitioners and other key stakeholders (eg, magistrates) were prepared for, and supportive of, the evaluation and the wider research; all necessary resources were made available to the evaluator (a private work space, access to case files, etc); and the evaluation findings were disseminated throughout the service and published on its website.

In the women's project, an unstable organisational context resulted in a turbulent research experience from the outset. Initial negotiations with WPT were disrupted by its dissolution and the establishment of Wales CRC and the NPS. There were a number of false starts to data collection as the women's project faced significant issues with resources, including the allocation of premises being delayed, difficulties regarding file access and problems with staff designation related to the restructuring of probation and division of employment under the CRC and NPS. Once underway, the evaluation process continued to be subject to the unpredictability of organisational instability, necessitating the extension of the data collection period from one week to six months.

Planned sessions for observation failed to take place; most interview appointments were either forgotten, ignored, cancelled or rescheduled at the last minute; and bidding issues preoccupied partner agencies (one of which was dissolved without warning during the evaluation, another shortly after). As discussed, dissemination of the evaluation findings was also complicated by managerial concerns regarding the potential of the Unsatisfactory categorisation to diminish the project's status in the open market, the resulting suppression of findings having ethical as well as practical implications for the research.

Differences in attitudes and responses to evaluation between the women's project and JPACS, and the reasons for these differences will be explored in Chapters Four and Five, which relate the findings of Phase One and Phase Two of the study.

Conclusion

In conclusion, a pragmatic methodology was applied to address the research questions raised at the end of Chapter Two. The study was able to incorporate the positivist methodology of the CPAI-2010, take into account qualitative measures regarding the lived experiences of evaluation and strengthen the validity of findings through triangulation of data due to its independence from a single, fixed epistemological or ontological paradigm.

As a result of the characteristics of the sample services, access and ethical considerations were guided by agency protocol, risk and vulnerability, gender issues and the turbulent external environment at a time of major change.

Whilst limited by the small sample size, research such as Mawby and Worrall's 2013 study (see also Deering, 2011; Robinson et al, 2014) has shown that there is homogeneity in terms of probation cultures in England and Wales and the women's project sample appears to be unexceptional in this aspect, with probation officers in Jersey also demonstrating the identified elements of probation culture to a large degree.

It seems reasonable to extrapolate that evaluation systems and approaches that meet the demands of the sample projects will have wider applicability.

Table 3.2 Study Participants

Phase One				
	<i>Interviews</i>	<i>Case file-reads</i>	<i>Observations</i>	<i>Total participants (minus duplications)¹⁹</i>
<i>Women's Project</i>	Project manager (F) Research officer (F) Team manager (F) 5 Probation officers (2 NPS, 3 CRC, all F) 2 Partner agency managers (F) 3 Service-users (F) Partner agency practitioner (F) Total: 14 (F)	20 active and inactive cases (F) Total: 20 (F)	5 Probation officers (F) Team manager (F) 3 Probation support officers (F) 14 Service-users (F) 7 Partner agency practitioners (6F, 1M) 3 Partner agency managers (F) 3 Administrative staff (F) Total: 36 (35F, 1M)	55 (54F, 1M)
<i>JPACS</i>	Chief Probation Officer (M) Assistant Chief Probation Officer (M) Team leader (M) 5 Probation officers (3F, 2M) Research & Information Officer (F) 2 Practitioners (1F, 1M) Magistrate (M) 6 Service-users (M) Total: 18 (5F, 13M)	20 active and inactive cases (3F, 17M) Total: 20 (3F, 17M)	Team leader (M) Assistant Chief Probation Officer (M) 7 Probation officers (5F, 2M) 2 Practitioners (M) 20 Service-users (4F, 16M) 4 Administrative staff (3F, 1M) Total: 35 (12F, 23M)	56 (17F, 39M)
<i>Phase Two Interviews (subset of Phase One participants)</i>				
<i>Women's Project</i>	Project Manager (F) 4 Probation officers (2 NPS, 2 CRC, all F) 1 Partner agency manager (F) Total: 6 (F)			
<i>JPACS</i>	Chief Probation Officer (M) Assistant Chief Probation Officer (M) Probation team leader (M) 4 Probation officers (all F) Total: 7 (4F, 3M)			
				TOTAL: 111 (71F, 40M)

¹⁹ Where an individual has participated in more than one element (eg, interview *and* observation), they have been counted once for the purpose of the total figure.

CHAPTER FOUR

Findings from Phase One:

Correctional Program Assessment Inventory-2010 (CPAI-2010) Evaluation Results

Introduction

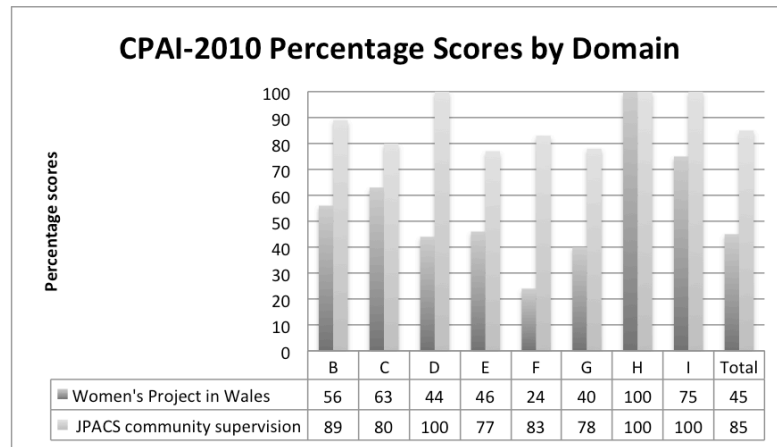
As discussed in the methodology, CPAI-2010 evaluation involves the scoring of the following domains:

B. Organizational culture	9 items
C. Program implementation/maintenance	10 items
D. Management/staff characteristics	18 items
E. Client risk/need practices	13 items
F. Program characteristics	25 items
G. Core correctional practice	45 items
H. Inter-agency communication	5 items
I. Evaluation	8 items

Domain A of the CPAI-2010, programme demographics, is descriptive only and does not contribute to its scoring. As mentioned in footnote 19, p. 77, there are consequences for assuming that descriptive information can be included uncritically and this will be explored further on pp. 166-168.

As far as the scored domains are concerned, the evaluations undertaken for this study revealed significant differences in each service's adherence to the principles of effective intervention.

Fig 4.1 CPAI-2010 Evaluation Scores by Domain



As can be seen, Jersey Probation and After-Care Service (JPACS) out-performed the women's project in all but one domain (H: Inter-agency communication, in which each service scored the maximum 100%). Overall, the women's project scored 45% and was categorised as Unsatisfactory, whilst JPACS scored 85% and was categorised as Very Satisfactory. In this respect, JPACS is somewhat exceptional in terms of international findings from CPAI evaluations. In their assessments of nearly 400 services, University of Cincinnati researchers categorised 75% as either Unsatisfactory or needing improvement, with only 7% being categorised as Very Satisfactory (Latessa & Smith, 2015, p. 391).

This chapter explores the findings of the CPAI-2010 evaluations of the women's project and JPACS by domain to discuss the key issues that were identified; contextualise these through reference to the literature relating to effective practice (particularly Latessa and Holsinger's 1998 summary of findings from 51 CPAI evaluations undertaken in the US); and explore the sample agencies' responses regarding the usefulness of findings and recommendations generated by the evaluation.

CPAI-2010 Results, Recommendations and Agency Responses by Domain

Domain B: Organisational Culture

	Available points	# of points	N/A	% score	CR
Women's project	09	05	0	56%	4
JPACS	09	08	0	89%	4.6

It is widely acknowledged that the collective values, principles and behaviours of organisational members are crucial in the implementation of services (Harris & Smith, 1996). In this domain, the women's project achieved a Satisfactory categorisation, whilst JPACS was categorised as Very Satisfactory.

The average confidence ratings (CR) indicate how reliable the scores are in each domain according to the quality of evidence supporting the scoring for each item at the time of data collection. In order to be given the maximum CR of 5, an item would need to be evidenced thoroughly. Where, for example, an item requires the service to use a written document (eg, protocol for procedures, assessment tool, code of ethics, etc), the CPAI-2010 evaluator would need to see that document in order to give a CR of 5. If interviewees state that the document exists but it is not available to be seen at the time of data collection, the evaluator may still score the item "yes" but would lower the CR for that item according to the available evidence. If an average CR of 3 or above is given to a domain, this would indicate that the rater is at least moderately confident about the reliability/validity of the information gathered. A CR of below 3 would require the evaluator to be very cautious in their final conclusions as to a service's effectiveness and a CR of below 3 across all items or domains would result in an incomplete assessment. In this domain, the evaluator was able to review sufficient supporting evidence in both services to give a high confidence rating overall.

Both the women's project and JPACS were found to have clearly defined their goals and mandate in formal documents, a crucial component of implementation since, as Harris and Smith (1996) point out, a lack of specificity about goals "leads to confusion,

co-optation, inappropriate decisions, and an inability to assess the success or failure of a policy or program” (p. 204; see also Casper & Brereton, 1984; Kelling et al, 1986; MacKenzie et al, 1993). However, neither service had a documented code of ethics. This was identified as a common weakness across the US programmes assessed by Latessa and Holsinger (1998) and the findings from this study indicate that this appears to be a neglected aspect of service development in the British Isles, with individual staff members subscribing to different professional standards of conduct, performance and ethics depending on their registration (as social workers under the Health & Care Professions Council, for example) or governance by a wider body (eg, The Civil Service Code). In the multi-agency context of women’s services and with the mixed practitioner base of social-work, probation, States and third-sector workers of JPACS, a recommendation of both evaluations was that there is the specific need for disparate agencies and staff members to be formally united under a shared, organisationally-specific value base, providing guidance and accountability for all staff working within, or in partnership with, each service. The management teams at both the women’s project and JPACS were receptive to this recommendation, describing the lack of ethical guidelines specific to their service as an oversight that needs to be addressed:

“The evaluation did highlight some things...that we hadn’t thought about developing, like the ethical code, for example, which is really useful.” (WPM).

“I’ve actually been saying for a long, long time that we actually haven’t looked at having a service statement, an ethics statement... We’ve got some people who’ve come across from a social work background, we’ve got some people who were probation ... People do have different sets of values and beliefs ... and actually talking about the philosophy of the service ... the ethics, I think that was a really, really valid point because I think, unless we’re singing from the same hymn sheet it’s very difficult to actually go forward.” (JTL).

Both the women’s project and JPACS were found to respond to new initiatives within a reasonable length of time. They also had procedures in place for information sharing and self-evaluation and involved external researchers to advise on aspects of programming, though the evaluator had less confidence in scoring for the women’s project in these areas, particularly regarding researcher involvement. As an established service, JPACS’s collaboration with external consultants was embedded into service delivery and assessment, with several internal quantitative assessments (largely facilitated by the Research and Information Officer working in collaboration with

external researchers) made readily available on the service's website. The women's project, however, was being piloted at the time of the evaluation and so the reduced confidence scores reflected that plans to use external consultants to advise on programming and the establishment of designated internal evaluation and research services, whilst identified by the evaluation as a distinct strength, were not yet fully developed or in place.

Where JPACS clearly out-performed the women's project was in items relating to organisational harmony, staff turnover and response to problematic issues, items that are shown to be inter-related as well as inextricably linked with management characteristics in each evaluation. The CPAI evaluation report for JPACS identifies a "hands-on" managerial approach to be integral in facilitating a collegiate atmosphere in which problems are addressed effectively, resulting in high levels of staff satisfaction and retention:

There are clear procedures for responding to problems within the team, which are routinely dealt with using a non-confrontational, non-crisis approach. In general, problems are dealt with through discussion and mediation, enabled by a supportive and hands-on managerial approach by all members of senior staff, whose offices share corridors with practitioners, operating an open-door policy. In interview, all staff were aware of the formal procedures for grievance and disciplinaries, etc, and relevant documents are readily available, though at the time of the evaluation there had been no formal grievance for a number of years. There is a collegiate atmosphere in staff meetings and day-to-day practices. A number of practitioners identified the support, knowledge and experience of their colleagues as crucial resources in their daily work.

...High levels of organisational harmony are reflected in extremely high levels of staff retention, with 100% of staff surveyed having remained with JPACS for more than two years. In fact, over 70% of staff members involved in the evaluation had worked in the Service for over ten years. (Kerr, 2016, pp. 10-11).

Conversely, the CPAI report for the women's project found that low morale amongst practitioners and a high staff turnover resulted from a lack of coherent managerial support:

There is no clear, coherent process for dealing with problems, resulting in organisational disharmony when difficulties arise. This is due, in part, to a complicated and unwieldy management structure ... At present the programme is at risk of low morale amongst practitioners and high staff turnover, which would reduce the effectiveness of services. (Kerr, 2014, p. 13).

In stark contrast to the shared spaces, “hands-on” approach and “open-door policy” of management in JPACS, National Probation Service (NPS) and Community Rehabilitation Company (CRC) practitioners in the women’s project were line-managed by staff unattached to the project and based off-site at a local probation office, where all support staff (admin, IT, finance, etc) were also based. The evaluation notes that:

This lack of on-site management, supervision and administrative support is contributing to the team’s isolation from the service as a whole, a lack of oversight (regarding hours worked, working practices, programme logistics, etc), limited managerial support, the absence of a clearly delineated structure for addressing problems and low morale amongst staff members. (Kerr, 2014, p. 17).

Referring to staff dissatisfaction expressed throughout the evaluation; to an incident in which a member of staff was assaulted by a service-user on project premises (during which a lack of alarms or procedures for such incidents resulted in staff in other areas of the building being unaware that the incident was occurring); and to managerial factors emerging as a recurrent deficit across the domains pertaining to organisational culture, programme implementation/maintenance and management/staff characteristics, the report concluded that deficits in the management structure “have the potential to irreparably damage the cohesiveness of the organisation and possibly result in further serious incidents” (Kerr, 2014, p. 31).

Recommendations made within the domain relating to organisational culture were made within a broader context, which required the organisation to significantly re-structure its managerial approach:

To stabilise the programme, ensure organisational harmony and consolidate processes, the management structure must be cohesive and clearly defined, with appropriate policies and procedures in place to respond to problems in a non-confrontational, crisis-free fashion. (Kerr, 2014, p. 13).

In order fully to contextualise the issues and agency responses to recommendations regarding management of the women’s project, these will be discussed in further detail under domain D (management/staff characteristics).

Domain C: Programme Implementation/Maintenance

	Available points	# of points	N/A	% score	CR
Women's project	10	05	2	63%	4.4
JPACS	10	08	0	80%	4.3

One of the main conditions associated with successful implementation is that “there is a close fit between the programme and the environment in which it is enacted” (Harris & Smith, 1996, pp. 194-196), an environment which must be stable as well as supportive to maintain programme effectiveness. Additionally, effective programmes are based on strong theoretical models derived from the treatment literature and are begun on a pilot basis to work out the logistics.

In this domain, the women's project achieved a Satisfactory categorisation, whilst JPACS was categorised as Very Satisfactory. Two items were voided as not applicable in the evaluation of the women's project. These related to established programmes and were concerned with years in operation and changes in staffing levels so could not apply to a pilot project.

Both the women's project and JPACS were faced with far-reaching and contentious issues at the time of the evaluations and these reduced the stability of both settings, though not to an equal extent. JPACS had seen both reduced funding from the States and a proposal that the service should be brought under the jurisdiction of Jersey's Community and Constitutional Affairs Department, threatening its long-established governance by the courts. However, the staffing levels of probation officers had been prioritised and made resilient to funding cuts by the management team and a sustainable budget for the service had been agreed for four years, providing some stability in terms of the service's fiscal responsibility.

The women's project, on the other hand, was set up at a time of significant upheaval in the field of criminal and community justice, due to the division of probation services into the NPS and Wales CRC. The evaluation found that the organisational context of

the women's project, whereby staff had been relocated to premises without on-site supervision, managerial or administrative support, was unable to provide the stable and supportive working environment necessary to proof the service against external pressures. Furthermore, the quality of the project was dependent on the stability of stakeholder agencies, which were operating in a context of funding cuts and increased competition. Funding for these services, and indeed for the project itself, was not guaranteed beyond the short term and one partner service was, in fact, dissolved during the evaluation period. Where the management team and governing board of JPACS were recognised by the CPAI for taking a strong stance in negotiations, priority setting, resourcing and commissioning for the service, evaluation recommendations for the women's project acknowledged the unique challenge in the administration of integrated, holistic services, recommending that the project must be "responsive to the differing pressures on third sector agencies in particular", whilst accepting the potential for increased sustainability if third sector agencies' "presence in a competitive environment is strengthened by their participation" in the project through, for example, collaborative bidding processes (Kerr, 2014, p.15).

In their evaluations, Latessa and Holsinger (1998) found that the majority of services were consistent with existing values in the community or institution in which they resided. Both the women's project and JPACS were found to be well-fitted to their environment in this respect. JPACS has developed a culturally-responsive range of services based on consultation with community agencies and statutory bodies in Jersey and the women's project embodied the principles of Integrated Offender Management (IOM) Cymru with its focus on the re-configuration of existing local resources to manage all women in the criminal justice system in partnership. Additionally, in line with the IOM principle of providing "a local solution for local problems" (Home Office, 2013), the women's project was actively engaged in the collection of individual-level survey data on the need for the service and had formed focus groups of local professionals, academics and service-users to facilitate this. The recommendation in the evaluation report for JPACS that it "may benefit from taking a more structured approach to identifying gaps in service provision and/or emerging issues through the use of, for example, focus groups and/or surveys" (Kerr, 2016, p. 14) was met with

enthusiasm and resource planning, demonstrating the responsiveness made possible by coherent organisation:

[Reading from the CPAI report] ‘Using focus groups or surveys to identify gaps in service or emerging issues’ ... Interestingly, [a member of the team is] going on to a management course and she has to have a project and that’s most probably the one that we will focus on. So, again, you identified that and that was useful.... (JTL).

The twin flaws commonly evident in programme implementation and maintenance are: scant regard given to the treatment literature; and a failure to pilot treatment components before full implementation (Latessa & Holsinger, 1998). In fact, both JPACS and the women’s project were found to have a strong commitment to piloting interventions (the women’s project itself was being piloted at the time of the evaluation), though only JPACS had undertaken a review of the international literature relating to effective practice. Independent from strategic and political developments in probation in England and Wales, the development of JPACS’s service has been shaped as a result of its longstanding relationships with academics, researchers and professionals in the field of criminology, which have engaged managerial staff in an ongoing review of the literature regarding ‘what works’ in offender treatment. Though a literature review had been undertaken by the women’s project, this was focused on the demographics of the female offender population rather than on programme materials directly related to effective treatment using cognitive, behavioural or social learning modalities.

Domain D: Management/Staff Characteristics

	Available points	# of points	N/A	% score	CR
Women’s project	18	08	0	44%	4.4
JPACS	18	18	0	100%	4.7

Latessa and Holsinger identified “a great deal of variation from program to program regarding staff quality” (1998, p. 26) and this was evident in the current study, with the women’s project being categorised as Unsatisfactory in the domain relating to management and staff characteristics, whilst JPACS received a score of 100%. As with the majority of Latessa and Holsinger’s evaluations, the current study found educated,

knowledgeable and experienced managerial and practitioner staff working in both the women's project and JPACS. Some weaknesses in the women's project were also common amongst the services evaluated by Latessa and Holsinger. For example, there was no schedule in place to provide regular training in effective practice, and there was no clinical supervision of practice. Of their evaluations, Latessa and Holsinger observed that:

...we rarely found staff who had received sufficient training on the interventions and treatments utilized by the program. Clinical supervision was not provided routinely, and staff were rarely assessed on service-delivery skills. (Latessa & Holsinger, 1998, pp. 26-27).

In this respect, JPACS is an exceptional service. A comprehensive programme of training is provided to staff, with training in the theory and practice of effective interventions being delivered two or three times annually by a consultant forensic psychologist. Additional training is sourced through JPACS's research relationships with criminologists in Swansea University and the Cognitive Centre Foundation as well as links with eminent academics (eg, Professor Chris Trotter of Monash University, Melbourne) who provide specialist training in specific skills. Training in pro-social modelling and problem-solving is delivered to all members of staff, including the administration team and all members of senior management, demonstrating JPACS's considerable commitment to evidence-based practice throughout the service. There is an annual Performance Review and Appraisal process and monthly supervision of probation officers. JPACS has also implemented use of the Jersey Interview Skills Checklist (Raynor et al, 2009) to undertake a structured assessment of practitioner skills in video-recorded intervention sessions as part of the clinical supervision process.

The predominant issues revealed by evaluation in this domain again relate to inadequate organisational structures and processes in the women's project. In Latessa and Holsinger's evaluations, a strength of services was that programme directors "tend to be involved in the hiring and training of staff; in many instances they provide some direct services to offenders" and that "often staff were selected on personal characteristics, such as life experience, fairness, firmness and problem-solving skills" (1998, p. 26). This was patently the case in JPACS, where The Chief and Assistant Chief Probation

Officers and Team Leader all have an intimate knowledge of the service and staff and are involved in the hiring and supervision of practitioners and direct service delivery (for example, covering office duty on occasion) and all staff are selected for employment based on assessment of their relationship skills and skills specific to evidence-based practice (eg, pro-social modelling) using vignettes.

Staff at every level in JPACS identified the strong operational structure, facilitated by consistent and supportive management, as enabling and enriching practice and fostering resilience and job satisfaction:

It's very rewarding working here. There's lots of backing from staff, I'm very well supported. You're never out on a limb here – there's always advice available and the management are great. It means I can work in creative new ways when I identify a need, like, for constructive activities. (JP1).

We're very supported by management and we're all there for each other – it's a very strong team. We give each other assurance but also different viewpoints. There's a vast array of knowledge, experience and expertise here and we all benefit from it – from everything everybody brings. (JPO3).

I hope that people feel that they have a clear framework that they feel supported to make mistakes within. How you learn is by making mistakes. (JCPO).

In the women's project, however, there was no on-site member of staff with a managerial role in the selection, training or supervision of practitioners. In fact, the division of the probation service under Transforming Rehabilitation (TR) was so extreme that probation practitioners within the project were being line-managed by different, off-site CRC and NPS staff with little or no knowledge of the project and its service-users. Moreover, the evaluation report recommendation that “[the project] practitioners must be managed as a team, which the current management structure cannot achieve” (Kerr, 2014, p. 18) was seen as “not physically feasible” by the Project Manager, who described having little authority or influence in this respect in a context of even greater disruption and division across probation services as a whole:

I absolutely agree that [the project CRC and NPS practitioners] should be [managed together as one team]...but...that's just something to do with just how life is with those organisations at the moment...it's not physically feasible...it doesn't do what we would want it to do really but actually compared with the status quo for work in other projects...it's actually ... quite good...I would still like us to work towards a situation where they are managed more holistically. Even though right now we can't do that, it's something I keep harping on about, you know, all the time. (WPM).

Some research studies have found correlations between organisational culture, leadership behaviour and job satisfaction. For example, Tsai (2011) found that “when the interaction between the leadership and employees is good, the latter will make a greater contribution to team communication and collaboration, and will also be encouraged to accomplish the mission and objectives assigned by the organization, thereby enhancing job satisfaction” (p. 2). Moreover, it is unarguably the case that the organisational context of services had a knock-on effect across all areas of practice in these evaluations. For example, staff at the women’s project identified their previous experience (under Wales Probation Trust) of the Skills for Effective Engagement, Development and Supervision (SEEDS) training programme as positive, key to which is the development of practitioner skills and practice in working with service-users as well as ongoing supervision of practice, though this was discontinued during the restructuring of services and no elements of the programme appeared to have survived the TR transition to NPS and CRC.

We completed part 1 of the SEEDS training but part 2 never happened. It’s not useful just knowing the stuff – there’s nothing to say you’re doing the work you should be... There’s no... assessment of your practice... it’s just assumed you’ve got this level of knowledge and ability and they’re like “just go, go, go and do it!” (WPO3).

Despite their shared, unequivocal belief in rehabilitation and consistently high level of investment in the project, both CRC and NPS practitioners working in the women’s project had little confidence in its organisation, which appeared to lack stability, structure, resources and direction:

I don’t even know who runs [the project] any more. (WPO3).

I understand the vision, it’s just how we’re going to physically do it. It’s trial and error which is a bit concerning. I was fearful they’d try to put groups over here straight away when we need to do the bread and butter stuff first – day-to-day offender management. Have we got enough rooms? That kind of thing. We’re finding people are going to [the probation office] when they’re supposed to be coming to [the women’s project] for their appointments. Inductions from court have been sent there instead of here. (WPO2).

My concern is, we’re so short-staffed and we’ve hit the wall at 100 miles per hour – we didn’t get it up and running pre-launch. (WPO1).

Furthermore, managerial staff in partner agencies identified the organisational context of the women's project as alienating and oppressive due to imposed processes and inadequate communication:

My concerns are that a lot of it is being process driven – done to us rather than with us...They do ask for our feedback. I commented on things, but I don't know whether the feedback was acted upon. (WPA2M).

It feels sometimes that the process is a little bit top down. So...even though we are a partnership, I don't feel that we're always treated as *equal* partners, sometimes it feels like we're shared *certain* information, we have *certain* decision making powers.... (WPA1M).

The lack of coherent organisational context was also felt by service-users to varying degrees. During informal observations, several service-users commented or complained to each other, practitioners or administrative staff about unclear arrangements, delays, changes in workers, misunderstandings, lack of communication, etc (nine separate instances were noted in the observation records for the evaluation period). Interestingly, this did not appear to have a significantly detrimental impact on service-users' view of the project. In fact, all service-users consistently identified the women's project as comfortable, accessible, friendly and safe in contrast to their previous experiences of the probation office, which they described as institutional, uncaring and even unsafe in some cases, due to the presence of male probationers. Nevertheless, the organisational issues experienced by service-users, such as upheaval under TR and a lack of communication/information, were frequently mentioned in evaluation interviews, sometimes as fundamental to their perception of the project. In response to the question "what does [the women's project] mean to you?" one service-user replied:

That means to me that, erm...there's different kinds of support, it's meant to be, erm...it's probation...but they've changed it to the public sector [*sic*], haven't they? Which I haven't been explained a lot about. (WSU1).

Conversely, there was no ambiguity about (or discussion of) JPACS's organisational context in service-users' responses in Jersey. Instead, when questioned about JPACS, service-users focused on their relationships with staff and the nature of support provided to them. When asked specifically whether JPACS as an organisation (of staff beyond individual case workers) was supportive, service-users were unanimous in agreement, and their perceptions of organisational support were framed by experiences of good communication practices within a coherent service:

Yeah, [JPACS staff are] helpful...if I phone them up and [say] “can you put me through”, or [my probation officer] might be out at the prison or another meeting or whatever, then they’ll let her know and then she phones me back straight away when she’s back. (JSU3).

Everything’s been great so far. I met...[my probation officer’s] boss one day, when he thanked me for doing a recorded interview. (JSU2).

Domain E: Client Risk/Need Practices

	Available points	# of points	N/A	% score	CR
Women’s project	13	06	0	46%	4.6
JPACS	13	10	0	77%	4.2

A special strength of the women’s project was its commitment to diverting low-risk offenders from the Criminal Justice System (CJS). This reflects the evidence base, which indicates that successful offending behaviour programmes target participants presenting a medium to high risk of re-offending, at the same time taking a humanitarian and holistic approach to ensuring that services are in place to respond to the needs of women outside of the CJS, ie, not ‘diversion to nothing’. However, CPAI evaluation found that practice in assessing and classifying service-users was Unsatisfactory. This was by no means uncommon in the services evaluated by Latessa and Holsinger, who observed that:

Strengths: The vast majority of programs studied have stated criteria for admissions, receive appropriate clients, and have a rational/legal clinical basis for the exclusion of certain types of offenders. We also found that, in general, most programs attempt to assess some offender characteristics related to risk and need.

Weaknesses: While many programs did indeed attempt to assess offenders regarding risk and need, doing so did not involve incorporation of a standardized, objective, actuarial instrument...Even when a standardized assessment is being performed at some point in the offender’s entry/progress, it is seldom found that the information gathered is being used to distinguish offenders by risk. In other words, even when proper (and potentially beneficial) assessments are being performed, the information is not influencing the decision-making process, let alone service delivery. In addition, it is generally found that staff assessments of offenders are based on a quasi-clinical approach that does not result in a summary score. Likewise, it has been very rare to find that programs are routinely measuring with standardized instruments responsivity characteristics [*sic*], such as levels of motivation, intelligence or psychological development. (Latessa & Holsinger, 1998, p. 26).

In fact, statutory service-users of the women's project (ie, those entering the project on Court Orders or prison licenses) *were* assessed using a valid instrument (the Offender Assessment System, OASys), incorporating an actuarial predictor (the Offender Group Reconviction Scale, OGRS) which summarises offenders' personal characteristics as to the level of risk they present. OASys was re-administered periodically, and service-users appropriately re-classified. However, the project was unable to score on some related items because no assessment of risk of re-offending was undertaken of non-statutory service-users. Instead, these were referred to services based on an assessment of their eligibility for diversion (which did not necessarily preclude those presenting a higher risk of recidivism), and areas of need related to re-offending (the project had developed its own evidence-based needs assessment). In short, neither the decision-making process in terms of case management and referrals nor the delivery of services was consistently influenced by the risk levels of service-users. Conversely, JPACS's categorisation as Very Satisfactory in this domain was largely due to its *consistent* and periodic use of a valid instrument (the Level of Service Inventory – Revised, or LSI-R), which summarises clients' personal characteristics relative to the level of risk they present.

Again, it appears that the organisational context of services was a predominant factor in these differences. Independent from the National Offender Management Service (NOMS; now Her Majesty's Prison and Probation Service), JPACS has adopted and maintained its assessment instruments and approaches supported by ongoing independent consultation with academics, engagement with the international literature regarding effective practice and allocation of funds and resources to acquiring and training in use of the LSI-R. In the women's project, the periodic administration of OASys is built into National Standards for the core agencies' (NPS and CRC) management of cases (though it should be noted that NOMS research suggests that practice in the re-assessment of offenders in England and Wales is "sub-optimal in many cases", with assessments neglecting to reflect changes resulting from external events or supervision (Howard & Moore, 2009)). With OASys licensed and available solely for use on statutory cases under NOMS, a recommendation of the evaluation was that "a validated risk assessment tool such as the Level of Service Inventory – Revised –

Screening Version (LSI-R-SV) should be used on all admissions...which should then be used to inform the level of interventions delivered” (Kerr, 2014, p. 20). Again, this was considered desirable but unfeasible, due to a lack of resources: “It would be great if we could do it, but the money just isn’t there” (WPM).

In line with Latessa and Holsinger’s findings, responsivity assessment and matching were neglected areas of practice in each service, though some consideration was given to the matching of service-user to service, with JPACS undertaking some in-house assessments of, for example, basic skills, and utilising external providers and partner agencies for more advanced screening of, for example, mental health and psychometric testing in specific cases. Central to the women’s project was a commitment to gender-responsivity through the provision of women-only spaces and holistic interventions delivered in partnership with local providers, especially focusing on mental health, substance misuse, experiences of abuse and trauma, parenting and sex work. Beyond this, both JPACS and the women’s project monitored ethnicity, disability and language, though, in practice, these were largely actioned in an *ad hoc* and informal way. The outcomes of individual responsivity assessment were more closely applied in adjustments to aspects of service delivery in JPACS where, for example, assessment of motivation levels informed the provision of work to address substance misuse related offending, with sessions corresponding to the appropriate stage of Prochaska and DiClemente’s transtheoretical model (1982). Where relevant, assessments undertaken by external consultants also informed service delivery and were considered during decision-making processes regarding, for example, whether a group environment would be appropriate for a specific service-user.

Though practitioners in the women’s project demonstrated some awareness of the need to respond to different learning styles and personal characteristics to engage individuals and enhance service provision, assessment of these was consistently described as informal and intuitive (“[service-users are] assessed at induction stage, based on our experience...issues that we observe...” (WPO3)) and, in spite of practitioners indicating that they responded to diverse characteristics in their work (“If someone’s learning a certain way, you go with it.” (WPO1)), they were unable to provide examples of how

this was done (in contrast, practitioners in JPACS referenced the use of flip-charts and flash cards with visual learners as an example) and there was, in fact, little differentiation between individual service-users in service delivery. Initiatives to enhance services with, say, the development of mindfulness interventions with a local healthcare provider, were seen as appropriate for *all* women.

One element that stood out in evaluations, both for its absence in practice and agencies' resistance to related recommendations, was the formal matching of staff both to services and to service-users based on personal characteristics. Both agencies indicated that some staff were allocated to certain activities and/or service-users according to factors such as professional interest/experience in a relevant area or language needs. However, practitioners were most often assigned to service activities and service-users to staff according to caseload capacity and availability; or using experience, intuition and perception of skill and need rather than any formal assessment of staff characteristics and matching process. Research has found that the assessment of characteristics such as the conceptual and interpersonal maturity levels (I-level) of both staff and service-users can facilitate a process which ensures the most appropriate match for effective differential treatment (Andrews, Bonta & Hoge, 1990; Gendreau, 1996a; Palmer, 1963; 1965; 1967; 1971; 1996; Warren, 1971). Whilst a rating inventory for the selection and matching of practitioners was developed by Palmer for the California Youth Authority's Community Treatment Project (a large-scale project from which significant findings regarding I-level matching emanated), this is not easily accessible and there appears to be a significant lack of resources available to agencies wishing to pursue matching according to such an assessment. Similarly, the conceptual-level matching model (CLMM) in corrections expounded by, for example, Reitsma-Street and Leschied (1988) is dependent upon use of instruments such as the Paragraph Completion Method (PCM, Hunt et al, 1978), which must be administered and interpreted by trained raters. It was certainly the case that both JPACS and the women's project had very little point of reference from which to engage with the terminology or processes of I-level or CLMM assessment, and their responses to recommendations that "staff should be assigned to service-users with whom they can work effectively – eg, staff would be assigned to service-users with a similar conceptual level" (Kerr, 2014, pp. 20-21; 2016,

p. 18) identified case allocation according to formal assessment as entirely impractical and not necessarily desirable:

Basically, our first referral point is a Social Enquiry Report from the court where very little information is known about that individual...so you base [allocation] on the offences...there might be some factors [which influence allocation]...there is a certain numbers game...We also try the continuity thing, so the person who writes the Social Enquiry Report, we do try, from beginning to end...there is something in the SER writer's relationship....

...All I can say is there is a process, it's not based entirely on numbers and there is a trying to fit people [but] we don't have specialisms in the office. I think we have what I would describe as areas of interest...rather than areas of expertise...and everybody says they want to remain [generic officers]. So, if...a person is saying they want to remain generic that means that basically whoever comes in through the door they'll need to deal with but I do accept that there are client attributes so, for instance, if I had a female client specifically saying "I don't want to work with a man", I would take that into account. (JTL).

Potentially, you could have some, or a couple, of probation officers with a certain style and lots of offenders with a similar one and you'd end up with them being overloaded... and what would happen to the ones that weren't [matched]?...I just don't know how realistic, given our [service-user] numbers and our staff numbers...how that would work. (JPO4).

Cases are assigned based on risk levels, child protection issues...a lot depends on capacity at the moment. That's just the way it is. Things are tight, we're understaffed...I can re-allocate if there's a problem but...all of our officers need to be able to build good relationships with a range of women. (WTM).

The services' emphasis on the development of generic skills and processes aimed at developing and maintaining positive relationships (which will be explored further under Section G) is also supported by the literature. Palmer's 1996 review of research into staff characteristics found that, whilst some differences did occur on matching, the results of studies strongly favoured the greater effectiveness of practitioners with high empathy/socialization personality dispositions (Andrews & Kiessling, 1980) or relationship-expressive I-levels (Palmer, 1965), leading psychotherapists Traux and Carkhuff to conclude that "To generalize, it would seem likely that the matching of patient and therapist types plays a critical role in cases where the therapist is quite restricted in his ability to show understanding, warmth or genuineness to all but a narrow range of human beings" (1976, pp. 166-167).

Domain F: Programme Characteristics

	Available points	# of points	N/A	% score	CR
Women's project	25	06	0	24%	4.6
JPACS	25	20	1	83%	4

Domain F: Programme characteristics investigates whether services adhere to the principles that:

- services should be intensive and behavioural in nature;
- behavioural programmes should target the criminogenic needs of high-risk offenders;
- programme contingencies and behavioural strategies should be enforced in a firm but fair manner;
- programme structure and activities should be designed to disrupt the delinquency network by placing offenders in situations (people and places) where prosocial activities predominate;
- relapse prevention strategies should be provided in the community to the extent possible (adapted from Gendreau, 1996a).

In this domain, the women's project was categorised as Unsatisfactory, whilst JPACS achieved a Very Satisfactory categorisation. One item (F7) was voided as not applicable in the evaluation of JPACS. This related to a series of four items (F4-F7) concerned with the treatment strategies employed by the evaluated service. Where any of the appropriate treatment strategies (behavioural, social-learning or cognitive-behavioural) identified in items F4-F6 are employed by the service, F7 is not applicable. F7 is only applicable where the criminogenic needs of service-users are targeted by a structured intervention not listed in items F4-F6 (as in the evaluation of the women's project).

An additional consideration in the evaluation of the women's project was that the expectations of the CPAI in this domain relate specifically to interventions for higher-risk offenders and so it was scored based on the characteristics of services within the project that aim to reduce the recidivism of statutory participants. The report noted that Domain F "is not applicable to diversion services, since exposing women who present a low risk of re-offending to the probation-type interventions itemised in this section of the evaluation is ineffective practice and it is not possible to delineate voluntary

participants according to their risk level due to the absence of assessment and categorisation in this respect” (Kerr, 2014, p. 21).

The results for the women’s project in this domain mirrored the findings of Latessa and Holsinger, who, whilst acknowledging that a strength of many services is the targeting of criminogenic needs, reported that “offenders typically have not been spending a significant percentage of their time in structured programs [and] treatment provided has not been varying by risk and need levels” (1998, p. 26). Moreover, they describe a disregard for the evidence base in the design of interventions as a common flaw in CPAI-evaluated services, resulting in “a lack of a consistently applied treatment model in place. In general, major shortcomings found...include lack of programmatic structure; incomplete or nonexistent treatment manuals; few rewards to encourage program participation and compliance; ineffective use of punishment; staff being allowed to design their own interventions regardless of the treatment base; and a host of very obvious and definable, yet ineffective treatment models” (Latessa & Holsinger, 1998, p 26).

The evaluation of the women’s project found that it primarily targeted criminogenic needs such as addressing pro-criminal attitudes, reducing negative peer associations, enhancing constructive use of leisure time, reducing substance misuse, promoting family affection, communication and problem solving and increasing Education, Training and Employment (ETE) related performance. Positively, non-criminogenic targets, such as decreasing anxiety and increasing self-esteem, were given less attention, though were accounted for as important responsivity issues, specific to each case. The project’s partnership model was particularly effective in enabling staff from all services to monitor closely the whereabouts and peer associations of service-users and to share information in this regard. Evidence suggests that these behaviours are especially indicative of risk of re-offending, so close monitoring enabled effective risk assessment and management throughout service-users’ engagement with the project. Additionally, the evaluation report noted that the project’s “strong participatory ethos” was reflected in practitioners’ commitment to a “mutual and collaborative approach” (Kerr, 2014, p. 22). Nevertheless, the major weaknesses common to services assessed by Latessa and

Holsinger were readily apparent in the evaluation of the women's project. For example, the percentage of service-users' time spent engaged in prosocial activities was not monitored and the report noted that:

...services do not routinely employ behavioural, social learning or cognitive behavioural strategies as treatment modalities, instead employing techniques based on restorative practices, focused discussion, outreach, mentoring, practical activities and...corporate development strategies. Whilst NPS and CRC practitioners did mention some techniques associated with effective practice, no manual is routinely used for the statutory supervision of [service-users], and the emphasis is instead on client-centered supervision. The observed exception to this was the accredited Enhanced Thinking Skills (ETS) programme, which employs effective behavioural strategies through manualised groupwork. However, the two participants of ETS at the time of observation both presented a low risk of re-offending according to their Offender Group Reconviction Scale (OGRS) score. At present, the level of interventions delivered across the [women's project] is not determined according to the level of risk of reoffending and this limits the effectiveness of services. (Kerr, 2014, pp. 22-23).

By its nature, the women's project was a magpie organisation – drawing on local resources and initiatives according to perceived need or advantage. The provision of an holistic approach is the bedrock of women's services and it is important to note that the project's acknowledgement of the gender-responsive literature was not recognised by the CPAI evaluation of programme characteristics, since the elements it embraced (women-only space, trauma-informed work, parenting interventions, etc) were not sufficiently linked to the gender-neutral effective practice that is the basis for the CPAI items (eg, they didn't employ learning behavioural treatment modalities).

Notwithstanding arguments for greater consideration of gender-responsive factors in evaluation (which will be looked at in greater detail in Chapters Five and Six), deficits in the organisational context of the project were a key underlying issue. Without coherent leadership or direction, there was no clear overview of the different treatment strategies used by individuals or agencies across the services that comprised the women's project, nor clear referral processes for these, leaving the project vulnerable to ineffective practice and duplication. The report noted that "there are a number of services and group work programmes available across the [project] that aim to reduce reoffending, and to which statutory cases can be referred (and in some cases required to complete) as part of their Order. However, there are no discernible differences in criteria for inclusion in [these]. As a result, it appears relatively arbitrary which services

women are referred to” (Kerr, 2014, p. 22) As the women’s Project Manager pointed out, these deficits were inextricably linked to the organisational context of the project:

There...is duplication, there are gaps, we [are] working towards being able to provide a more comprehensive...provision of different interventions for different needs and levels of risk [but] we can’t [get] to that place right now...because there’s so many things that impact on that in terms of resources...the project doesn’t control all of that. We’re reliant on what individual agencies deliver, ...what funds deliver, ...what commissioners want, ...what [the] skill bases of staff [are], you know, that even if there’s [effective interventions] out there we haven’t necessarily got staff trained up today to be able to deliver them. (WPM).

There are also clear links between JPACS’s organisational context and its achievements in this domain. The service adopts social-learning and cognitive-behavioural treatment strategies through the delivery of evidence-based manualised programmes and the use of external consultants. Also in line with the evidence regarding effective practice, JPACS predominantly targets higher-risk offenders, with approximately 80% of service-users being assessed as presenting a moderate to high risk of re-offending at the time of the evaluation. Programmes with the highest intensity and duration are reserved for higher-risk offenders, for example, the Self Management and Rational Thinking (SMART) programme targeting high-risk offenders runs for thirty-five sessions, whilst the Offending Is Not The Only Choice (OINTOC) programme for medium-risk offenders comprises twenty. The development of these services has been part of the same conscious organisational strategy for the enhancement of probation practice that included spearheading use of the LSI-R in the British Isles and partnering with academics and consultants (including the Cognitive Centre Foundation, which provides both SMART and OINTOC) to base services on strong theoretical models (Heath et al, 2002; Miles et al, 2009).

A comparison of the women’s project and JPACS’s responses to the CPAI recommendation that all service-users should spend at least 40% of their time (or 50 hours per week) in pro-social tasks exemplifies the greater possibilities for well-resourced, coherently-led organisations to mobilise to implement effective practice:

One of the things that stood out for me in [the report] is [the comment] about women spending at least 40% of their time or 50 hours per week in pro-social tasks ... I don’t really see how that would work, we haven’t got the manpower, we haven’t got the...you know, to sort of set that up at this point but maybe that will change. (WPO3).

I chair our risk management meetings ... with other agencies and quite frequently [after receiving the CPAI report] I'm now saying "ok, so how are we filling up this guy's 50 hours?" ... and that's actually been a really good one because people have actually turned around and sort of thought "yeah, you know, actually we could be doing more" and it's actually, we've been able to sort of say to, to adult social services or to other agencies "look, we really need to get this person doing extra time" and people seem to understand it so... That's just a sort of ... a real simple sort of thing which ... came out of [the evaluation]. I know it wasn't a major one but it was just, it was just a nice one just to show you just some practical, some, er, some practical, er, thing about it. (JTL).

Domain G: Core Correctional Practice

	Available points	# of points	N/A	% score	CR
Women's project	45	18	0	40%	4.4
JPACS	45	35	0	78%	3.5

The core correctional practices (CCPs) assessed in domain G of the CPAI (see Table 4.1) have evolved as a result of ongoing empirical evaluation since their inception in the 1980s and have been validated in studies correlating the scoring of over 700 programmes with service-user recidivism. (Latessa et al, 2013; Lowenkamp, 2004; Lowenkamp, Latessa, & Smith, 2006; Matthews, Hubbard, & Latessa, 2001).

Based on observations of supervision and programme sessions, the women's project was categorised as Unsatisfactory in this domain, and JPACS as Very Satisfactory. The reduced CR for JPACS was due to some disparity in delivery, with some practitioners utilising more components of effective treatment interventions than others, whilst practitioner approach and level of therapeutic skills observed in the women's project was generally of a similar standard. Again, the scoring for the women's project is in keeping with the international literature. In their 2004 meta-analysis, Dowden and Andrews found that programmes incorporating elements of CCPs were more effective at reducing recidivism but that CCPs were used infrequently – even the most commonly used CCPs were found in only 16% of the interventions studied (though it is important to note that this may have been the result of services failing to describe their staff and treatment procedures in detail rather than an actual lack of skills within the services included in the meta-analysis).

Table 4.1 Core Correctional Practices (CCPs)

Appropriate use of authority	<ul style="list-style-type: none"> • Staff focus their instruction and feedback on the service-user's behaviour, without being personal • Staff are direct and specific • Staff do not raise their voices • Staff specify choices and attendant consequences • Staff are encouraging and respectful • Staff support words with action • "A firm but fair approach"
Appropriate modelling and reinforcement	<ul style="list-style-type: none"> • Staff use a coping model – giving a concrete and vivid demonstration of desired behaviour • Staff define and model the skill before engaging the service-user in progressively difficult practice (ie role-play) • Staff are a source of reinforcement rather than punishment • Reinforcement involves giving or encouraging immediate and detailed description of the benefits of the desired behavior • Staff immediately give or encourage detailed description of the undesirability of anti-social behaviours
Skill-building and problem-solving strategies	<p>Staff members help service-users to:</p> <ul style="list-style-type: none"> • Identify the problem • Implement a plan • Clarify goals • Evaluate options • Generate alternatives • Evaluate the plan
Relationship factors	<p>Staff are:</p> <ul style="list-style-type: none"> • Warm • Genuine • Humorous • Enthusiastic • Confident • Empathic • Respectful • Flexible • Committed to helping the service-user • Engaging <p>Communication with service-users is:</p> <ul style="list-style-type: none"> • Directive • Solution-focused • Structured • Non-blaming • Contingency-based
Motivational interviewing	<p>Staff:</p> <ul style="list-style-type: none"> • Develop discrepancy • Express empathy • Amplify ambivalence • Roll with resistance • Support self-efficacy (Miller & Rollnick, 2013)
Cognitive restructuring	<p>Staff:</p> <ul style="list-style-type: none"> • Encourage descriptions of problematic situations (eg, thinking report) • Draw out descriptions of thoughts and feelings • Identify thinking deficits • Guide practice of alternative thinking • Put in place opportunities to practice less risky thinking (Spiegler & Guevremont, 2013; Gendreau et al, 2010)

Items 1-4 are adapted from Andrews & Kiessling (1980) and Dowden & Andrews (2004). The original items also included the use of community resources as a CCP, though Gendreau et al redefined this as an individual dimension of effective practice, 'inter-agency communication' (assessed in domain I of the CPAI), and added the elements of motivational interviewing and cognitive restructuring as CCPs in their 2004 summary of key elements.

Breaking down the scoring further can be helpful in exploring where deficits lie and the implications of this for treatment effectiveness. In the Jersey Supervision Skills Study (JS3), staff practice was observed and scored against a checklist of CCPs. Raynor et al (2014) found that, on average, staff scored more highly in skills designated as “relationship or responsive” skills than in “structuring” skills intended to prepare for or promote change. Additionally, staff varied considerably in the skills they were observed to use and higher-scoring staff used a wide range of skills with their service-users, with high scores in structuring skills as well as relationship skills, while the lower-scoring staff tended to score particularly low on structuring skills. In fact, most staff scored quite highly on relationship skills and there were no staff who combined a high level of structuring skills with a low level of relationship skills, perhaps indicating that “good relationships may be a precondition for the most effective use of structuring skills rather than a strong source of change in themselves, at least in the short term” (Raynor et al, 2014, p. 244). Finding that higher-scoring officers were more effective in reducing re-offending, Raynor et al also point out that it is likely that a combination of skills will produce the best results, as indicated by Paparozzi and Gendreau’s 2005 study, which found that practitioners taking a “balanced approach” in their supervision of parolees were more effective than those with either strong “social-casework” or “law-enforcement” orientations.

Borrowing from the JS3 methodology, the results from domain G of the CPAI evaluations of the women’s project (WP) and JPACS can be broken down into relationship and structuring skills as below:

Table 4.2 Domain G Scoring of Relationship and Structuring Skills

Skills	#Items	WP	JPACS
Relationship or responsive skills			
Effective use of authority	10	10	10
Relationship practices	7	4	7
Total/percentage score	17/17 100%	14/17 82%	17/17 100%
Structuring skills			
Anti-criminal modelling	4	2	3
Effective reinforcement	4	1	4
Effective disapproval	4	1	2
Problem-solving techniques	6	0	4
Procedures for skill building	5	0	0
Elements of cognitive restructuring	5	0	5
Total/percentage score	28/28 100%	4/28 14%	18/28 64%

It can be seen that, in both JPACS and the women's project, practitioners demonstrated a high level of relationship skills. It should be noted here that elements of Motivational Interviewing (MI) are incorporated in the CPAI as relationship practices but accounted for as structuring skills in the JS3 study. Had these been scored as structuring skills in the CPAI evaluations, the women's project would have achieved an even higher score of 14/15 or 93% for the relationship skills of practitioners. The authors of the JS3 suggest that the high scores in relationship skills in their study may be related to most of the probation officer sample qualifying as social workers, with relationship skills tending to "feature strongly in social work training, whereas several of the structuring skills are more likely to be encountered, if at all, in post-qualifying or in-service training" (Raynor et al, 2014, p. 240). However, there was only one social-work trained practitioner in the sample for this study's CPAI evaluation of the women's project, indicating that relationship skills may not be solely the result of social-work training.

As discussed, the development of positive relationships with service-users was a predominant theme in interviews at both the women's project and JPACS, though the exact mechanisms by which practitioners develop relationship skills is unclear. For example, it may be that more innately socially-oriented people are attracted to probation as a profession; that relationship skills feature similarly in the various different routes to qualification and/or employment in probation – eg, higher-level education in social-work, criminology, social care, criminal justice, etc; that hiring processes favour candidates with relationship skills; or that practitioners develop relationship skills in-post through their interactions with service-users and/or according to the expectations of the organisation. It seems unlikely that the latter is the only factor since there was no obvious disparity between the relationship skills of more experienced practitioners and those with less experience. Of course, it may also be that a high level of relationship skills is a dominant feature of the women's project because all the practitioners are female, or because people with innate relationship skills are more likely to be drawn to working with women – as discussed in Chapter One (p. 33), relationship building has been identified as especially important in work with women, and the JS3 study found that female practitioners achieved higher CCP checklist scores on average than their male counterparts (Raynor & Vanstone, 2016, p. 1142). It is beyond the scope of the current study to reach any conclusions in this respect since there were not enough male practitioners included in the sample to draw any meaningful comparisons, though it should be noted that the importance of positive relationships was not only considered significant by the women's project service-users, but by male service-users in JPACS who identified their workers predominantly as “someone to talk to” (JSU3, JSU5, JSU1), “like a counsellor” (JSU6) or “my firm rock” (JSU4).

Clearly, fewer structuring skills were observed in practice at both agencies, though practitioners at JPACS were observed to use a wider range of skills than at the women's project. As identified by CPAI evaluation, the employment of incomplete and ineffective strategies in the women's project is likely to have resulted partly from the lack of standardised use of an evidence-based manual. The thinking skills groupwork programme delivered by the project was a notable exception. The observed session drew on an evidence-based, accredited manual and the employed treatment strategies

were in line with effective core correctional practice. As discussed, there were also deficits in training and clinical supervision in the women's project (and thus a lack of opportunity to review and improve CCPs) and it is also likely that workload pressures (resulting from understaffing, increased caseloads and other organisational issues) impeded the delivery of structuring work in some cases. Three months into a twelve-month Community Order, one service-user pointed out:

I was meant to be, well, when I was sentenced at court, doing an alcohol [programme] with [the project] but I haven't done it, I haven't done one session...it hasn't been [mentioned]...it's in my sentence plan...I'm meant to be doing it on a one-to-one basis...but they haven't done any work with me on that. (WSU1).

JPACS's greater use of structuring skills in sessions was clearly linked to staff training, clinical supervision and the use of evidence-based manuals for effective practice. Practitioners positively reinforced pro-social behaviour, explaining why they liked the behaviour exhibited by the service-user and encouraging service-users to consider the benefits of continued use of this behaviour. They also engaged them in some problem-solving techniques including focusing on the antecedents, behaviour and consequences of their offending, clarifying goals that could resolve or prevent future negative consequences and generating alternative solutions. The evaluator also observed some practitioners using elements of cognitive restructuring by developing descriptions of problematic situations, producing descriptions of thoughts and feelings, generating discussions and identifying risky thinking, establishing alternatives and setting up additional practice opportunities.

However, whilst the skill-building procedures of modelling, role-play and practice were referred to in programme manuals and interviews with practitioners and service-users, the evaluator observed no practitioner using skill-building techniques and it does appear that this element of effective practice may be somewhat underused. Service-users recalling their involvement in role-play did not understand it as a skill-building technique but as a way of illustrating or 'playing-out' offending behaviour to increase understanding:

Evaluator:	Do you practice new skills in those groups?
JSU6:	No, they just tell you about it...

- Evaluator: Now I just heard...it said in your report that you were really good at role-play... is that a kind of practice?
- JSU6: No, it's like a situation. It's like...so, like the girl didn't cook the tea on time. I was the fella and someone else was the girl so I walk in "what the fuck are you doing? Why isn't my tea cooked? Blah blah blah..." That was the role-playing thing...How could I deal with that better?
- Evaluator: And then, do you get a chance to...do they show you how to deal with it better or anything?
- JSU6: No...you just have to, like, brainstorm it and all that but they don't actually show you, they just write it down on the board.
- Evaluator: Right...do you get a chance to try it again, doing it a different way?
- JSU6: No...

It is possible that practitioners are under-confident in using these techniques. Commentators on the use of drama in English Language Teaching have found that facilitation of role-play in particular may evoke a sense of dread in teachers who feel inexpert, and may fear losing control of participants or looking foolish (Ebong, 2004; Royka, 2000), whilst Hewish suggested that there is an unwarranted fear of the "opening up" of an offender through drama unleashing "some uncontrollable demon", which he contended is used as an excuse not to engage offenders in emotional work (Hewish, 1999). Whether the under-use of skill-building techniques in probation services is the result of practitioner fear, inexperience, excuse or any other factor, it seems likely that specific training would help to address this deficit.

Domain H: Inter-Agency Communication

	Available points	# of points	N/A	% score	CR
Women's project	05	05	0	100%	4
JPACS	05	05	0	100%	4.8

The domain on inter-agency communication is concerned with the "referral of offenders to community-based services that provide quality services applicable to offenders and their problems" (Gendreau, 1996a, p. 125) and the quality of advocacy and co-ordination involved with multi-agency practice. Harris and Smith (1996) observe that "to the

extent that the program interacts with other agencies around the same clients, interorganizational adaptations need to take place that include new information systems, the development of new procedures, and mechanisms for resolving conflict” (p. 200). In this domain, both the women’s project and JPACS achieved 100%.

The women’s project was founded on the principle of holistic work and the CPAI report noted that:

The [women’s project] has the potential to be an exemplar of a multi-agency service. Cases are managed co-operatively and referrals discussed in weekly multi-agency case meetings, ensuring that services are delivered across agencies according to the presenting needs of women. Additionally, ... staff regularly have contact with relevant agencies outside of the [project] (eg, schools, hospitals, housing providers, etc) and advocate for service users where appropriate. (Kerr, 2014, p. 27).

Plans to develop monthly multi-agency training events and a shared database to record contacts with service-users across the different agencies of the project were also highlighted by the report as having the potential to increase the effective co-ordination of treatment and, though it appears that organisational issues prevented full implementation of these plans following the evaluation, the project would still have been considered to be working to a Very Satisfactory level in this domain.

Multi-agency working is a crucial feature of women’s projects. The value of the ‘one-stop-shop’ model of community provision for women advocated by Corston (2007), whereby women’s centres provide female-only environments, support and a range of programmes and interventions tailored to their individual needs, has been reiterated by subsequent HMIP inspections (HM Inspectorate of Probation, 2016b) and the development of local whole-system approaches is a key strategic objective for the government Advisory Board for Female Offenders (Ministry of Justice, 2013b; HM Inspectorate of Probation, 2016b). Diversion schemes such as that provided by the women’s project have been singled out as particularly promising examples of innovative practice in inspections (HM Inspectorate of Probation, 2016b). Of course, as a gender-neutral evaluation instrument specific to practice with *offenders*, the CPAI does not take account of diversion services, nor does it specify the need for gender-responsive practice in this domain – eg, the need for community resources to respond to

substance misuse, mental health (including experience of trauma), emotional management, family contact and social capital as priority needs (NOMS, 2015) and so it is, arguably, inadequate in its evaluation of inter-agency communication in a women's project (this will be explored further in Chapters Five and Six).

Inter-agency practice in the gender-neutral service of JPACS was more comprehensively captured by CPAI evaluation, which found that:

JPACS works in a multi-agency context, including involvement in JMAPPA and RAMAS processes for the multi-agency management of risk and Child Protection processes alongside Social Services. Additionally, JPACS is formally linked to a number of agencies that offer services relevant to the needs of offenders, some of which (eg, victim support, mental health, substance misuse, etc) are represented amongst the staff team at JPACS. Staff routinely advocate with outside agencies on behalf of their clients, and the evaluator observed several examples of this as well as evidence of effective inter-agency communication, including telephone conversations, email correspondence and meetings with staff from other organisations. Leaflets and posters in the waiting room and practitioners' offices provide clients with links to other organisations and the community, and referral documents in JPACS's policy book are easily accessible to staff members. Staff also have the opportunity to engage in multi-agency training across the various organisations (eg, Child Protection training delivered by Social Services), which enables effective co-ordination between JPACS and other agencies in regards to shared clients. (Kerr, 2016, pp. 24-25).

Evidence regarding the effectiveness of advocacy and brokerage in terms of reduced reoffending is mixed (Palmer, 1996, pp. 136-137), it being one of just two CCPs not significantly correlated with reduced recidivism in Dowden and Andrews's 2004 meta-analysis. However, the most likely factor in determining its success is the *quality* of services to which offenders are referred. As Gendreau points out "a high level of advocacy and brokerage should be attempted *as long as* community agencies offer appropriate services" (1996a, p. 125; [My italics]). He goes on to advocate use of the CPAI since "it is vital that community services be assessed in this light in as objective a manner as possible" (1996a, p, 125).

Domain I: Evaluation

	Available points	# of points	N/A	% score	CR
Women's project	08	06	0	75%	3.5
JPACS	08	08	0	100%	5

Latessa and Holsinger observed that “programs that study themselves tend to be more effective than programs that do not” since “data provide insight into program and offender performance, help identify who is successful and who is not, and allow adjustments to be made” (1998, p. 27). In their 1998 summary of findings from CPAI evaluations in the US, they found that most services did not have adequate evaluation and quality control processes in place. Whilst services usually conducted file reviews and case audits, most failed to measure programme or offender performance over time, did not track service-users after programme completion and nor did they undertake formal evaluation involving comparison groups (Latessa & Holsinger, 1998, p. 27). Evaluation also appears to be a weakness in services across Europe. Miles et al note that there are hundreds of probation services across the world, with Van Kalmthout and Durnescu’s 2008 survey of probation work in Europe alone covering 32 countries, but that very few services “are able to document the outcomes of their work or to specify what difference they make to offenders” (Miles et al, 2009, p. 2). A recent HMIP inspection of community services for women in England and Wales found that:

There was a lack of focus on outcomes for women, both strategically and operationally. Some responsible officers monitored the progress of individuals but very few received sufficient information about outcomes. There was little evidence of any follow-up to the progress of women beyond completion of their statutory orders or licences. This was compounded by the fact that once a case managed by the CRC was closed, the responsible officer could no longer gain access to the electronic record to track any progress made.

Overall, ... The monitoring and evaluation of the progress of women, both during and beyond their statutory orders or licences, needed much more rigour, in order to know what was effective in working with those women. This was a concern, given the recommendation from our 2011 inspection that Probation Trusts should: ‘continue to maintain an additional focus on women in order to embed strategic developments in operational delivery through the development of effective outcome measures supported by monitoring, evaluation and managerial oversight’. (HM Inspectorate of Probation, 2016b, p. 21).

It is therefore an indication of exceptional practice in both JPACS and the women’s project that they each achieved Very Satisfactory categorisations in this domain of the CPAI.

Miles et al identified Jersey as “one of the very few” probation services that can document effectiveness, attributing this to the “conscientiousness of its staff and

managers and the quality of data that they provide” (2009, p. 2), indicating that, once again, the organisational context of the service is crucial in its success:

There is a clear strategic focus on supporting research which has implications for practice and then applying research findings to improve effectiveness. (Miles et al, 2015, p. 23).

The CPAI report outlined an approach to evaluation and quality control that, as Miles et al point out, “has attracted international attention” (2009, p. 2):

JPACS takes a conscientious and rigorous approach to evaluation, involving: quarterly checks reviewing files and monitoring treatment progress; video-tapes of programme sessions being sent to the Cognitive Centre Foundation to ensure adherence to the manual; use of the Jersey Supervision Interview Checklist to evaluate the integrity of probation supervision sessions; the distribution of client feedback questionnaires, with results published annually; the standardised re-assessment of clients on target behaviours; and the gathering of follow-up reconviction data.

The fourth of a series of reports evaluating the effectiveness of JPACS’s community supervision services was published in November 2015 and the Service has also been the subject of a number of peer-reviewed journal articles regarding the Jersey Supervision Skills Study (JS3). (Kerr, 2016, pp. 25-26).

The comparatively low confidence rating in this domain for the women’s project is due to its status as a pilot at the time of the CPAI evaluation. The report cautioned that:

... the low confidence score reflects the lack of evidence available for practices that could only be undertaken once the [project] is established such as, for example, publication of an article detailing its effectiveness. This section has been scored according to the processes being put in place for future practice. (Kerr, 2014, p. 28).

Several processes were being put in place at the time of the CPAI assessment to enable effective evaluation of the women’s project. The report noted that:

The [women’s project] will be subject to ongoing evaluation of both statutory and diversion services. Programme checks will monitor treatment progress, women will be surveyed as to their satisfaction with the service, assessment scoring regarding target behaviours will be analysed, and (re-)arrest, (re-)conviction and/or (re-)incarceration data will be gathered for a substantial amount of time following programme completion and compared with a risk-control comparison group for analysis. This dedicated approach to evaluation will enable the production of a formally-written document detailing the effectiveness of the programme that can be used to inform commissioners, policy makers and service providers. (Kerr, 2014, pp. 28-29).

It is important to add, however, that Phase Two of this study revealed that some elements of these plans had not been implemented. Dedicated research staff in place at the time of the CPAI assessment had been redeployed elsewhere and some evaluation

processes had apparently been abandoned without explanation or were suffering from a lack of coherent leadership. As the manager of a partner agency recounted:

There [seemed] to be quite a strong focus on evaluation when the pilot ... first started ... so [the project researcher] was at ... every meeting, she spent time here visiting women ... and then all of a sudden ... she was out of the picture.

...The other thing that I felt was missed out is... even though there was ... there is a co-ordinator, there seems to be a lack of leadership in the sense of properly monitoring the data. So, for example, at one point, ... I've looked at all the prison releases, I've cross-referenced the lists as [a worker] would send them to me from prison and then looked at how those were allocated from the minutes ... just so that we can say... you know, since the pilot's started, this many women were released, this many were allocated, these are the agencies who picked up the cases. And I did that on my own ... off my own back because it was just driving me crazy ... and I was thinking there should have been... somebody ... collecting and monitoring that data and also, not just to that point that I did it, but also even further to look at [whether] those women who were allocated to [us] from the [project] meeting: Were they supported? Where are they now? There's ... no review element embedded in the process ... so cases are allocated but they're not reviewed so you don't know how successful ... [the project] is and we potentially – I'm not saying that this is happening – but potentially cases could be allocated and nobody would be working with them. How would we know about it? How would we know that ... it's effective? ... Who's looking if we are reducing reoffending? I know that a lot of focus has gone on the diversion scheme and I think the data there is monitored quite well but I think, again, that's down to [another partner agency] and the work that they've been doing, so.... (WPA1M).

Additionally, external researchers described difficulties in accessing data from the women's project, partially because of a low response rate from project staff and partially because data relating to measures of interest (eg, whether service-users had experienced domestic abuse or were diagnosed as having mental health issues, etc) were not recorded in a standardised way.

Taking all the above factors into account, it is possible that a later CPAI evaluation of the women's project may have produced a lower score in this domain.

Conclusion

Clearly, evaluations of two services are not sufficient to draw conclusions regarding the effectiveness of interventions across the British Isles and, indeed, that is not the aim of this study, though we can reach two general conclusions from its first phase. Firstly, the CPAI-2010 appears broadly relevant in the context of services in the British Isles. Factors identified as strengths or weaknesses by Latessa & Holsinger in their US

evaluations were evident in the evaluations of both the women's project and JPACS. JPACS's exceptionally high score is supported by findings of a commitment to evidence-based effective practice that is exceptional to the British Isles, as well as internationally. In line with evidence from the US that high CPAI scores correlate with reduced recidivism amongst service-users, reconviction rates for community sentences in Jersey are lower than in England and Wales and, in over two thirds of probation cases managed by JPACS, the risk of re-offending, as measured by the LSI-R, is reduced during supervision (Miles et al, 2015).

Whilst recidivism scores for the women's project are not available to facilitate meaningful comparison, it is certainly the case that the women's project's Unsatisfactory score is entirely in keeping with the international literature regarding CPAI evaluations. Summarising the findings from 282 CPAI assessments across the US and Canada (Gendreau & Goggin, 1991; Hoge et al, 1993; Latessa & Holsinger, 1998, 1999), Gendreau et al write: "the blunt truth is that 70 per cent of all programs ... 'failed' according to the CPAI" (2001, p. 260).

This study has found that JPACS's exceptional achievements relate to its exceptional organisational context. Unlike its counterparts in England and Wales, the service is not an agency of a government department (instead being accountable to the judiciary) and, as Raynor and Ugwudike observe, this has protected it from undue political pressures (2013, p. 50). In the absence of the top-down, target-driven processes that have become a feature of probation services in the UK under government departments including NOMS (see, for example, Cowe et al, 2007), JPACS has collaborated with academics and other experts to develop its services, earning it "a global reputation as a pioneer of evidence-based practice" (Miles et al, 2015, p. 24). Conversely, organisational deficits such as the division of probation services and turbulence in the third sector under TR have been seen to have had a significantly detrimental impact on a number of features of the women's project, limiting its potential to practice effectively.

This leads to this chapter's second conclusion: that organisational context (as assessed in Domains B, C and D of the CPAI-2010) is key across almost every other aspect of

service. For example, the use of structuring skills in sessions (assessed in Domain G) was clearly linked to the provision or absence of staff training, clinical supervision and evidence-based manuals; effective programme characteristics were linked to effective leadership, ideological commitment and strategic direction; and strong risk/need assessment processes required efficient resource management. The clear and overarching impact of organisational context and management issues in this study emphasises the inadequacy of evaluation and inspection approaches that take no or little account of these. It can be argued that an evaluation approach which looks solely at service-level practice is likely to be of little use in helping an organisation to identify and address the core issues.

Whilst these conclusions speak to the CPAI's *relevance* in the contexts of the British Isles, the findings also indicate that the CPAI is *not sufficient* to capture the breadth of practice in, for example, desistance-informed and gender-responsive contexts. The inability of the CPAI to address the diversion practices of the women's project was picked up in Domain F, and the narrowness of its focus on advocacy and brokerage was noted in Domain H as being insufficient to capture the full picture of work to engage women holistically. As discussed in Chapter Two, the RNR evidence base is predominantly concerned with work towards the psychological or personal rehabilitation of service-users, the area for which there exists comprehensive empirical support. However, this means that the CPAI overlooks desistance-focused efforts to engage with service-users' legal, moral and social rehabilitation. As such, incorporating only the quantitatively empirically-supported principles of RNR may do a disservice to agencies that refer to a wider evidence base to meet the needs of their users (pp. 67-69).

Beyond these general conclusions, the findings from CPAI-2010 evaluations of the women's project and JPACS indicate some areas for further work beyond the scope of this study. There appears to be a need for specific training in the use of skill building techniques (and particularly role-play) to increase practitioners' use of this CCP. The availability and desirability of responsivity assessment for the matching of practitioners to service-users would also bear further investigation.

This chapter has looked at the findings from Phase One of the study, ie, the CPAI evaluations of the sample programmes. Chapter Five will discuss findings from Phase Two of the study, which is concerned with the qualitative experience of evaluation and will pick up on some of the points raised here regarding factors relating to gender-responsive practice.

CHAPTER FIVE

Findings from Phase Two: Themes from Semi-Structured Interviews

Introduction

This chapter explores the seven themes identified through inductive analysis of transcripts of semi-structured interviews with evaluation participants from the women's project in Wales and Jersey Probation and After Care Service (JPACS), excluding service-users, who were unavailable in each service in Phase Two of the study. Though inter-related, these seven categories (whether *a priori*, eg, cultural and ideological issues regarding language, or emergent) appear to be distinct and the most important, given the research objectives identified in Chapter Two. Similarities and differences across sub-groups of participants (eg, practitioners/management, National Probation Service (NPS)/Community Rehabilitation Company (CRC), gender-responsive service/gender-neutral service, etc) and relationships between categories will also be explored in order to reach generalisable conclusions about the experience and usefulness of Correctional Program Assessment Inventory (CPAI) evaluation.

Reflection, Not Inspection: Holistic, Independent, Inclusive Evaluation

A significant theme in interview was the value of a participatory approach to evaluation, which compared favourably to experiences of inspection. The purpose of evaluation was seen as enabling a reflective service:

The important thing is that it's made us reflect on ourselves as a service, our practice and where we're going. I think that's been the most important thing. (JTL).

If you don't reflect on something then you can't learn from it...if it's not evaluated, then how can you tell that it's working? (WPO1).

To a great extent, this was facilitated in three ways:

- The evaluator's independence from the services was seen as providing an unbiased perspective ("it's having an independent pair of eyes coming in and looking" (JPO3); "it's that objective view" (WPO1); "that impartial look" (WPM)), and a mutual space for evaluation, safe from coercion or imposition:

I suppose we haven't been influenced by above, so I think what you've collated and put together is probably a true reflection. I wouldn't have liked it to be enforced too much from above. (WPO2).

I remember we spent quite a bit of time [on the CPAI interview] and that's because I felt really comfortable and...able to be honest about things. And I think what's really useful with having [an external evaluator] in that role is that [they're] not going to be defensive and trying to justify why, if things have gone wrong, why that's happened. [They're] independent, all [they're] looking for is the comments. (WPA1M).

- The evaluator's inter-personal style was identified as needing to be "approachable, accommodating...warm..." (JPO1) and "very clear" (JPO3) in order to reduce anxiety around evaluation ("oh my god, what questions is she going to ask...? Am I going to be able to answer them...? Am I going to be letting the organisation down...?" (JPO3)), make participants "feel at ease" (JPO3), "comfortable" (WPO4) and "valued" (WPA1M), and facilitate an honest and engaging process:

I never thought, "oh, I felt uneasy" or "I felt awkward". I think it's important to be open and honest and I think I was. (JPO3).

Your personal style was important. You were available, able to listen and gave staff confidence. Sometimes you meet professionals and academics who are keen to tell you what you should be doing. That's a turn-off in terms of relationship-building and taking on ideas. You were thoughtful, respectful and insightful. (JACPO).

Additionally, pro-activity and flexibility were highlighted as particularly important qualities by commissioners:

With these kind of evaluations of this type of service there is a need for a lot of proactivity and flexibility in my experience, so I've seen...evaluations previously where they've needed to interview x amount of people and only managed a tiny fraction of that because...they...plan for everyone to be there on the day and they weren't there on the day and that was the end of that.... [The CPAI evaluation process] worked really well because...if [participants] missed one opportunity, they had another opportunity.... (WPM).

The whole [evaluation] process was, I thought, quite inclusive and empowering for participants...I think that's one of the differences between the CPAI approach and the more traditional style of inspection and audit...I have to say it was partly because of the way [it was done]. I could see that you could actually use the CPAI in quite a...mechanistic and unhelpful way. But the way it was done was actually very helpful, people felt engaged with it. (JCPO).

- An holistic approach, embracing all stakeholders, with a particular emphasis on the contributions of service-users:

It was nice to have your chance to sort of speak about how you felt things were going and I think it was good for the women who have come through the [project] to be involved as well...it gave a bit of the whole picture, I suppose. It wasn't just interviews it was observations and things... You've got the women's perspective, you've got the workers' perspective, you sat in on bits and pieces – I think you're getting the 'whole'. (WPO1).

[Compared to inspections] it's a lot more thorough, actually. I think that you've taken a lot more time to discuss things with us, as practitioners, and our clients than other processes have. I think there's been more focus on individual feedback...as opposed to a lot more sort of case readings and looking at things from that perspective, just purely about what's documented as opposed to actually, you know, the verbal feedback...I think [that's] a strength 'cause I think it's always nice to...share your views on something because we all have views about how we work and operate...and where those things perhaps need improving, and it's nice to be asked those questions. (JPO1).

What do I think is most important when evaluating intervention programmes? I suppose to get everybody's perspective really. To make sure that everybody feels, you know, an active participant and that their voice matters, especially with service-users...'cause ultimately we're making decisions that may or may not benefit [them]...A lot of the feedback I've had [from service-users] was "oh great" – you know, they were happy that...their voice and...their opinion was being heard. (WPO2).

Taken together, these three factors were identified as enabling honest responses and greater engagement with the evaluation process, including and empowering participants rather than imposing upon and judging them. This participatory approach was considered essential in producing a culturally-sensitive and accurate reflection of the services, which would actively engage management and practitioners in implementing recommendations:

I recognised our service in the report...The (CPAI) methodology was more helpful than, for example, [an earlier] risk audit [which] didn't involve anything other than looking at paper processes...so there's no involvement from staff, it's done *to* them and whatever comes out of that, the staff aren't going to be engaged with to the same degree as with [the CPAI]. It's

just done to you isn't it and then, you know: "Under financial direction such and such, you're going to have to do x"...or whatever, and people are going to grumble about doing that if they don't understand the rationale behind it and I think that [the CPAI] process involved conversations with people so they'll...know how it got to where it got to....

It was a good, positive process...it's robust but inclusive and I think that's so important...I actually found with a [previous] review...that people have almost forgotten that report and yet the changes that it produced were huge...as a result of the conversations which took place during that process of preparing and delivering that report and...that shows the real value of this sort of process, rather than...coming in from on high, doing a paper exercise and issuing a document. People might remember the document but actually the change it produces is maybe less lasting and less significant, whereas, if you have an inclusive process, it's more likely to stick, even though people may not remember the report so well. (JCPO).

Sometimes...evaluation...can be...academic – done by academics for academics, really...but we need to be able to easily relate to what that means for practice...if it doesn't do that, it gets lost quite quickly. (WPM).

A consistent finding from interview was that the quality of participant engagement was established pre-evaluation, therefore the process of contracting-in was essential to participant experience, regardless of the approach taken during data collection. The pre-evaluation paperwork sent out to each service in the weeks prior to evaluation included a letter to management providing an overview of the study (Appendix 6), a checklist of the evaluation requirements (Appendix 7), written information for research participants (Appendix 8) and a message for service-users (Appendix 4) as well as a staff survey, request for information form and copies of the informed consent documents that would be completed with participants in person on initiation of on-site data collection (Appendices 1, 2 and 9). This easily-underestimated stage of the process was, in fact, critical in establishing a participatory ethos:

It's not really been that inconvenient in any way...it's been quite open and transparent...I've not been concerned about anything. I think it's all been very clearly laid out from the beginning when you...emailed all of the documentation so we knew exactly what we were looking at and what we were gonna expect. (JPO1).

We had the HMIP inspection and the process was torturous – the appetite for pre-inspection information was voracious. I think it sapped the energy, certainly of managers, before they came. I still felt when we got that report, it wasn't culturally sensitive. I felt... "scarred" is too strong a word, but I can still remember some of those demands. I really felt that we were being done *to*. The pre-visit process for the CPAI was very, very good and could be underestimated. This has felt more in the 'critical friend' relationship than the previous inspection. Staff have felt that it's an extension of normal service – a sense of normality. The pre-evaluation paperwork was essential – very, very useful, helpful. The sending of the consent forms was evidence of the ethos and ethics behind evaluation, which sat well with service staff. (JACPO).

The full impact of pre-evaluation engagement on participant experience is revealed in contrasting responses from the women's project, where the CRC and NPS were at different organisational stages at initiation of the pre-evaluation process. In one of these agencies practitioners were engaged at an earlier point and had time to read, absorb and discuss the pre-evaluation paperwork prior to meeting with the evaluator to complete informed consent processes. Interestingly, they saw their involvement as fully participatory:

We've been involved in the evaluation process from the start. You've kept us up to date with...what's gone on. We've engaged with you so, yeah...I've felt involved. (WPO1).

I think you had a good understanding of what we were trying to achieve so...I felt like a full participant in the study. (WPO2).

On the other hand, practitioners in the second agency were unavailable in the early pre-evaluation stages due to organisational issues (including discrepancies over role designation, lack of premises, etc). The evaluation was seen as imposed and exploitative of the service, in spite of their engagement with informed consent processes at the outset of their involvement:

[Participation in the evaluation] was just like something that was expected of us...we were told just to go in the [project] and then it wasn't just [the CPAI evaluation] – there were a multitude of people contacting [us], wanting a piece of the pie and so we were just kind of “ok, we gotta do this, we gotta do this ‘cause that's expected...rather than sort of actively thinking “oh yeah, I wanna be involved in this process”. (WPO3).

Language: Terminology that “Got in the Way”

As hypothesised in Chapter Two (pp. 64-66), some of the language of the CPAI-2010 was identified as culturally and ideologically inappropriate by interviewees. In fact, findings regarding the need for cultural transposition initially emerged in Phase One of the study, during the evaluation interviews. Fieldwork diary entries for both evaluations note that at this point participants' understanding of the words 'programme' and 'treatment' in particular was in conflict with the intended meaning of the CPAI-2010 and required clarification from the evaluator:

18th July 2014...The CPAI scripts for interview questions relating to almost every item have the word “program” in them and this has been a point of contention in every single interview, to a surprising degree. Having clarified at some length that the CPAI-2010 definition of ‘program’ covers all elements of the service, staff still query the relevance of questions relating to ‘programmes’, so entrenched is their understanding of programmes as specific requirements of an Order. ‘Treatment’ is also proving to be a problematic term. It’s quite common for practitioners to ask what I mean when I ask about “treatment groups” or “treatment strategies”, the concept of treatment appears incongruous to them and, in this case, they understand the questions if I simply remove the word. (Fieldwork diary for the women’s project evaluation).

14th September 2016...Staff are really struggling with the CPAI’s use of ‘treatment’ and ‘program’ – a number haven’t completed sections of the staff member survey, including ‘Program name’ and ‘Before coming to this program have you worked for other treatment programs with offenders?’ telling me in interview that’s because they didn’t understand what was meant or expected in these sections. (Fieldwork diary for the JPACS evaluation).

Participants in Phase Two interviews summed up:

The language is difficult to understand. It’s very American. I’m thinking about programme as being group programmes and interventions rather than supervision. ‘Treatment’ – we don’t really use that. (JTL).

What do you mean by programme? Am I doing a programme? Treatment programmes? I don’t do treatment! That’s medical. (JPO2).

This particular linguistic issue was identified across both services and all levels of staff, indicating a shared cultural terminology, regardless of the service’s or individual’s exposure to the literature relating to effective practice (which itself often uses the terminology of ‘treatment programs’).

The use of technical terminology (relating, for example, to cognitive-behavioural strategies and research methodologies) was also found to have impeded engagement with both CPAI interviews and reports, though not to the extent of culturally inappropriate terminology. Participants identified that they required clarification, repetition and/or additional time to engage meaningfully with more technical terminology, indicating that difficulties were not the result of conceptual misunderstanding but that the articulation of concepts could be simplified to improve accessibility:

I struggled [with] some of the terminology that you used [in CPAI interview]...but you were really clear and explained what you wanted from me...You lost me when you were talking...and I think you had to give me clarity...something about treatment strategies, I think and I was like [*mimes going over their head*] whoop!.... (JPO3).

I did understand [the CPAI interview questions] but...sometimes I needed a bit of time to reflect on a question, just to get my own understanding of it...So, I suppose sometimes...things could be a bit plainer....Sometimes the language could be a bit...simpler. (WPO1).

It wasn't a simplistic [*sic*] report, put it that way...it took me a while to break it down. Like with this statement: "The [project's] intention to generate recidivism norms on its own population reflects..."....I have to read it about three times and say "right, ok, what are you trying to say here because I don't understand...." (WPO3).

Though practitioners and managers emphasised the importance of service-user engagement in evaluation processes, it is notable that little consideration was given to the accessibility of the report to service-users. In the women's project, the summary report was disseminated only to staff and project partners and, whilst JPACS made the report available on their open-access website, it was generally accepted that service-users were not its intended audience. When asked about accessibility, participants referred almost exclusively to the understanding of practitioners ("It was very clear, the wording of the report, and generally how it was presented was in a useful way for practitioners" (WPM)). In fact, only one participant in the study identified service-user access as a consideration, noting that, due to the report's technical language, "I couldn't hand this to one of my offenders and say "read and absorb"...", and pointing out some sections as less relevant to service-users in any case:

I don't know how interesting all of it would be to them. I suppose the [service] delivery parts may be useful [to them] but then staff training [for example]...you know, there's certain things [in the report] that are more focused on the office. I don't know whether they'd be interested in that realistically. (WPO3).

Ideological concerns related primarily to the use of the word "correctional". Participants saw this as distancing, authoritarian and punitive, ill-fitted to services which aim to engage, support and empower their users:

The thing I've...learned...working with women is, if you come from a sort of punishment, correctional angle, you're already putting up a barrier.... (WPO1).

"Correctional" – that word...It's not very caring. It's quite superior and 'know-it-all'. (JPO2).

I do hate the word 'corrections'...it's not who we are...We're about helping people make the changes they need to make and there's a monitoring and enforcement function as well, an accountability function. But...we don't make people change and you can't make people change and 'corrections' implies...that you're applying corrections to them. (JCPO).

Ideological issues regarding the term “punisher” were also raised in a team meeting at JPACS to discuss the report:

‘Punishers’ was a controversial term...no-one in the team likes that term... ‘Reinforcers’ everyone’s happy with and they know what...you mean by punishers and...there was discussion about how we can express disapproval without unnecessarily having destructive confrontation...[but] the word ‘punisher’...sort of implies physically rapping someone across the knuckles...or applying a curfew to them or something in a way that’s not very helpful....They recognised what you were saying, there was a useful conversation about it but initially the term got in the way. (JCPO).

From a commissioning perspective, however, it was the categorisation of services that emerged as the most significant linguistic issue in the use of the CPAI-2010. The women’s Project Manager identified the ratings scale range (ie, ‘Unsatisfactory’ to ‘Very Satisfactory’) as inadequate in the context of local measures using ‘Very Good’ and ‘Excellent’, etc, noting that “...even if [the project] scored ‘Very Satisfactory’ people wouldn’t necessarily associate that with very good or best practice” (WPM). The idea that the CPAI’s categorisation terminology may not be meaningful was also highlighted in Jersey, where the ‘Very Satisfactory’ categorisation for JPACS’s community supervision services was described by the Chief Probation Officer (CPO) as “great...as long as you know where it lies on the rating scale”. Moreover, the potential impact of the label ‘Unsatisfactory’ was a barrier to report dissemination in the women’s project, where feedback on categorisation during Phase Two interviews was limited to the Project Manager, since, at the request of the senior managerial team, all other evaluation participants had access only to a summary report in which the scoring and categorisation were omitted:

I was keen to get the good learning out of [the evaluation] but if I had shared the report with partners the only thing that would have come back was ‘it scored Unsatisfactory - it doesn’t work, it’s not good’, and then to try to build on that would have been very difficult. (WPM).

Similarly, in JPACS, the potential of a lower categorisation was seen as a barrier to evaluations of other services, particularly in a turbulent political climate:

Words are so difficult, aren’t they?...And I know...because...this year we were looking to evaluate our prison stuff because it’s an area where I think we can do better...Now, I could use [the CPAI]. It’d be great to come back and use it specifically on that. I’d be quite cautious about doing that bearing in mind the threat to us at the moment in case it came

out...I suspect it would still come out ‘Satisfactory’...whether it would make the ‘Very Satisfactory’, I don’t know....” (JCPO).

Of course, issues around negative evaluation findings are not solely to do with the way in which they are articulated, as JPACS’s CPO explained:

I guess...that’s not a CPAI issue, that’s more of a political issue around inspections and reviews, isn’t it – that the...areas that really need inspection and review are the ones least likely to be done if you then have a commitment to publish it. (JCPO).

Nevertheless, the language of the CPAI was seen to label and stigmatise, rather than motivate and facilitate, and this was identified as an impediment to service growth and, ultimately, a barrier to its use in both services:

You would really want to use [the CPAI] where you suspect something might be, at the moment, unsatisfactory and you want to progress it and you want to...start a process of measuring improvement...yet you don’t want to expose yourself to the funder with an unsatisfactory service, do you?...I wonder whether you could...have a...baseline measure which is less...prejudicial, perhaps?... (JCPO).

[An ‘Unsatisfactory’ label means that] good practice might not look like good practice because it’s not perfect practice, and then we might lose the opportunity...to consolidate and build on that good practice. Also, it’s quite disheartening for people...so I think having...more of a differentiation between absolutely bare minimum...and then...silver standard and gold standard, if you like, might be useful for projects to be able to meet some of it but not all. (WPM).

Depending on Context: Evaluation is a Risky Business

As the findings regarding the language of categorisation indicate, the political context was a predominant concern for evaluation commissioners. It was only staff with operational responsibility for their agencies who identified this as an issue with regards to the function of evaluations, but the impact of concerns at this level was definitive, influencing the decision to commission evaluation and/or disseminate findings.

The women’s Project Manager identified the need to justify services in the Transforming Rehabilitation (TR) environment as the impetus for evaluation, since:

With Payment by Results, everything’s outcome measured...everything relates to preventing or reducing reoffending and evaluation helps us to demonstrate that we’re doing that or shows us what we need to do better. In the current climate, that’s pivotal for us. (WPM).

Similarly, JPACS's high CPAI-2010 score and 'Very Satisfactory' categorisation was seen by the CPO as justification for the service at a time of unprecedented threat to its governance:

...the scoring's been quite useful to us 'cause, you know, I'm waving it at every opportunity – anyone has a pop at us...it's one of the things I can use, because, you know, there are these questions about governance at the moment...the question is... 'does the Jersey Probation Service become part of the executive rather than the judiciary? Does it become a States of Jersey department?' and...in discussions with officials...the CPAI is one of the things which is being trotted out to say, "well, actually, ...it's doing very well under its current stewardship". (JCPO).

However, the same political contexts were acknowledged as increasing the risks of evaluation and provided a strong impetus to suppress unfavourable findings and avoid scrutiny of services that were less likely to achieve the highest categorisation. In Jersey, CPAI evaluation of JPACS's services for prisoners was seen as too risky in its contemporaneous political context:

....the CPAI would be a great tool, I think. It would help us focus on...delivery of services to prisoners...and I think it could be really helpful for us and I suspect we'll probably do it once the dust has settled...Normally it wouldn't worry us – we'd go ahead. It's because we've got such a long history and a good track record [for evaluating services] that we'd want to do it and then, you know, produce an action plan out of it 'cause that's what you hope to get from it, is an action plan about how to improve, and then act on it and then do a follow-up and show how you've improved. But at the moment, we're in difficult political times so I'm gonna defer that one this year...until we're clear where we are, and that's just a political move. (JCPO).

In the women's project, the CPAI scoring and categorisation were suppressed largely due to concerns that a project labelled 'Unsatisfactory' would not be able to sell itself effectively to the existing and potential partners necessary to ensure its survival in the free market of the TR environment. This conceptualisation of the function of evaluation as promotional was reinforced by the responses of the manager of one of the project's partner agencies who identified a critical function of evaluation as informing third sector agencies about whether to enter into partnership with the service and what to dedicate in terms of resources and commitment:

As a third sector provider...I really wanted to have our voice in [the evaluation] because...if there was another third sector organisation thinking of being involved in something like this, it's really useful to be able to see what other people's experiences have

been like...I think we might have approached [the project] differently if we knew how things were going to pan out...I wouldn't have done things the same way...If somebody asked me now "oh, I'm thinking of getting involved in [the project]"...I'd say to them "I'd be careful of how much I give to it"...Being able to see [the evaluation] would be really useful...for other people to see what the reality is. (WPA1M).

Critical Timing and Dynamic Evaluation

The need for well-timed and dynamic evaluation emerged as a recurrent theme in interviews in both services. In the women's project this related to the evaluation taking place during the early stages of its pilot, when practitioners felt they were not yet well-enough established to engage meaningfully and perform to potential:

I think that we've been so busy over here that we haven't really maybe given enough time to the process and, because it's all new, we're finding our feet anyway. So maybe it would have been good to have been a bit calmer over here and when you were observing a group that it would've been a bit better set up and that kind of thing. (WPO1).

For the Project Manager, evaluating at this early developmental stage raised issues around findings which measure current performance against ideal standards (particularly the stigmatising category labels) and do not take into account the distance travelled by a service:

Because we [were evaluated during] the development phase...the [CPAI] methodology is quite challenging, because...we might be developing good things [that are] better than what we've had before... but [that] might not come across...It's a wasted resource...to have somebody come in and say "this is not there yet" when we know it's not...the [CPAI] doesn't necessarily allow for taking...context into account...You might have had absolutely...nought before and now you might have kind of 50% and it looks really rubbish...when actually from where we were [it] might be a huge step in the right direction....To have something that tells us [the project's] "unsatisfactory" is...not in line with where we're getting to...in terms of where we've been, or what we're trying to achieve, it might be absolutely satisfactory that we have achieved half of an ideal world. (WPM).

For the heads of both services, useful evaluation was seen as enabling "a process of continuing improvement" (JCPO) and the potential to measure distance travelled was seen as a possible solution to the problem identified by the women's Project Manager:

More weight [could have] been given to...the context or the background of where the project was coming from and what it was hoping to achieve – to put it in...that context...This would be a complete, perhaps, re-working of the [CPAI-2010]...[but] it would be really useful to have known that...previously, services for women were at, like, this level and now we're there. So that would...have been more useful at this stage of the

project for us than [being] looked at...against this gold standard and it's not there...It would be useful to have some of those high standards in there...even when they may not be likely to be achieved in a UK adult offender context...at the moment...I think that [a] way of balancing it out [would be] to say...what level is this project coming from? (WPM).

Repeated evaluation was considered essential in enabling this process of continuing improvement. In the women's project, where interviews for Phase Two took place approximately one year after data collection for Phase One (which was conducted during the first months of the pilot), participants were keen for re-evaluation to track change, to "see [if] recommendations [are] the same, if they're different, if there's more, if there's less..." (WPO1) and to include issues that had arisen since the first evaluation:

I don't think, at the time, I don't think anything was missed out. What I would say is that because it was so early on...It's interesting, now I'm thinking "oh, there's x, y and z issues" that we probably didn't discuss then because it wasn't an issue back then...so it would have been useful if there was kind of more ongoing evaluation...if it was more regular. (WPA1M).

In JPACS, participants identified repeated evaluation as an essential process to undergo in order to maintain accountability and momentum in the implementation of recommendations, though longer intervals of three years between evaluations were suggested for its established services:

[Repeating evaluation] will show what areas have been tackled and whether it's working or not. And for the...recommendations...that haven't been tackled...questioning "why not?" I would probably say...three-yearly...would be useful...because...with our own service demands...that could be...a bit heavy...it being too regular.... But...you need to keep it alive, don't you, because it's an ongoing thing. It doesn't stop just because the report's done....you want to avoid complacency, don't you. (JPO3).

We're about to launch...services...in a new area.... In that setting you would plan perhaps annual CPAs to measure the distance travelled until you get to the point where you're Very Satisfactory in your new organisation...and then you'd want to do it periodically in order to maintain your effectiveness.... I would guess in our current cycle [with JPACS's community supervision services], it's probably every three years or so, given where we are as an organisation. (JCPO).

A Point to Points: Scoring as "Adding Value"

Notwithstanding related issues regarding the language of categorisation, political contexts and developmental stages of services, scoring was nevertheless seen as a

valuable and important aspect of evaluation. In terms of repeating evaluations, scoring was seen to provide a benchmark measure as a framework for interpreting progress:

We will be looking at supervision [again] in the next 2-3 years and currently the CPAI provides a great fit for that so, you know, why wouldn't we? And then we've got the benchmark scores from last time and we'd be looking to maintain or improve on those...so...it gives you that measure. (JCPO).

Though qualitative findings and recommendations were perceived as the *substance* of the report, scoring was identified as “[adding] value because, if you’re in the right mindset, it motivates you” (WPO1) and “[making] people take a little bit more notice” of findings (WPO4). Meanwhile, JPACS’s CPO also highlighted the potential for complacency without scoring:

[The CPAI] gives you a score and I think [that’s] useful. [We’re interested in] percentages for compliance with standards. We’d also be trying to pull out some more qualitative measurements and getting stakeholder feedback into it. [Without scoring, though] people might just get complacent and not do a lot and we’ve certainly seen that [in some local services] over the years. [JCPO].

An Evidence Base Fit for Purpose

The value of evidence-based evaluation was clear in both services. As stated by JPACS’s CPO, the CPAI was seen as a “great fit” for its community supervision services, which was, perhaps, unsurprising given the service’s familiarity with and commitment to the Risk-Need-Responsivity (RNR) evidence base. Use of the CPAI-2010 was seen as refreshing and consolidating a well-established practice approach and represented a natural progression for the service:

We’ve struggled with finding a validated toolkit for [evaluation] and that’s something the CPAI gives you, of course – it’s evidence-based...It reminded people of the evidence and it started them thinking again about “well...how do we achieve that?” and there have been some great discussions about that. (JCPO).

In the women’s project, evidence-based evaluation was seen as providing a framework upon which to build an effective and justifiable service:

[It's important] to know that...we're doing stuff that works...looking at existing research and knowing that we're delivering that. (WPO2).

It's public services, it's public money and we're in a domain now where we need to know that what we're doing is the right thing...and I think...evaluation's the only way for us to do that really, is to have that evidence base.... Some of the detail within the recommendations has helped to inform thinking about...what the shape of [effective services] might look like... and that can provide a helpful framework.... Things tend to get watered down from a...pure, academic "What Works". When we come down into the world sometimes...because of tight resources...things can get watered down slightly, so I think...having a kind of very best practice standard is really useful then to be able to work towards. (WPM).

However, the gender-neutral base of the CPAI was found to fall short of the needs of women's services, where practitioners' efforts to respond to the distinct needs of women were at times frustrated by a knowledge deficit regarding effective practices:

For me, [a gender-responsive evaluation] would be useful...because ...when you've worked specifically on a women-only caseload, you do notice massive differences in their need, their risks, their reasons for non-attending...their reasons for going out and shoplifting and things like that, they do tend to be different to men's and I think we work quite well sometimes to the gender, but then other times we don't. (WPO1).

Furthermore, the knowledge base for gender-responsive services was seen to be incomplete and in flux, once again necessitating a dynamic approach to evaluation which both incorporates and adds to a burgeoning evidence base:

With women's services...there's not a massive wealth of relevant evidence that you can pinpoint to say exactly what will work with this particular one group of women...so you're contributing to the evidence base as well as working from it all of the time... But if there isn't any...evidence base linking it back to why it might reduce reoffending...if it's not related to achieving outcomes, then should we be doing it...? We need to know that what we're doing is the right thing to be doing.... We need to be evidence-based. (WPM).

Organisational Context: Dissemination, Direction and Accountability

Regardless of all else, the consensus across both services was that evaluation can only be useful if it is used - if recommendations are acted upon. In this regard, the organisational context of services emerged as a crucial component of the evaluation experience.

In JPACS (which scored particularly highly in CPAI-2010 domains related to its organisational context), a clear managerial direction was taken in order to mobilise

resources around dissemination of findings and implementation of recommendations, a process which, crucially, involved practitioners in reviewing the findings and formulating plans:

We had a whole...staff meeting on the report. We went through it, we discussed the conclusions and recommendations and in a number of areas we've taken things forward.... So, for example, we're making enquiries about using other measures to assess responsiveness.... We discussed using the LSI-R more often.... We've talked about establishing a user-group and how we can do that...[the evaluation] hasn't gone away.... It's not huge stuff and it's not finished yet but there's no point having a review if you don't look to act on the areas where you could improve. (JCPO).

For practitioners, there were two key ingredients in a worthwhile evaluation. Firstly, they needed to accept its recommendations (a key outcome of inclusive, participatory evaluation in contrast with paper-process inspection) and, secondly, they needed confidence in the organisation's ability and commitment to implement them. This was immediately evident in responses from JPACS practitioners, who were able to describe a tangible course of action, expressing belief in the organisation's efficacy:

It was useful...[The evaluation] recommendations...are valid and I think it is the case that the service will take them forward and take note of them. So that's all good...We had a practice meeting to look over the report...we went through...each step. I think it was [the Assistant Chief Probation Officer and Team Leader] who did the presentations.... Some [requirements] we could implement straight away and others, I think we're gonna work towards looking at in a bit more detail in team meetings...and I think [taking recommendations forward is] the whole purpose, isn't it, of when something's evaluated. I mean, it makes it [worthless] otherwise, doesn't it? (JPO3).

For the manager of the women's project, the value of the CPAI report's recommendations was clear:

The recommendations are really, really valuable, and I think probably the most useful bit of [the evaluation]...because [they give] some direction, so even if you aren't able to take all of them forward...even if you'll just be able to take some of them forward and to be able to articulate what it is that we're doing that meets those or not and think about practising that way, I think really helped to develop the project in an evidence-based way. (WPM).

However, the organisational deficits identified in the evaluation, including those concerned with management characteristics, resulted in every other women's project participant in Phase Two interviews expressing doubt that the report would be acted

upon and, ultimately, perceiving the evaluation as a futile exercise. Three major deficits were identified by practitioners:

1. Lack of participation in dissemination, planning and implementation processes: for the manager of one partner agency, this was due to her not having received a copy of the summary report (possibly due to an administrative error or managerial oversight), which resulted in low morale and an inability to improve:

It is disappointing that we weren't given the [evaluation report]...On a personal level, it's [bad for] morale because you feel you're not being valued but, secondly, there might be information in there that would impact on our service. For example, there might have been comments on how we play our part in the [project] or how we deliver the service that we would have welcomed to make necessary changes and if we don't get that feedback we can't...deliver the changes. (WPA1M).

For those participants who *did* receive copies of the summary report, this did little to make them feel included (nor to make them feel the evaluation was worthwhile). Instead, the lack of discourse following evaluation deepened perceptions of division and inequity between the managerial team and front-line staff:

I think [the evaluation report recommendations] are only useful if the gods that be are gonna implement them because if they...don't choose to implement them, you know, most of these recommendations are not things that...we can take forward. So, you know, certainly it hasn't been discussed at any meeting...not at our level.... It might have been [by managers] but certainly not to us...we don't even get invited to meetings. (WPO3).

I suppose it would be nice just to have had a meeting, all of us who were initially involved...just to say..."here's the report, the feedback...what's your feeling on this?" or "we've decided we're just gonna take the report and leave it there now and not take it any further" or "actually, we've looked at the feedback but we think that some of the recommendations might be worth..." but the only person I see who's involved is you...No one else has really, yeah, which is sad, I suppose...I just feel...I'm just like the minion involved in it and I don't really have a say in anything. (WPO1).

2. Lack of confidence in organisational direction, resources and support:

So, it has to come top down and across the board.... [This evaluation report] is only ever going to be useful if the person in charge takes it and says, "Ok, out of these we can implement at least these, and let's roll with it, and this is what you need..." and then a direction is set but it's not useful

otherwise because, you know, no disrespect to probation but probably we'll never see this document again or hear about it.... It's a case of, I think, out-of-sight, out-of-mind. So, it's useful and relevant when we read it and we receive it and you think, "Right, ok, yeah," and then, it validates the [evaluation] process that we had to go through, so the evaluation makes you think, "Ok, so something came of it, you didn't waste your time," but then...nothing.... It's just...unless somebody does something with it...we're just dealing with whatever we're dealing with day-to-day. (WPO3).

I'd like to think that [the report's] being used appropriately, I suppose – that it's worth the paper that it's written on! I think [evaluation's] useful if it's used.... If I'm honest, I don't know if anybody's really taking it on board and is gonna move forwards with it...I really haven't had any feedback from anybody about what was highlighted – if they think that we need to do something about it...and I haven't really seen any change so, to me, things have been highlighted but they haven't really thought, "Ok, well maybe we need to work on that."... There's stuff in there that I could do myself on the recommendations, but I also need support.... You know, I'm only a lowly [practitioner] at the end of the day, I can't do stuff without management. (WPO1).

3. Lack of accountability:

[Evaluation is necessary and important] if it's gonna be used for something. Because I have seen evaluations just been done and they just sit somewhere or people just use them to kind of go, "Well, we think we've done great. We've got an evaluation that says we've done great. Thank you very much!" and...it needs to inform practice doesn't it...there needs to be some accountability. (WPA1M).

What would be interesting is if there is a report and there are recommendations, who is accountable for that report and following those recommendations up? And will that person...be held accountable by someone to make sure they've done it? 'Cause I'd be interested if you came back in a year's time and we re-visited the report with a list of recommendations – have those been implemented? Or have you just been paid a sum of money to go and do a job and then it's forgotten about, you know? That would be really frustrating. (WPO4).

Interestingly, for practitioners in the women's project, the potential for evaluation to address issues relating to deficits in the organisational context was key to its utility:

[I hope] the way...those of use who are working on [the project] feel has come across in the report.... I suppose it's just a bit of confirmation that...the way that [the project's] been set up isn't really the best way...that it was set up quite quickly and there wasn't maybe enough thought into the little things that needed to be in place. [I'd want] that information to be used so [in future] staff who get involved don't have to go through [what we did]. (WPO1).

Conclusion

When drawing conclusions from Phase Two of the study, it is important to note that participants in feedback interviews were professionals within each service. In all of the participant responses, positive evaluation was viewed as a participatory and inclusive process, with service-users as key players, though the outcomes (ie, the report) remained the domain of professionals and academics. Due to the absence of service-user participation in Phase Two of this study it is not possible to reach any conclusions regarding their interest and/or needs/preferences in accessing evaluation findings. It would not be helpful to make any assumptions here and further research needs to be done specifically to engage service-users in discussion around this. Nevertheless, the purpose of this study is to develop evaluation approaches and instruments that are useful and meaningful for agencies. The input from managers and practitioners explored in this chapter is crucial in informing this development.

A significant finding is that holistic, participatory evaluation (requiring an independent and interpersonally-skilled evaluator) engages services in a process of reflection that is the foundation for change through acceptance of recommendations. However, in order to progress this through to effective implementation, the organisational context itself must be participatory, efficient and accountable. This raises the question of how an organisation can move forward to action recommendations regarding its own context. In other words, if the CPAI finds a service's managerial direction to be deficient, how can the evaluation help to enable that service to improve? The answer appears to lie, at least in part, in understanding the shifting, turbulent, political contexts of organisations and the impact of language used in the evaluation, particularly regarding categorisation.

The potential political impact of its Unsatisfactory CPAI classification significantly delayed the women's project's dissemination of evaluation findings (the summary report was issued by email several months after the CPAI data collection period ended and was not sent to all participants). This alienated most participants from the process while senior leads within the National Offender Management Service (NOMS) suppressed aspects of the original report (specifically the scoring and categorisation) and decided which aspects should be carried forward. For practitioners, who were

excluded from these discussions (and, indeed, the full findings of the evaluation), the evaluation was seen to have atrophied, becoming a waste of paper and a waste of time, reducing their confidence in and enthusiasm for both the organisation and the evaluation process. There appears to be little hope that the evaluation recommendations could be used effectively to develop practice in this context, regardless of the intentions of management – the importance of participatory evaluation cannot be overstated based on the findings of this study.

Had the language of categorisation not been perceived as stigmatising, and had the evaluation taken account of distance travelled as well as measuring services against the benchmarks for effective practice, the report would have been less volatile politically (at the same time retaining its evidence base). This would have given the potential for earlier and wider dissemination and discussion, enabling the participatory and supportive approach necessary to implement recommendations and ensuring greater accountability for organisational leads through shared planning and agreed responsibilities. Of course, this is not a solution to organisational deficits in and of itself – an organisation still needs to commit to evaluation (and re-evaluation) as an ongoing process of improvement and to dedicate resources where needed – but it increases the possibility of engaging an organisation and of securing that commitment. It is also the case that timing re-evaluation at an appropriate frequency and taking account of distance travelled would help evaluation to address the wider context of services through reflecting the turbulent political environment.

Furthermore, issues around language were, in fact, fourfold. In addition to the stigmatising and labelling effects of categorisation and use of overly-complex, technical language (emanating from cognitive behaviourism and research methods), CPAI terminology was found to be culturally and ideologically inappropriate, with certain terms being misunderstood in the context of services in the British Isles (eg, ‘treatment’ and ‘programme’) and others conceptualised as punitive (eg, ‘correctional’ and ‘punisher’), which were anathema to the sensibilities of inclusive, empowering services.

These findings, together with arguments for the inclusion of gender-responsive evidence and the benefits of scoring, will be taken forward into the next chapter's discussion of the development of implementable, culturally- and ideologically- sensitive evaluation approaches and instruments for services in the British Isles.

CHAPTER SIX

Amendments, Adaptations and Developments: The Swansea Service Evaluation Inventory – Women’s Projects (SSEI-W)

Introduction

The findings described in Chapters Four and Five indicate that the Correctional Program Assessment Inventory (CPAI) offers a significant contribution to evaluation, not least because of its holistic approach, but they also reveal some clear limitations to its usefulness in the British Isles and particularly in women’s projects. It could be argued that amendments to the language used in interview questions and the phrasing of the report might be made to address some of these issues without significantly deviating from the format of the CPAI. The growing evidence base for gender-responsive work, however, is a vital consideration when developing meaningful evaluation processes and requires the development of a bespoke instrument for women’s projects.

Copyright issues prevent straightforward adaptation of the CPAI and present further obstacles to its use in under-resourced women’s services, for which the cost attached to licensing and administering the tool is likely to be prohibitive. In any case, it seems pertinent to return to the evidence base with a focus on women in order to develop a gender-responsive evaluation instrument from the ground up which will guide women’s projects in developing effective practices within safe spaces and ensure that the instrument is accessible to all. This chapter discusses the development of such an instrument, the Swansea Service Evaluation Inventory – Women’s projects (SSEI-W, Appendix 13), drawing upon findings from this study and the Risk-Need-Responsivity (RNR), desistance and gender-responsive literatures.

Dimensions and Items of the SSEI-W

The SSEI-W retains the holistic, participatory approach of the CPAI, involving the collection of data through interview, observation and document-reads to address 152 items across 8 dimensions:

A. Description of Service	8 items (none scored)
B. Leadership & Innovation	23 items
C. Quality Assurance	8 items
D. Staff Attributes & Development	13 items
E. Assessment Practices & Impact	13 items
F. Service Characteristics	18 items
G. Collaborative Working	6 items
H. Practice Skills	63 items

Similar to the CPAI, the first dimension, Dimension A: Description of Service, is descriptive only and contextualises the evaluation. The scoring system for the remaining dimensions is also modelled on that of the CPAI – each item is scored yes (1) or no (0) and a verification score (VS) is given on a scale of 1-5, according to the evidence available to verify scoring for each item.

Dimension A: Description of Service

The equivalent domain in the CPAI (Domain A: Program Demographics) has received little attention in the literature relating to the instrument. Domain A (CPAI) and Dimension A (SSEI-W) do not relate to elements of effective practice and are not scored, so have been included here as “descriptive *only*” (see p. 107 and above). However, given the role of Domain A in contextualising the CPAI evaluation, and the centrality of context to this study, the development of Dimension A bears further examination here.

The completion of Domain A of the CPAI is dependent upon the agencies’ responses to the pre-evaluation Request for Information Form (Appendix 2) and the evaluator’s own

observations of the service premises and environment. As acknowledged in footnote 19 (p. 77), description of the demographics of service-users in this study relied on agencies' responses to the question:

10. Presenting problem

When participants initially begin the programme, what types of problems do you most often encounter (e.g. substance misuse, sexual offending, history of violence, anger problems, anti-social values, etc.)? (CPAI-2010 Request for Information, Appendix 2).

The agencies' responses to this were duplicated in the reports and, for the women's project, this included the statement "...the majority of...involvement with women service users is related to violent and/or acquisitive offending behaviour" (Kerr, 2014, p. 11). On reflection, this requires analysis since it does not fit with what is known about women as predominantly low-level, non-violent offenders (Gelsthorpe & Wright, 2015; Ministry of Justice, 2018c; 2018e; Prison Reform Trust, 2017b), nor with the individual cases reviewed by the evaluator during data collection. Although many involved crimes of "violence" (eg, Common Assault), these were, on the whole, minimally harmful offences causing minor, if any, injury. Moreover, women's mental or emotional health issues, substance misuse and problematic relationships (in their personal and public lives) were overwhelmingly at the core of their offending behaviour, as opposed to intrinsic violent tendencies or pathologies. Uncritical inclusion of the statement that service-users had mostly committed violent crimes compromises the study in three distinct ways:

1. Women's projects face a number of challenges when seeking to define the parameters of their service as a result of their holistic nature, comprising a wide range of services for both offenders and non-offenders and embracing a broad spectrum of public, private and third-sector providers with a concomitant, varied range of objectives. Following his propensity score analysis of a London community resettlement programme for women prisoners which found an increased rate of recidivism amongst participants, Sutherland observed that "[the] heterogeneity of those worked with more broadly by the programme, coupled with large variations in practice paint a picture of an intervention that did not know what it wanted to achieve, with whom or how it would do this"

(Sutherland, 2017, p. 11). Evaluation that reflects without scrutiny inaccurate data relating to the demographics of service-users both fails to address the implications of this regarding the potential misfit of services to service-users, and actively reinforces the inaccuracy, with the report itself becoming a source of misinformation for misguided practices.

2. Inaccurate and unnecessary labelling of service-users is likely to have a negative impact on perceptions of the service and its users, whilst, as discussed in footnote 16 (p. 65), using stigmatising labels such as “violent” can contribute to service-users’ self-perception, and consequent behaviour (Tannenbaum, 1938; Link et al, 1989;1999). In employing the label of “violent” in the report, the evaluation of the women’s project runs counter to desistance approaches, which require services to engage social mechanisms to reverse labelling effects and enable a service-user to have a positive role in the community (Trice & Roman, 1970, p. 539).
3. There is a clear and pressing need for research into gender-responsive approaches to support the emerging evidence-base for women. Without it, Van Voorhis argues that they “are doomed to be caught in the senseless cycle of no research, no evidence, and ongoing adherence to ‘best available evidence’ where the ‘male is norm’” (2012, p. 129). Evaluation has a critical role to play in providing evidence of effectiveness. The pressure upon the data for women’s services to be accurate and useful could not be greater.

The consequences of such a lack of attention given to unscored descriptive information may therefore be significant and far-reaching, and there is potential for evaluation approaches to encourage a more iterative and critical engagement with the data on the part of the evaluator. The SSEI-W, therefore, requires that the evaluator completes a draft of Dimension A on the basis of pre-evaluation information provided by the agency (the same process as the CPAI) and then, crucially, reflects on this information with all evaluation interview participants (including, of course, service-users) before writing up

the final report. This enables participants and the evaluator to challenge assumptions about the service and its users, opens up the evaluation to take account of different perspectives on the service's demographics, physical setting and fiscal arrangements, and engages the evaluator in a critical process across *all* dimensions of the SSEI-W.

It must be said that the CPAI does not itself preclude this iterative approach, and that its developers would likely encourage, or even expect, it from the evaluators. The uncritical inclusion of data is an error on the part of the evaluator, rather than a fault in the instrument. However, explicit guidance towards an iterative approach across the dimensions of the SSEI-W facilitates holistic evaluation, and the lack of attention given to this in the CPAI may have some serious consequences, as was demonstrated here.

Dimension B: Leadership & Innovation

Dimension B is largely framed around Ellickson and Petersilia's investigations into the correlates of successful implementation in criminal justice, which find, in short, that there must be a close fit between the service and its environment, that all stakeholders must be committed to the service at every level of its operation and that the service must be resourced and supported at an appropriate level (Ellickson & Petersilia, 1983; Petersilia, 1990). Additional components of the dimension are derived from further work exploring the processes and organisational settings correlated with success (Gendreau et al, 2001; Harland et al, 1979; Harris & Smith, 1996) and evidence from the gender-responsive literature which indicates that a commitment to prioritise women's issues must be embedded in the organisational structure (Bloom et al, 2003). This combination of evidence is the foundation of items such as:

B1. The service's goals include reducing re-offending, reflect the evidence base for effective work with women (eg, diversion from custody, the provision of social support, family reunification, protection from abusive relationships, recovery from victimisation, etc) and are clearly documented. (SSEI-W V1.4).

Dimension C: Quality Assurance

Evaluation, as experience, testimony and research have demonstrated, is a crucial component in the development of effective, sustainable services. In women's services, for which the evidence base is relatively new, incomplete and evolving, "documenting the effectiveness of practice addresses the need for empirical research on the outcomes of gender-responsive programmes" (Bloom et al, 2003, p. 89). Dimension C incorporates the evidence relating to effective quality assurance and evaluation practices, such as the need for the collection of recidivism figures twelve months or more after completion of engagement with the service (Item C5), since the recidivism of "treatment groups" is known to be delayed during that first year (Lösel, 2001, p. 80). It is also a key feature of gender-responsive practice to "go beyond traditional recidivism measures to assess the impact of specific program attributes on pathways to female criminality" (Bloom et al, 2003, p. 90). Item C2 requires service-user feedback to be sought, C3 requires "data relating to changes in targeted behaviours and human/social needs" to be "collated and analysed" and C4 addresses the need for standardised data-recording processes across the multiple agencies involved in delivering a holistic service to women, responding to the finding that differences in practice have been impeding researcher access (see Chapter Four, p. 140; and Jolliffe et al, 2011):

C4. Data relating to measures of interest (eg, whether service-users have experienced domestic abuse, whether service-users are diagnosed as having mental health issues, etc) are recorded in a standardised way, to facilitate external and internal evaluator access. (SSEI-W V1.4).

Dimension D: Staff Attributes & Development

The need for staff to be selected and trained appropriately is crucial to the delivery of effective services. As Covington and Bloom point out "programming designed for women can only be as good as its staff" (2007, p. 28). As with the other dimensions of the SSEI-W, Dimension D draws on the principles of effective intervention identified by research into exemplary services (Gendreau, 1996a; Gendreau et al, 2004) and the gender-responsive literature, which indicates, for example, that appropriately selected

and trained core practitioners are women (Item D1), are conversant in, and committed to, the evidence base for work *with* women (Item D9), and receive specialised training in this respect (Item D11) (Covington & Bloom, 2007).

Dimension E: Assessment Practices & Impact

Dimension E relates to the fundamental principle that “gender-responsive screening and assessment tools are utilized, with appropriate treatment matched to the identified needs and assets of each client” (Covington & Bloom, 2007, p. 21). Whilst it is a requirement of the SSEI-W that, in line with the RNR evidence base (including and especially for women – see Chapter One, pp. 19-24), service-users are categorised as to the level of risk that they present (eg, Item E1), assessing and responding to the strengths of service-users is also key, since:

Many women already are struggling with a poor sense of self because of the stigma attached to their addictions, their parenting histories, their trauma, or their prison records, for example. It may be non-therapeutic to add another problem to the woman’s list of perceived failures.

A strength-based (asset) model of treatment shifts the focus from targeting problems to identifying the multiple issues a woman must contend with and the strategies she has adopted to cope. (Covington & Bloom, 2007, p. 20).

Acknowledging the, perhaps unnecessary, conflict between strengths-based and risk-based approaches, the phrasing of Item E3, for example, aims to appeal to the common ground, since the “protective factors” of the RNR approach are the “capabilities” of a strengths perspective:

E3. Service-users’ capabilities/protective factors are assessed as predictors of the reduced likelihood of reoffending or of desistance. (SSEI-W V1.4).

Dimension E also emphasises the importance of engaging women in assessment and case-planning through “reciprocal, respectful and empathic practitioner interactions” (Item E8) and responds to findings from the pilot of the CPAI-2010 relating to the matching of practitioner to service-user, in which case allocation according to formal assessment was viewed by participants as entirely impractical and not necessarily

desirable (see Chapter Four, pp. 123-124). Returning to the evidence indicates that practitioners skilled in relationship-building will be able to work effectively with the majority of service-users (Palmer, 1996; Traux & Carkhuff, 1976; Raynor et al, 2014), though it is clear from the gender-responsive literature that the assessment process and delivery of services needs to be individualised, for example, to respond to the service-user's "language, culture, literacy level and cognitive functioning" (Covington & Bloom, 2007, p. 22). To facilitate this without demanding formal matching through separate additional assessment processes, the SSEI-W asks the responsible member of staff, "How do you allocate practitioners to service-users?" in order to address Item E11:

E11. Consideration is given to characteristics that may facilitate or impede an effective working relationship (eg, personal regard, learning/facilitation style, conceptual level, language, culture, etc) when allocating practitioners to service-users. (SSEI-W V1.4).

Dimension F: Service Characteristics

Chapter One outlined the evidence for social learning and cognitive-behavioural work with women (pp. 29-32), and the literature regarding behavioural principles and procedures (eg, Spiegler & Guevremont, 2009) is the basis for items concerning cognitive-behavioural strategies (Item F6), etc. This dimension also incorporates key findings from the gender-responsive literature. For example, the service is required to provide a women-only space (Item F1) "to foster safety and a sense of community" (Gelsthorpe et al, 2007, p. 54). Item F2 requires the service to "provide women with practical help with transport and childcare so that they can maintain their involvement" (Gelsthorpe et al, 2007, p. 54) and other items also relate to the provision of services to meet gender-specific needs, such as the integration of services for substance misuse, mental health and trauma/victimisation (Item F9) (Blanchette & Brown, 2006; Covington & Bloom, 2007).

Items in Dimension F also require services to attend to women's social and moral rehabilitation (see p. 68). In addition to being "interested in and aware of service-users' social contexts" (Item F8), the service must "provide...opportunities to improve [the]

socio-economic conditions” of service-users (Item F10), “[promote] and [support] healthy relationships with children, families, significant others, and the community” (Item F11), provide and support opportunities “for criminalised service-users to access...services...together with non-criminalised women” (Item F12), build “long-term community support networks for service-users” (Item F16), and work with service-users *and their support networks* to formulate relapse prevention strategies (Item F17). A woman’s social reintegration is incorporated alongside her personal rehabilitation as a measurement of effectiveness of the service:

F15. Service-users’ success in the service is measurable and visible using criteria such as reduction in risk category, engagement with community services, employment, child reunification, managed substance misuse, etc. (SSEI-W V1.4).

The SSEI-W departs from the CPAI’s recommendation for the use of token economies as a behavioural strategy. In spite of their sound theoretical foundation in social learning theory and operant principles (Spiegler & Guevremont, 2009) and widespread use in social settings, including education, healthcare and criminal justice (especially youth justice residential facilities in the US) (Mohr et al, 2009; VanderVen, 2016), contingency management tools such as token economies or point and level systems have been subject to a number of criticisms. The majority of these relate to poor implementation, such as their inconsistent use by poorly trained staff or their misuse as a punitive function (Mohr et al, 2009). The quality of implementation is, of course, the concern of the CPAI, indicating that evaluation itself has the potential to address these criticisms. However, the token economy/point and level emphasis on staff control of disruptive behaviours may come at the expense of service-users’ real and perceived autonomy (and the individualisation of their treatment), and at the risk of re-traumatising them through coercive and punitive treatment (Mohr et al, 2009; Rautkis, 2016; VanderVen, 1995; 2016). Therefore, the SSEI-W draws instead upon female-responsive, trauma-informed approaches, such as the Trauma Informed Effective Reinforcement System (TIER; Rothschild & Selvaggi). Covington (2018) describes this as an “empowering paradigm” that “reduces staff focus on traditional compliance methods and teaches skills that are more effective in motivating positive behaviour than traditional points and level systems”, through a positively-stated focus on behaviours

contributing to personal, interpersonal and community safety (Selvaggi et al, 2013) (see SSEI-W item F14, discussed on p. 176).

Since the evidence base for gender-responsive services is incomplete and in flux, the question of what exactly constitutes effective service has not yet been fully answered and remains open to discovery, interpretation and debate. For example, the concept of a safe “women-only” space is yet to be clearly defined in the context of transgender issues (which raise concerns about what constitutes safety for women and trans women alike). The SSEI-W must be used reflexively to respond to practice-level developments as well as the wider literature. The current version was developed to incorporate the best available empirical evidence at the time of writing, but its developers acknowledge that it is not the “final word” on effective gender-responsive practices.

Dimension G: Collaborative Working

Multi-agency working is required, or alluded to, in multiple dimensions of the SSEI-W. For instance, Items H46-48 require practitioners to make referrals and advocate for service-users where appropriate and items referring to the multiplicity of services available to women (eg, D9; F9-12) will, in practice, require practitioners’ knowledge of and involvement with a range of providers. Furthermore, the need for holistic delivery of a range of services, exemplified by the “one-stop shop” approach (Corston, 2007), is critical to the development of gender-responsive services and so is the focus of Dimension G. For example, Items G1 and G2 require “wraparound services” that will be delivered involving “a holistically and culturally sensitive plan for each woman that draws on a coordinated range of services within her community” (Covington & Bloom, 2007, p. 14). This enables services to “take a holistic and practical stance to helping women to address social problems which may be linked to their offending” (Gelsthorpe et al, 2007, p. 54).

Dimension H: Practice Skills

The final dimension, focused on the assessment of the relationship and structuring skills of practitioners, has been adapted from a pre-existing instrument, the Jersey Supervision Interview Checklist (Raynor et al, 2009). Developed and validated in the British Isles (Raynor et al, 2014), this addresses seven skill sets: interview set-up, non-verbal communication, verbal communication, use of authority, motivational interviewing, pro-social modelling, problem-solving, cognitive restructuring, and overall interview structure. Whilst these skill-sets largely correspond to the core correctional practices itemised in section G of the CPAI-2010 (see p. 57), more attention is given to the quality of service-user engagement and each practice principle is broken down into more specific (and more easily observed) items. For example, CPAI items G39-G42 (a total of four items) require that practitioners “are open warm, and...respectful” (G39); “are non-blaming, empathic, and genuine” (G40); “are flexible, use humour, and are engaging” (G41); and “are enthusiastic and express optimism” (G42). To measure these qualities, the Jersey Checklist incorporates 15 items (SSEI-W H5-H24), including specific instructions relating to the quality of non-verbal communication (eg, “open posture/arms legs uncrossed”, SSEIW item H6) as well as verbal communication (eg, “mostly open questions”, SSEI-W item H10). Importantly, caveats regarding the need for flexibility in creating “genuinely collaborative relationships” and “making active listening choices” (Nelson-Jones, 2011, p. 51), and for cultural responsiveness in communication, are provided in the Checklist’s accompanying manual (Vanstone & Raynor, 2012).

The evidence base for these items is supported across the RNR, desistance and gender-responsive literatures. For example, relationship skills, characterised by empathy, respect and warmth, etc, are a key feature of the core correctional practices identified by Andrews & Kiessling (1980) and Dowden & Andrews (2004), whilst McNeill identifies “sustained and compassionate support from a trusted source” as the bedrock of desistance processes (2006, p. 49), and the gender-responsive literature calls for “mutual, empathic and empowering” relationships to “foster growth in women” (Bloom et al, 2003, p. 55). Egan’s skilled helper model describes the importance of non-verbal

communication such as facing the service-user (SSEI-W Item H5), maintaining an open posture (H6), being attentive (H7) and appearing relaxed (H9), since “a respectful, empathic, genuine and caring mind set might well lose its impact if the client does not see these internal attitudes reflected in your external behaviours” (Egan, 2002, p. 70).

A further departure from the CPAI is the inclusion of items relating to the structure of the session. Four Checklist items (SSEI-W H1-H4) require due attention to be given to privacy, confidentiality, seating arrangements and potential distractions in the set-up of sessions. A further eight (SSEI-W H56-H63) relate to the evidence regarding task-centred casework which indicates a need for work to be facilitated through communication (about purposes, problems and plans), agreement (on the problems and how and when work will be undertaken) and guidance (through exploration, questioning, prompting, etc) (Reid & Epstein, 1972; Taxman et al, 2004; Vanstone & Raynor, 2012, p. 26).

Language: Cultural, Technical, Ideological

As discussed in Chapter Five, there were a number of issues with the language used in the CPAI interviews and reports (pp. 148-152). The CPAI’s use of the word “program” most closely translates in the British Isles to “service”, since this encompasses every level and aspect of a project’s involvement with women, including the project itself. Where the CPAI-2010 asks, for example, “How did the organization establish a need for the program before it was initially implemented?” (Item C1), the SSEI-W asks, “How was the need for this service identified?” (Item B3).

There is no need for the SSEI-W to use the culturally inappropriate language of “corrections” and “treatment”. For example, instead of, “What treatment strategies does the program use?” (CPAI-2010, Items F4-F7), the SSEI-W asks, “What strategies does the service use to encourage and empower service-users to transform their behaviour?” (Items F4-F6). As well as using more culturally-appropriate terminology, this linguistic approach focuses on the service-user as the owner of the process of desistance and the service as an enabler, encouraging and empowering rather than imposing and controlling. Additionally, describing the specific aims of strategies may ameliorate

some of the issues around use of unnecessarily technical language. The SSEI-W is being developed with accessibility as a key consideration, avoiding excessively formal phrasing and onerous questioning as well as the use of terminology found to be ideologically or culturally inappropriate in the CPAI pilots. For instance, Item F16 of the CPAI-2010 uses language perceived as authoritarian, punitive and technical, whereas the equivalent item of the SSEI-W re-articulates the same concept as a simplified description to emphasise the active participation of service-users. It draws upon the key principles of gender-responsive approaches such as the TIER System (see p. 172) which “intentionally include safety, respect and dignity at every level” and are “not built upon teaching compliance”, but instead “negative, destructive behaviors are diminished through supportive techniques that teach [service-users] the necessary skills to manage their own challenging feelings, thoughts and attitudes” (Covington, 2018).

CPAI-2010	SSEI-W V1.4
F16. Appropriate punishers	F14.
Program has identified a range of appropriate punishers that include: (a) fines, (b) loss of tokens, (c) time out, and (d) social disapproval. Punishers should be individualized to ensure they are meaningful to the clients in order to suppress behaviour. Punishers should be brief, consistent and high intensity.	Disruptive and/or unsafe behaviours are immediately and consistently dealt with using supportive measures that have been influenced and approved by service-users, are proportionate to the behaviour and appropriate to the individual service-user (eg, verbal warning, coaching, temporary privilege/community restrictions, etc).
<i>Please describe how punishers are used. Are there written procedures for applying punishers?</i>	<i>How are disruptive or unsafe behaviours dealt with?</i>

Language and Progress: “The Power of Yet”

It is resoundingly clear from the findings in Phase Two of this study that the language of the evaluation report must be positive and motivational for the purposes of removing barriers to engagement, enabling dissemination and facilitating change. An inter-related finding was that meaningful evaluation needs to locate a service on its continuum of change. In the field of education psychology, Dweck (2017) describes labels that judge students, demonstrating that such labels imply permanent traits and inculcate a “fixed mindset”, in which the student will invariably avoid challenges and resist negative feedback, prioritising instead their record and reputation. This appears to be relevant

and significant in this study, where the women’s project suppressed the full report due to its ‘Unsatisfactory’ categorisation; and JPACS embraced its ‘Very Satisfactory’ categorisation but resisted evaluation of a service it was less sure would perform well. In order to facilitate learning, Dweck advocates the use of encouraging feedback focused on process, strategies and effort, rather than judgement, to develop a “growth mindset” in students, “based on the belief that your basic qualities are things you can cultivate through your efforts” (2017, p. 7). A “growth mindset” is characterised by embracing challenges, persistence in the face of setbacks and learning from criticism leading to higher levels of achievement:

If you get a failing grade, you think “I’m nothing, I’m nowhere” but if you get the grade ‘Not Yet’ you understand that you’re on a learning curve. It gives you a path into the future. (Dweck, 2014).

However, a factor that will need to be taken into consideration is that positive language of categorisation must not diminish or obscure the evaluation’s ability to highlight practice which is detrimental, risky or lacking potential. It may be pertinent to include more categories than the CPAI in order to distinguish between services that are meeting very few of the items (and therefore require urgent attention) and those which are making progress but not yet meeting the majority of items.

Table 6.1 CPAI-2010 and Potential SSEI-W Categories

Categories of the CPAI-2010 (Fixed Mindset)		Potential Categories of the SSEI-W (Growth Mindset)	
70% +	Very Satisfactory	70% +	Extremely high level of dedication to effective practices evident
50% - 69%	Satisfactory	50% - 69%	Dedication to effective practices evident
Below 50%	Unsatisfactory	20% - 49%	Working towards effective practices
		0% - 19%	Not yet working towards effective practices

The language of the SSEI-W report subheadings is being developed to focus content on the service’s progress on its continuum of learning and change.

Table 6.2 CPAI-2010 and SSEI-W Subsections

Subsections of the CPAI-2010 pilot reports	Subsections of the SSEI-W report (Growth Mindset)
Strengths	Progress
Areas for Improvement	Opportunities
Recommendations	Potential

Scoring

As identified in Chapter Five, scoring was seen to add value to evaluation (pp. 155-156) and it is the intention of this study that the SSEI-W will be scored. In its current version, the number of items in each dimension simply represents the number of elements from which that dimension of practice is comprised. However, reference to the available evidence, such as comparison with the CPAI-2010 (which has been validated in a number of research studies – see Chapter Two, p. 55), lends support to the weighting of the SSEI-W dimensions with regards to correlations of effectiveness. As can be seen from Table 6.3, each dimension of the SSEI-W accounts for a similar percentage of the total score to corresponding domains of the CPAI-2010 (items relating to organisational context (culture, implementation and staff) have been grouped together since these are distributed differently across the two instruments).

Table 6.3 CPAI-2010 and SSEI-W Scoring Distribution

CPAI-2010			SSEI-W		
Domain	Points	%Total (/133)	Dimension	Points	%Total (/144)
Organisational Context (B, C, D)	37	28%	Organisational context (B, D)	36	25%
Risk/Need Practices (E)	13	10%	Assessment Practices (E)	13	9%
Program Characteristics (F)	25	19%	Service Characteristics (F)	18	13%
Core Correctional Practice (G)	45	34%	Practice Skills	63	44%
Inter-Agency Communication (H)	5	4%	Collaborative Working (G)	6	4%
Evaluation (I)	8	6%	Quality Assurance (C)	8	6%

Gendreau and Thériault advise that close attention should be paid to risk/need practices and service characteristics (totalling 29% of the CPAI-2010 and 22% of the SSEI-W), because the “research literature indicates the strongest validity for these two dimensions

for program success” (2010, p. 5) and the emphasis on practice skills (34% of the CPAI-2010 and 44% of the SSEI-W) is supported by the burgeoning international evidence base for the significance of effective practitioner skills (Ugwudike et al, 2018). Nevertheless, the percentage divisions for SSEI-W categorisation will need to be validated through comparison between evaluation scoring and reoffending figures during pilots of the instrument.

Usability

Developing an instrument from the ground up also provides the opportunity to build in features to improve its usability. The CPAI-2010 is a paper document in which all 133 scored items and the corresponding questions are listed under the relevant domains across 47 pages. This format had an impact on the evaluation experience from the evaluator’s perspective: the flow of interviews was sometimes interrupted by the evaluator searching for the next relevant item, and the efficiency and accuracy of the process may have been impeded since the evaluator was transcribing from hastily written notes when compiling the reports.

The SSEI-W was developed in Excel so that, while it currently functions as a PDF paper document like the CPAI, there is scope for its future development as software which could incorporate a number of features such as being able to draw together relevant items for interviews with specific individuals and automatic calculation of percentage scores, etc. This would also enable evaluators to undertake the evaluation using a tablet or other electronic device. However, further research will need to be done to assess the desirability of digitisation in this context, as the potential for technology to depersonalise practice has been noted in many fields (see, eg, Higgs, 2001). It may be that its introduction here would be detrimental to the relationship between evaluator and participants, demonstrated by this study as pivotal to the efficacy of evaluation (see Chapter Five).

Conclusion

The SSEI-W is being developed to retain the holistic, participatory approach of the CPAI along with the inclusion of organisational context as a key feature (Dimensions B

and D). The evidence base for gender-responsive work (Covington & Bloom, 2007; Gelsthorpe et al, 2007) is integrated with the principles for effective practice from the RNR literature (Gendreau et al, 2004) and attention is given to language to complete its development as a response to findings from this study (Chapters Four and Five). A full version of the SSEI-W is appended to this thesis (Appendix 13), which must now be piloted, and feedback collected from services to facilitate its future development.

Chapters Seven and Eight will reflect on the aims of this study and contextualise its findings and outcomes in the current political, academic and practice context to explore the potential for future evidence-based evaluation.

CHAPTER SEVEN

From Research to the Real World Part II - Bricks for the Bridge

Introduction

Chapter Two presented the CPAI-2010 as an instrument with which issues of implementation can be addressed in the real world and set out the need for this study with a view to establishing the importance of contextual transposition and translation to women's services and to the jurisdictions of the British Isles. In short, the study set out to determine how knowledge about effective practice is best packaged and applied when trying to improve practice in agencies. This chapter systematically addresses the research questions presented at the end of Chapter Two (p. 69) in order to fulfill this task.

What Processes of Evaluation are Meaningful to Services?

There are, of course, a multitude of different approaches to evaluation, ranging from paper processes overseen by off-site assessors who may never interact directly with evaluation participants (perhaps, in this instance, more correctly defined as 'subjects'), to fully participatory action research, undertaken *by* and *for* participants. Covington and Bloom (1999) advocate the action research approach of Appreciative Inquiry (AI) as a positive, strengths-based approach to evaluating women's projects, to "assist the community of participants to take an active role in guiding its own evolutionary transformation" (Doshier & Terry, 1993). This involves appreciating "the best of what *is*" before envisioning "what *might* be", dialoguing "what *should* be" and innovating "what *will* be" (Shear & Smith, 1999, p. 4; [My italics]).

Of their research in staff-prisoner interactions in a British prison setting, Liebling et al (1999) identified the possibilities for AI, both as an “instrument for change” (giving participants “a sense of being valued” and making “a significant difference to the way at least some prison officers conceived of their role”), and as a research method in the search for interpretive understanding (1999, p. 92). G. Robinson et al’s use of AI in British probation settings to explore ideas about ‘quality’ in practice was similarly successful, where participants were “empowered” by a process “which centred on voicing strengths, achievements and moments of pride” and “by virtue of its power to divert interviewees from negativity, [AI] maximized the extent to which participants were able to explore, focus on and share...their views” (2012, pp. 16-17).

As Covington and Bloom point out, AI “is a different approach than using quantifiable individual measures and creating an aggregate. When one creates an aggregate based on individual measures, the system (i.e. context) learns nothing about itself” (1999).

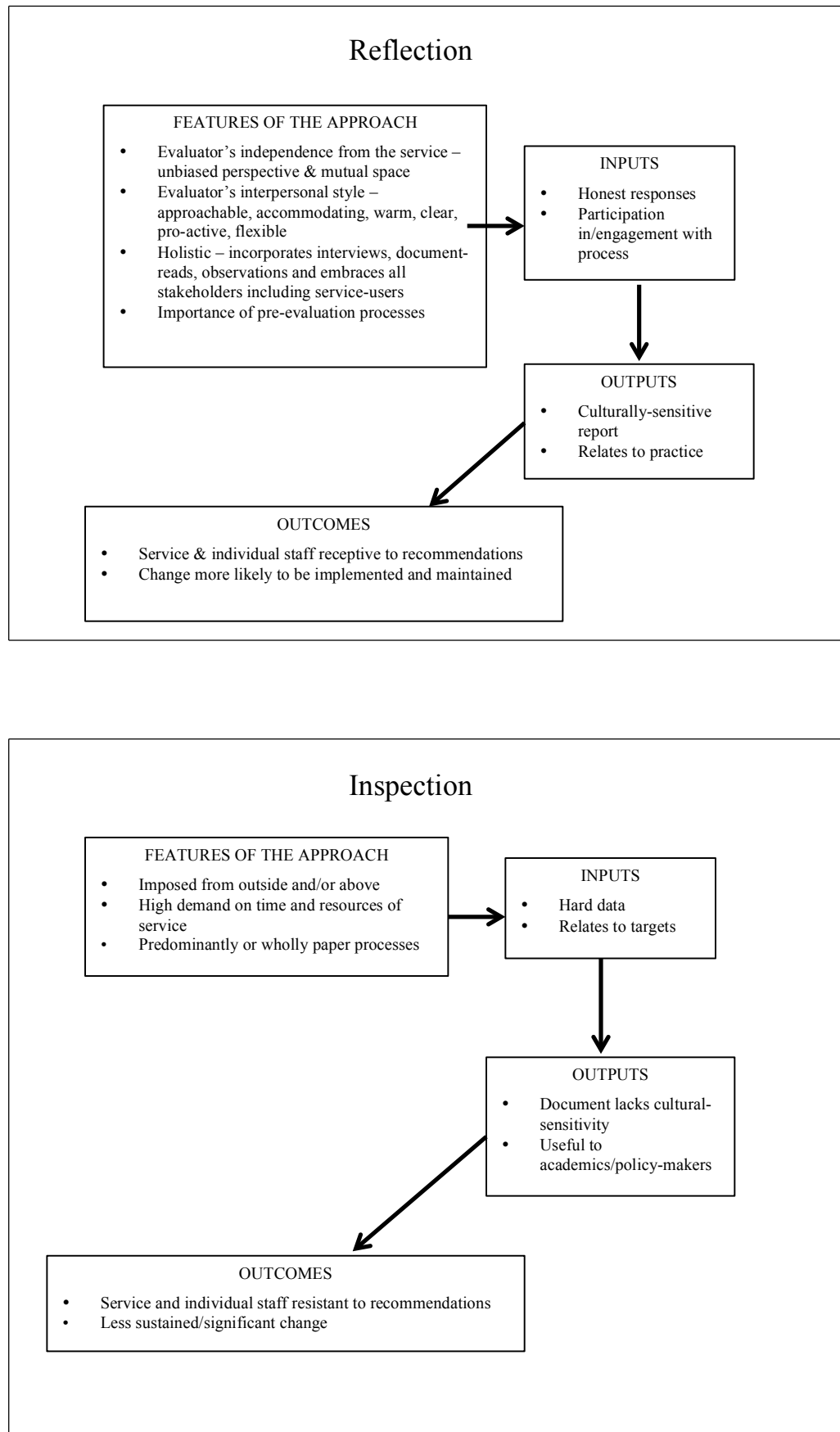
However, G. Robinson et al caution that “not all research questions can be framed...as ‘affirmative topics’”, and that the adoption of AI as a “principle mode of inquiry...is only advisable when the fit is a good one” (2012, p. 18). AI is essentially a generative approach (Bushe, 2007; Cooperrider & Srivastva, 1987) and is less useful where a pre-existing evidence base provides the framework for investigation. Furthermore, AI is not likely to be successful at uncovering unknown problems (Rogers & Fraser, 2003), this being a key aim of evaluation that seeks to determine whether services have been implemented as intended. It is worthy of note that Covington & Bloom’s later process evaluation instrument – the Gender-Responsive Program Assessment Tool (GRPA; 2008, revised 2017) – incorporates several items across seven elements of practice, each of which is scored on a five-point Likert scale from “No, not at all” to “Yes, functions very well”, according to how closely practice adheres to the “guiding principles” for gender-responsive practice (Bloom et al, 2003). Whilst it may be that creating an aggregate from individual quantifiable measures (such as changes in service-users’ recidivism or on pre- and post- tests of various individual scales) is of limited use in terms of contextual evaluation for organisational change, there is clearly room within a gender-responsive framework for quantitative assessment instruments by which “program administrators, program staff, program evaluators, and agency monitors

can...obtain feedback that can be used to improve the quality of a program's services" (Covington & Bloom, 2017). Nevertheless, the findings discussed in Chapter Five of this study support Covington and Bloom's concluding thoughts regarding the benefits of AI:

The system must reflect upon itself if it is going to self-actualize and create an environment that can create change. (Covington & Bloom, 1999).

In the study that is the focus of this thesis, a dichotomy of evaluation processes was identified by participants – that of engaging services in meaningful “reflection” against subjecting services to superficial “inspection” (see Fig. 7.1 below). The study's findings are supportive of the use of an instrument such as the CPAI-2010, which was perceived as enabling an inclusive, holistic approach to evaluation to engage services in a process of reflection that is the foundation for change through acceptance of recommendations. It is notable, though, that significant features of the approach identified by participants as facilitating reflection depend on evaluator competencies independent of the instrument itself. In a cross-comparison of evaluator competency taxonomies (*Essential Competencies for Program Evaluators* (Stevahn et al, 2005); *Competencies for Canadian Evaluation Practice* (Canadian Evaluation Society, 2010); and *Evaluator Competencies* (International Board of Standards for Training Performance & Instruction, 2006; Russ-Eft et al, 2008)), King and Stevahn (2015) identified interpersonal skills (effective and respectful interaction and communication) as an essential competency across taxonomies, alongside professional and reflective practice; the application of an appropriate methodology; situational analysis; and project management (2015, p. 27). Whilst the CPAI enables competent evaluation practice in terms of the professional, technical, situational and management focuses, effective interpersonal practice requires evaluators to possess expertise in relationship skills.

Fig. 7.1 Perceptions of Evaluation Dichotomy, “Reflection” vs “Inspection”



Are the Aims of the CPAI-2010 Appropriate to Services in the British Isles?

Participants in this study were strongly supportive of the need for evaluation in general, and recognised the value of evidence-based evaluation in particular. As Chapter Eight will explore, this demonstrates the resilience of the probation values of vocationalism and professionalism discussed in Chapter Three (Mawby & Worrall, 2013), given practitioners' often negative experiences of 'what works' initiatives in England and Wales historically (see Chapter Two, pp. 48-51).

The CPAI-2010 developers list eight purposes for the instrument (Gendreau & Thériault, 2010, p. 3) and the relevance of each in the context of the British Isles is explored below, with the addition of the ninth aim of encouraging re-assessment to reinforce positive practice change, since this has also emerged as a key strength of the CPAI (Latessa, 2004).

1. To Evaluate Funding Proposals

Though the potential to allocate resources better should be of interest to all commissioners and policy-makers (Latessa, 2018, p. 184), the potential for evidence-based evaluation of funding proposals is of particular value in the context of Transforming Rehabilitation (TR). TR's promise of innovative practice in reducing reoffending requires Community Rehabilitation Companies (CRCs) to commission services that are, by definition, not yet proven, particularly when working with minority offender groups such as women, for whom well-established rehabilitation services attracting specific funding are unlikely to cater. The House of Commons Committee of Public Accounts has noted that incentives to do this are low and that service delivery is likely to be shaped by CRCs' dependence on payment-by-results (PbR) outcomes (2016, p. 6). In the current study, the need to justify services in the TR climate was the impetus for evaluation in the women's project and the ability to evaluate the potential effectiveness of proposed services would be a clear advantage, reducing the risks of innovating in a PbR environment.

2. To Help Programs Articulate What They Do

Both the women's project and Jersey Probation and After-Care Service (JPACS) saw CPAI-2010 evaluation as an opportunity to identify elements of their practice that were consistent with the research. Moreover, the women's project sought positive representation in the turbulent environment of TR through explanation of its adherence to gender-responsive effective practices (in public launch and PR events, with criminal justice agencies, and with the service-users themselves), whilst JPACS sought positive representation at a time of increased scrutiny by the States of Jersey through articulation of its work in terms of effective practice under its current governance. Support for services is in part dependent upon stakeholder and public perceptions of effectiveness (Latessa, 2018, p. 185) and evaluation that enables services to articulate their work in terms of effective practice speaks to a key element of probation culture in the British Isles, that of "(re) presentation in the turbulent conditions of the external environment" (Mawby & Worrall, 2013, p. 145).

3. To Provide a Credible Rationale for Treatment

Of their work in evaluation, Van Voorhis et al opined, "before asking whether a program *does* work, we must ask whether there is any reason why it *should* work" (1995, p. 23). The need to provide a credible rationale for treatment was a concern for practitioners and managers alike in the current study. A sound evidence base was considered essential to ensure that the needs of service-users were properly addressed and that the services could justify themselves as value for public money. Put at its simplest, "[It's important] to know that...we're doing stuff that works..." (WPO2).

4. To Identify the Most Frequent Deficits

This study's findings support existing research regarding the most frequent deficits, such as underuse of the structuring elements of core-correctional practices in the British Isles (eg, Raynor et al, 2014; Ugwudike & Morgan, 2018). Within services, recommendations to address deficits were perceived by participants as the most useful element of the evaluation, providing an opportunity to address oversights and facilitate action and resource planning. In a wider sense this provides direction for further research and resource provision – for example, under-confidence and skills-deficits in the use of skill-building techniques (eg, role-play) indicate a need for the provision of specific training.

5. To Tabulate and Support what is Effective

Latessa's argument that scoring using the CPAI and related instruments "allows for comparison across programs [and] provides researchers with a numeric value that can be more easily disaggregated and analysed than a qualitative narrative" (2018, p. 185) is of patent value in the context of the British Isles, where there remains a paucity of research relating to practice integrity (Ugwudike & Morgan, 2018). The current study also reinforces the value to services of scoring to provide benchmark measures as a framework for interpreting progress, and for avoiding complacency in interpretation of, and responses to, evaluation findings.

6. To Evaluate External Service Contracts

It is worth noting that both JPACS and the women's project are operating in multi-agency contexts, each depending on elements of service being delivered externally or in partnership with external providers. Multi-agency working is of particular significance to gender-responsive approaches, which embrace a wide range of public, private and third-sector providers to offer holistic, women-centred community services. Additionally, a specific objective of TR reforms

was to open up the probation sector in England and Wales to a diverse range of rehabilitation providers (House of Commons Committee of Public Accounts, 2016). As with the evaluation of funding proposals, the capability of the CPAI-2010 to evaluate external service contracts is of value in the context of the British Isles, and particularly gender-responsive practice in the TR environment.

Interestingly, from a third-sector perspective, CPAI-2010 evaluation of a commissioning project was recognised as an opportunity to assess the potential for successful integration, cooperation and development.

7. To Provide Education for Staff in “What Works”

The need for staff engagement with education in ‘what works’ was a clear concern in this study. Practitioners in the women’s project lamented the loss of training provisions such as the Skills for Effective Engagement, Development and Supervision (SEEDS) programme under TR. In JPACS, where provision of a comprehensive programme of training in effective practices was linked to organisational harmony as well as superior supervision skills, the evaluation was welcomed as a refresher of staff education.

8. To Stimulate Relevant Research

This study found that CPAI-2010 evaluation has the potential to stimulate research at a micro-level, in terms of services’ responses to recommendations (“... we’re making enquiries about using other measures to assess responsivity.... We’ve talked about establishing a user-group and how we can do that...” (JCPO)), as well as at a macro-level where the instrument itself was found to be inadequate in gender-responsive services, stimulating further investigation and the development of the Swansea Service Evaluation Inventory – Women’s Projects (SSEI-W).

9. To Encourage Re-Evaluation

As discussed in Chapter Two (pp. 58-59), re-assessment using the CPAI can be embedded into institutional processes to reinforce positive practice change (Latessa, 2004, pp. 554-555) and repeated evaluation was considered by participants of the current study to be essential for the enablement of a process of continuing improvement for services (see Chapter Five, pp. 154-155). JPACS intends to commission CPAI re-evaluation of its community supervision services on a three-yearly cycle in order to maintain and/or improve on the results of its initial assessment, whilst annual re-evaluation was considered more appropriate in its less well-established services to monitor progress towards a Very Satisfactory categorisation. This potential for measuring distance travelled was identified as essential to meaningful evaluation of the women's project, but re-evaluation using the CPAI was undesirable for reasons relating primarily to its gender-neutral evidence base, and its lexicon.

Are the Items of the CPAI-2010 Appropriate to Services in the British Isles?

A significant finding of this study, discussed in Chapter Four, is that the items of the CPAI-2010 appear broadly relevant in the context of services in the British Isles. Factors identified as strengths or weaknesses in the international literature (eg, Latessa & Holsinger, 1998) were evident in evaluations of both the women's project and JPACS. The inclusion of items relating to organisational context was found to be especially important. Not only did the services achieve scores in domains focused on "organizational culture", "program implementation/maintenance" and "management/staff characteristics", reflecting both the stable and supportive management of JPACS and inadequate organisation at the women's project, but items in those domains were also found to be key across almost every other aspect of service, leading to the conclusion that an evaluation approach which looks solely at service-level practice is likely to be of little use in helping an organisation to identify and address the core issues.

Though a few items were perceived as unachievable in the context of the women's project due to restrictions on resources and organisational deficits under TR, only one item was rejected (by both services) as inappropriate in the context of services in the British Isles – the CPAI-2010's requirement for formal matching of practitioners to service-users according to, for example, assessment of conceptual level. In fact, the literature around effective therapeutic approaches (eg, Traux & Carkhuff, 1976) supports the services' arguments that, notwithstanding the need to respond to issues such as language needs and practitioner specialisms, etc, generic practitioners should be able to develop and maintain positive relationships with any service-user, mediated through strong relationship skills such as empathy, warmth and genuineness (these being evaluated in dimension G).

Whilst the vast majority of items of the CPAI-2010 were found to be relevant in both of the evaluated services, the gender-neutral evidence base fell short of the needs of women's services, where practitioners' efforts to respond to the distinct needs of women were at times frustrated by a knowledge deficit regarding effective practices. This knowledge base was recognised by participants to be incomplete and in flux, necessitating a dynamic approach to evaluation, both incorporating and adding to the expanding evidence base. Chapter Six discussed the need for development of a bespoke instrument for women's projects in the British Isles.

Is the Language of the CPAI-2010 Appropriate to Services in the British Isles?

Chapter Five dealt at length with issues arising from the language used by the CPAI-2010 (pp. 148-152), which was found to be inappropriate in the context of services in the British Isles, broadly falling into four main areas of concern: cultural, technical, ideological and motivational/reputational.

Table 7.1 CPAI-2010 Language Issues

CPAI-2010 language examples	Participant response examples
CULTURAL	
Treatment	"I don't do treatment! That's medical." (JPO2)
Program	"I'm thinking about programme as being group programmes and interventions rather than supervision." (JTL)
TECHNICAL	
"The [project's] intention to generate recidivism norms on its own population reflects..."	"I have to read it about three times and say, 'right, ok, what are you trying to say here because I don't understand..." (WPO3)
IDEOLOGICAL	
Correctional	"It's not very caring. It's quite superior and 'know-it-all'." (JPO2)
Punisher	"...implies physically rapping someone across the knuckles...or applying a curfew to them or something in a way that's not very helpful." (JCPO)
MOTIVATIONAL/REPUTATIONAL	
Unsatisfactory	"I wonder whether you could...have a...baseline measure which is less...prejudicial..." (JCPO)
Very Satisfactory	"...people wouldn't necessarily associate that with very good or best practice" (WPM)

Chapter Six discussed how these could be addressed in the development of the SSEI-W (pp. 175-178), though it would be possible to adopt some of these strategies without significantly deviating from the CPAI-2010 format. The word "program" might be replaced with "service" in interviews and reports, for example, without any significant impact on the integrity of the instrument. It is also possible that changes to the language of categorisation and subsections of the report could be trialled while at the same time remaining faithful to the CPAI-2010 domains and items. This might be achieved by avoiding the use of language perceived to be technical, punitive or stigmatising, simply rephrasing items in line with the examples given in Chapter Six (though permissions would need to be sought from the copyright holders). Given that JPACS has expressed a desire to continue using the CPAI-2010, this may be a useful line of enquiry. However, other barriers to the use of the CPAI-2010 should be taken into account when considering whether it may be practical to develop alternative instruments for use in the British Isles.

How Should Evaluation Findings be Presented?

The standard CPAI-2010 output of a report was found to be entirely appropriate to the evaluation commissioners. The norm for evaluations in the British Isles is that a report will be produced, and all agencies would expect this as a *sine qua non*. JPACS, for example, has a policy of publishing evaluation reports on its website. In this respect, the CPAI-2010 reports were perceived as appropriately “clear” and “useful...for practitioners” (WPM). However, service-users (considered to be key players in the evaluation process) were excluded from Phase Two of the study and so it is not possible to reach any conclusions regarding their needs, preferences and, indeed, interest in, accessing evaluation findings. Further research needs to be done specifically to engage service-users in discussion around this, including whether evaluation outcomes are best presented as, for example, a summary report, a presentation, or in the form of something more innovative for articulating research to diverse groups, such as a comic book (eg, Brimblecombe et al, 2014).

Additionally, dissemination was found to be a vital part of the evaluation process. The lack of service-user involvement in Phase Two of this study means that the potential for evaluation and dissemination to be “co-produced” (with the attendant benefits of providing “unique expertise to inform the development of a service” as well as acting as “positive involvement for someone engaged in a desistance journey, boosting confidence, developing skills and offering a vision for change” (Clinks, 2018, p. 6; See also Clinks, 2016)) was unexplored. What is resounding from the results of the study is that agencies are unable fully to benefit from evaluation without engaging their practitioners in reviewing findings and formulating plans. The provision of a report was not enough to stimulate this process alone. JPACS took the initiative to involve its practitioners in taking recommendations forward on an ongoing basis, which facilitated the participatory and supportive approach necessary to implement recommendations and ensure greater accountability for organisational leads through shared planning and agreed responsibilities. In the women’s project, however, practitioners were unable to participate in dissemination, planning and implementation; lacked confidence in organisational direction, resources and support; and expressed doubts about

accountability, ultimately perceiving the evaluation to be a futile exercise. Chapter Five concluded that considerable delay and partial redaction of the full report in the women's project reduced the potential for earlier and wider dissemination and discussion, and that this was due, at least in part, to the language with which findings were communicated in the report (pp. 161-162). Addressing the potentially stigmatising language of categorisation and placing the organisation on a continuum, taking account of distance travelled (as discussed in Chapter Six, pp. 176-178) would go some way towards enabling services to engage in participatory dissemination.

What Barriers Are There to Use of the CPAI-2010 in the British Isles and How Could These Be Overcome?

A definitive finding of this study is that the political contexts of organisations determine the potential usefulness of evaluation. Difficult political environments were the impetus *for* evaluation to promote and justify services, but also the impetus to *avoid* scrutiny of less well-performing services and suppress unfavourable findings. Upheaval under TR impacted upon the evaluation process itself in the women's project, where the timescale for CPAI-2010 evaluation had to be extended from the three- to five- day norm to six months to accommodate bidding issues, the dissolution of one partner service, delays, false starts, discrepancies over role designation, lack of premises, communication problems and participant availability, etc. These disruptions resulted in less comprehensive pre-evaluation engagement in one of the core agencies, which had a negative impact on practitioners' experiences of the evaluation. Clearly, the political context of services could be a barrier to the commission, implementation and dissemination of the CPAI-2010 evaluation, over and above services' abilities to respond to recommendations.

It is also noteworthy that the nature of different services may be inconducive to CPAI evaluation, regardless of their contemporaneous external environments. The relatively uncomplicated administration of the instrument in JPACS was in part because of JPACS's development according to the RNR evidence base – in other words, the CPAI was “a good fit” for the services it intended to provide (and the methods it employed to

provide them). JPACS is unusual in terms of its development, governance and relationship with the RNR evidence-base and so the potential appropriateness of the CPAI across other generic (ie, predominantly male) probation services across the other jurisdictions of the British Isles should not be overstated here. Equally, some of the challenges facing administration of the CPAI in the women's project were not related simply to struggles under TR, but relate to the nature of women's services as holistic in their attempts to engage with the social contexts as well as the personal capital of users. Whilst this study acknowledges the inability of the CPAI to capture desistance-informed and/or gender-responsive efforts as a limitation, it also endorses the holistic approach to evaluation afforded by the CPAI (see Fig. 7.1) in these contexts, since it is imperative to look beyond paper processes and hard targets to explore the quality of engagement at every level and with every aspect of a holistic service.

The cost of administering the CPAI-2010 is likely to be prohibitive for services with restricted resources. JPACS, which has expressed the intention to use the CPAI-2010 for future evaluations, is not representative of services in England and Wales, in that it has autonomy over evaluation and inspection processes. In the jurisdiction of the States of Jersey, the CPAI-2010 is perceived as inexpensive because it is considerably less costly than commissioning, for example, HMIP Inspection. The National Probation Service (NPS) and CRC services operating in England and Wales, however, are subject to obligatory HMIP inspections (it is unclear what direction Wales will take in this respect if justice is devolved) and are unlikely to have sufficient resources allocated for additional evaluations. Furthermore, to the best of this study's knowledge, there are only two accredited CPAI-2010 evaluators operating in the British Isles, which places clear limits on availability and increases associated costs.

Chapter Six proposed potential solutions to some of these issues, suggesting that using positive and motivational (rather than stigmatising) language and reflecting services' progress on their continuum of learning and change would make evaluation less politically volatile (pp. 176-178), while Chapter Five concluded that timing re-evaluation at an appropriate frequency and taking account of distance travelled would

help evaluation to address the wider context of services by reflecting the turbulent political environment (p. 162).

In the case of women's services, for which the items of the CPAI-2010 were inadequate, the development of a bespoke, gender-responsive instrument, the SSEI-W, also goes some way towards addressing the financial implications of the CPAI-2010 since no licence fee is associated with its use.

Can the CPAI-2010 be Developed Upon to Facilitate More Responsive Evaluation of Women's Services in the British Isles?

This question is not concerned with the *appropriateness* of the CPAI (or the RNR model upon which it is based) to women's services, since these matters have been attended to above. However, it is concerned rather with the CPAI's *adequacy* – ie, whether it is able fully to capture effective practice in integrated contexts, where desistance, gender-responsive and RNR efforts may be combined in the delivery of holistic services, both for offenders and non-offenders. As discussed in Chapter Two (pp. 67-69), desistance-informed and gender-responsive approaches aim to reach beyond personal rehabilitation to address the social contexts of women, and a fully integrative approach embraces evidence as to 'what works' to reduce recidivism alongside "theory and research on how and why the rehabilitative process works" (Maruna, 2015, p, 313).

This study has demonstrated that the CPAI-2010, with its foundations squarely in the RNR evidence-base, was unable fully to capture practices that engaged with the desistance of service-users in the women's project. For example, since the RNR framework is concerned with the targeting of higher-risk offenders and so does not specify the parameters for work with low-risk service-users, Domain F of the CPAI-2010, concerned with 'programme characteristics', was unable to respond to practices in the diversion scheme, which operated as part of the whole-systems approach indicated as effective for women (Holloway et al, 2017; Ministry of Justice, 2018c; 2018e). Evaluation of the full scope of the women's project's holistic approach was also limited

by the CPAI's narrow focus on advocacy/brokerage in Domain H, relating to inter-agency communication. Nevertheless, this study has revealed the value of the CPAI in the context of women's services, as an evidence-based evaluation that has the ability to incorporate every aspect of practice through interviews, review of documents, observation and the inclusion of all stakeholders (see Fig. 7.1).

Chapter Six discussed how this study's pilot of the CPAI-2010 has formed the basis for the development of a bespoke instrument that retains this holistic evaluation approach, whilst integrating the evidence-bases for RNR, gender-responsive and desistance-focused approaches in the British Isles, the Swansea Service Evaluation Instrument – Women's projects (SSEI-W).

Conclusion

Chapter Seven has addressed and answered the research questions relating to this study and signposts relevant materials for building the bridge between research and practice. Chapter Eight will consider the contribution made by the study, placing its findings in context and identifying areas for development.

CHAPTER EIGHT

Conclusion: The Study in Context

Introduction

This study's account of the application of the Correctional Program Assessment Inventory-2010 (CPAI-2010) in the British Isles has explored the limits of technology transfer between jurisdictional contexts.

The research includes analysis of evaluation results and participant feedback, and was the stimulus for development of a bespoke, gender-responsive instrument – the Swansea Service Evaluation Instrument-Women's Projects (SSEI-W). This conclusion will focus on the study's original contribution to the field by contextualising its findings in the existing literature and evaluating its methodology and limitations in order to identify directions for future work.

Context is Key

The context of this study, ie, the administration of the CPAI-2010 in adult services in the British Isles, is, to the author's best knowledge, a unique experiment. There are two key contextual issues to explore in this respect, each having a significant impact on the experiences of and potential for evaluation of services: the antecedent history of probation research in Britain; and the current approach to resourcing and organising probation in England and Wales.

Research and Evidence in British Probation

Raynor (2018a; 2018b) recounts a turbulent history of evaluative research in British probation, marked by periods of optimism and pessimism regarding both the value of research and evidence, and the potential for probation practice to have a positive impact

and reduce reoffending. For research to provide an evidence base for practice, he argues, there must be three “legs” of a social scientific approach: *understanding* usually requires qualitative investigation of “the aims of social actors, their beliefs about the processes they are involved in, and the meanings they attach to what they do and to what happens to them” (Raynor, 2018b, p. 60). Nevertheless, without *measurement* and *comparison*, the investigation, however worthwhile, cannot claim to be scientific:

[It] is not clear how social science can be *social* without an attempt to understand the meanings of social experience for the people involved. ... However, evaluation research in probation needs to go beyond [this] to develop more independent and objective ways to measure the impact of probation practice.... Here, we need to depend more on the quantitative procedures of measurement and comparison.... These are the scientific procedures that allow us to claim social investigation as a *science*, capable of generating reliable knowledge and building a cumulative knowledge base. Evaluation research...depends on getting these procedures right so that we can learn what works, how it works and in what circumstances, and how we might make it work better. Without understanding, we cannot get far, but without measurement and comparison it is difficult to turn understanding into evidence-based statements about the effectiveness of probation practice. (Raynor, 2018b, p. 60).

Framing the search for impact and evidence in this way helps better to understand some of the successes and failures of the research discussed in Chapters One and Two, and to contextualise practitioners’ and agencies’ responses to this study within their wider experiences of research and evidence-based endeavours. Arguably, one of the most unhelpful (and unnecessary) trends in probation research is a division between research that seeks to *understand* and research that seeks to *measure* and *compare*, exemplified by, though not limited to, the perceived qualitative/quantitative divide (see, eg, Bryman, 2012). Early attempts to evaluate rehabilitative services in Britain emerged from post-war social policy reform, with an emphasis on the production of “results in quantitative terms” (Home Office, 1958, p. 6). This emphasis on measurement and comparison has remained dominant in psychological research practices (Tafreshi et al, 2016), including those concerned with probation in Britain. As recently as 2014, the instruction for research applications to the then National Offender Management Service (now Her Majesty’s Prison and Probation Service) noted under “research prerequisites” that “research which focuses on outcomes through quantitative evidence of impacts is encouraged” (p. 9). In real terms, a research focus on measurement and comparison has often come at the expense of understanding (Raynor, 2018a; 2018b). Early efforts to evaluate probation work in Britain as well as the USA often found negligible effects

(eg, Wilkins, 1958; Folkard et al, 1976; Martinson, 1974), which findings contributed to the over-simplified but hugely politically influential notion that “nothing works”. However, as Raynor argues, these studies “tended to be relatively well executed with regard to measurement and comparison but they did not really examine, unpack and *understand* what social workers were actually doing” (2018b, p. 62; [My italics]). As discussed in Chapter One (pp. 10-12), this means that positive evidence of effective probation practice was obscured by inadequate research methodologies and the way was paved for increased punitive control in criminal justice (and a corresponding lack of governmental appetite for evaluative research (Raynor, 2018a; 2018b)).

Chapter One also discusses the significant breakthrough for probation evaluation made by meta-analyses (pp. 13-14) which were able to combine a robustly quantitative approach of measurement and comparison with some understanding of the “different inputs...being offered by a range of practitioners, and how well they were being delivered” (Raynor, 2018b, p. 64). North American researchers Andrews et al’s 1990 meta-analysis was particularly influential in drawing out factors that could improve recidivism rates, providing empirical evidence of the Risk, Need and Responsivity (RNR) principles that shaped thinking around probation practice in Britain as well as elsewhere. At the same time, British criminological research appears to have been suffering from a lack of international influence. In their 1990 citation analysis, Cohn and Farrington identified a scarcity of reference to British research in leading American criminology journals and books, and theorised that this was due in part to British criminologists valuing qualitative research (such as Cohen and Taylor’s ‘Psychological Survival’ (1981) or Cohen’s ‘Folk Devils and Moral Panics’ (1980)) whilst American criminologists mainly valued quantitative research (such as Hirschi’s ‘Causes of Delinquency’ (1969) or Elliott et al’s ‘Explaining Delinquency and Drug Use’ (1985)). This led the authors to conclude that “if British criminologists wish to influence their American counterparts, they should carry out high-quality quantitative research using the most sophisticated and up-to-date statistical techniques” (1990, p. 481). This is not to argue that measurement and comparison are more important than the in-depth understanding that may be gained through qualitative investigations. In quantitative research, understanding is *necessary* to make sense of what is being measured and the

significance of any comparison, whereas measurement and comparison are not pre-requisite to valid exploration of the motivations, intentions and beliefs of social actors. This means that qualitative research without measurement and comparison is, arguably, of far greater value than quantitative research that lacks understanding (this being, literally, meaningless). However, it seems likely that the under-representation of British criminology on the global stage meant that, whilst the RNR evidence base imported from the US and Canada was shown to be relevant to British services (McIvor, 1990; McGuire, 1995), British cultural issues were irrelevant in its development, and cultural context as an issue remains relatively under-explored (demonstrated in this study by, amongst other things, the culturally inappropriate language of “correctional programs”).

To compound these issues, the roll-out of RNR programmes in England and Wales was deeply flawed. As Chapter Two (pp. 48-51) discusses, a large-scale government initiative to pilot and evaluate RNR “Pathfinder” programmes in Britain was overly-ambitious with unrealistic targets, resulting in, amongst other deficits, unfeasible timescales and top-down management (Maguire, 2004; Raynor, 2004), failing to take into account Raynor & Vanstone’s earlier findings from a smaller mixed-methods (“three-legged”) evaluation of a Canadian RNR programme in Britain, which highlighted the importance of service-user and practitioner participation as well as the allocation of adequate time resources (Raynor & Vanstone, 1996; 1997; Raynor, 1998; 2018a; 2018b). Agencies and practitioners were alienated from the implementation process, put under pressure to meet unrealistic targets within unachievable timescales (the resulting shortcuts in assessment and targeting led to high attrition rates), and suffered from inter-related issues such as understaffing and poor communication between staff and agencies, only to face findings of null or limited effects (from evaluations with methodologies that had also been compromised by the rushed implementation). Under these circumstances, RNR (and, perhaps, quantitative research more generally) must have seemed like very bad science indeed. As Raynor (2004; 2018a; 2018b) argues, this is unfortunate, and the Pathfinder evaluations were, in fact, of considerable use to probation research. They indicated ways forward, produced some positive findings, and demonstrated the benefits of “three-legged” evaluation using, for example, programme manuals to clarify what practitioners *should* be doing, and video

recording of sessions to observe what they *are* doing. For Raynor, this was a crucial development in shaping probation research: “For the first time this gave researchers a clear grasp of what inputs were likely to be producing the measured outputs” (2018b, p. 65).

Chapter Two outlines how the ‘what works’ research focus has shifted to implementation and, crucially, practitioner skills – a key component of CPAI evaluation. Whilst this study shows that such a focus appears to be welcome (women’s project staff spoke highly of their experiences of the SEEDS programme (Skills for Effective Engagement, Development and Supervision: Rex & Hosking, 2013), and the JPACS team was continuing to use research instruments from the JS3 study (Jersey Supervision Skills Study: Raynor et al, 2014) for staff development), it is also the case that suspicion of and resentment towards ‘what works’ (at least in part resulting from experiences of poor implementation) has become part of the probation landscape in Britain. This has contributed to unnecessary conflict with alternative fields of enquiry such as desistance research, which has much to offer *understanding* through a focus on service-users’ own narratives of offending and desistance but does not (nor does it seek to) provide instruction for probation practitioners (Raynor 2018a; 2018b). As laid out in Chapter Two, this study proposes an integrative approach and goes on in Chapter Six to make some suggestions as to how to achieve this in evaluation instruments and approaches by, in the first instance, ameliorating the language with which evaluation is administered and communicated. It is the element of *understanding* that has historically been under-valued in positivist endeavours, the ‘what works’ enterprise included. However, that is not to underestimate the value of measurement and comparison. As Raynor points out “three is usually the minimum number of legs needed to support a stable structure” (2018b, p. 60).

Raynor’s 2018 outline of the different phases of British probation research usefully organises the issues explored in Chapters One and Two of this study: Pessimistic findings quelled post-war optimism in practice; the quest for ‘what works’ provided a framework for practice (RNR); and failings in implementation resulted in a shift in research focus to integrity (first to programme integrity and, most recently, to

practitioner skills, or “treatment integrity” (McGuire, 2004)), this being the phase of research within which this study is located. Raynor’s outline goes on to locate contemporary probation in “post-truth Britain” (Raynor, 2018a; 2018b), raising questions regarding the usefulness of research and evidence in a political climate within which the destruction of Probation Trusts and widespread privatisation of services was unapologetically evidence-free. The then Justice Secretary, Chris Grayling stated, “I don’t believe you need to pilot professional and operational freedom” (2014, cited in Raynor, 2018b, p. 70) and Michael Gove, in his role as Grayling’s successor (though referring, in this case, to a lack of evidence for Brexit), declared that “people in this country have had enough of experts” (Mance, 2016).

It has long been the case that the scope for public criminology (ie, that which calls for research-derived evidence to influence public understanding and criminal justice policy and practice) has been dictated by political inclination, depending on its subject matter and/or timing:

Research-derived evidence influences policy and practice in some places, at some times, and on some subjects. On some subjects, especially those that implicate important normative and ideological concerns, such as capital punishment or severity of sentences, evidence is seldom likely to make much difference. On other subjects, especially those that can be characterized as primarily technological or technocratic (e.g., the crime-prevention efforts of ignition locks), evidence is often likely to be influential. For subjects falling in the great middle, evidence might or might not matter depending on the political and policy weight of considerations that point in other directions. The presence or absence of windows of opportunity in specific times and places makes all the difference. (Tonry, 2010, pp. 793-794).

In Britain, Raynor notes the different status given to evidence in a changing political climate over even a relatively short period of time. From the use of evidence regarding ‘what works’ in the shaping of probation practice to the commissioning of research to support structural reform (such as the 2003 Carter review which was used to establish NOMS as part of a wider shift towards managerialism in public services (Easton & Piper, 2016, p. 246)), and to evidence demonstrating the impact of providing after-care services (Disley et al, 2015; Pearce et al, 2015) being used (erroneously) to argue for the success of Payment by Results (PbR) in criminal justice, the 21st century to date “shows a progression from being *guided* by evidence to using evidence as a *resource to support*

policy decisions already being made, to *creating* evidence to support policies, and eventually to dispensing with evidence altogether” (Raynor, 2018b, pp. 69-70).

In 2010 Tonry argued for a place for public criminology, even where there is little hope of tangible influence:

Even concerning the relatively impervious issues like capital punishment and sentencing severity, evidence sharpens the debate and clarifies what the issues really are. It is important to show and know, for example, that no credible empirical evidence suggests that capital punishment is an effective deterrent of homicide. If the evidence is absent or unclear, then people who support the death penalty for ideological reasons, but feel uncomfortable saying so, can hide behind claims that they support capital punishment as a means to saving innocent victims’ lives. If the evidence clearly shows that death does not deter, then the debate must be made in the moral terms that really motivate opponents and supporters. Likewise, if the evidence shows that California’s three strikes law has no discernible effects on crime patterns or trends, its defenders must be prepared to defend the savage sentences it requires in moral terms. That is a good thing. (Tonry, 2010, p. 794).

However, the place for evaluation research and, indeed, public criminology as a whole, is less clear in a brazenly hostile political context in which expertise is actively denigrated (Raynor, 2018a; 2018b). The current post-truth climate has offered refuge from being forced into transparent discussion of moral motivation, exemplified by US Counselor to President Trump, Kellyanne Conway’s use of “alternative facts” to describe demonstrably inaccurate statements made by then Whitehouse Press Secretary, Sean Spicer (NBC, 2017). In such a climate, that innocent victims are saved by the death penalty, for example, may simply be an “alternative fact” and evidence to the contrary may be dismissed as “fake news”.

For Raynor, the task for evaluation research in the contemporary context of England and Wales may simply be to survive and influence reform following the inevitable failure of evidence-free policies:

The emerging style of politics (perhaps not so much post-truth as post-Enlightenment) does not provide a promising environment for evaluation research or evidence-based policy in those countries where it is prevalent. However, it is not prevalent everywhere: probation research is flourishing in Europe, and in some of the devolved jurisdictions within the British Isles such as Scotland, Northern Ireland and the Channel Islands. Even within England and Wales there are examples of local research that engages directly with service providers. When post-truth policies fail, factual research on how to make probation more effective will be needed to support the necessary evidence-based reform. In the meantime,

in Britain, we face continuing uncertainty about how best to finance our research, where to place it and, in particular, how to communicate it effectively to people who may put it to some use. (Raynor, 2018b, p. 71).

From the outset, this study strove to take the “quiet iterative” approach suggested by Maguire (2004, p. 214) as a way to pursue evidence-based policy in the long-term. Its pragmatic (and, crucially, “three-legged”) methodology has produced results that indicate hope for the future, even in a post-truth climate, since a positive desire for evidence (the “culture of curiosity” identified by Raynor and Vanstone (2001) as necessary to drive organisational change) was immediately apparent at a practice level in the attitudes of staff, who were driven to *work from* as well as *contribute to* the evidence base (though this co-existed with a fear of research because of managerial excesses in the past, and the commercial market context in the present). This study demonstrates that, in order to sustain evaluation as a resource (for change at a practice level as well as for the development of the evidence base for possible future reconstruction of probation services), it will need to respond to and reflect the political context of agencies, integrate all relevant approaches and engage all key social actors in a process of reflection (rather than inspection – see Fig. 7.1).

A Market for Research?

In the turbulent political climates of this study, commissioners perceived (positive) evaluation as a tool with which to promote and sustain their services, and for JPACS there remains a reasonable possibility that evidence of success can achieve these ends. As Tonry (2010) pointed out, the influence of evidence on policy-makers depends, in the end, on the weight of any political, ideological and/or bureaucratic obstacles specific to place and/or time, and the States of Jersey appear, thus far, to have avoided what one Jersey magistrate described as the “political machinations” of England and Wales. The women’s project, however, was operating in Wales under Transforming Rehabilitation (TR) – an evidence-free enterprise within a “post-truth” culture. In this environment there is very little to suggest that policy-makers will respond to evidence rationally.

From its inception, TR was based on a media-friendly over-simplification of the “problem” of reoffending and lacked evidence for the PbR “solution” or its associated processes (Phillips, 2014). It is, then, unsurprising that Grayling’s stated intention for the “rehabilitation revolution” to “stimulate innovation and open the delivery of services to a wider range of providers with the skills needed to change an individual’s behaviour and reduce reoffending in future” (House of Commons debate, 2012, c.691) has not been realised. Instead, inspections of Community Rehabilitation Companies (CRCs) have overwhelmingly found that organisational deficits under TR (including understaffing and unmanageable caseloads, target-driven (rather than quality-driven) audits, and delays in the provision of essential IT systems), have severely compromised the delivery of services and that Courts and National Probation Service (NPS) staff may be reluctant to commission services from CRCs. As a result, voluntary sector involvement, long depended upon for the delivery of specialist services locally (including those for women) has diminished rather than flourished under TR:

...changes in sentencing and the nature of work coming to CRCs have seriously affected their income and indeed their commercial viability, causing them to curtail or change their transformation plans.... Having started with enthusiasm, many CRCs are now not commissioning the full range of specialist services that are needed to make a difference for people with particular problems.... With no single body responsible for the stewardship of valued specialist services (such as bespoke services for women), these are likely to continue to wane. (HM Inspectorate of Probation, 2017, pp. 6-12).

The impact on third-sector providers is clear in the current study, whereby one partner of the women’s project was dissolved during the data-collection period and another shortly after.

Based on the experiences of the inspectorate, Chief Inspector of Probation, Dame Glenys Stacey, questions the sustainability (and rectitude) of a system that neglects the relationship between service-users and services where face-to-face supervision contacts are “the heart of probation” (Toronjo & Taxman, 2018):

Funding aside, I question whether the current model for probation can deliver sufficiently well. Above all, a close, forthcoming and productive relationship between an individual and their probation worker is key. This is where skilled probation staff add most value.... Yet in some CRCs, individuals meet with their probation worker in places that lack privacy, where sensitive and difficult conversations must take place. Some do not meet with their probation

worker face-to-face. Instead, they are supervised by telephone calls every six weeks or so from junior professional staff carrying 200 cases or more.

I find it inexplicable that, under the banner of innovation, these developments were allowed. And I regret that the current national delivery model does not have at its heart the effective, joined-up local partnership work and other specialist services so much needed, for many who offend. (HM Inspectorate of Probation, 2017, pp. 6-7).

Though inspections have found the NPS to be performing relatively well (HM Inspectorate of Probation, 2017), issues around the imposition of TR are not limited to the CRCs. The participants across both core agencies in the women's project identified organisational deficits under TR as problems, and there is emerging evidence that practitioners in the NPS are suffering from the "high pressure and relentless work" of primarily high-risk caseloads, indicating a need for managerial support to prioritise and tier cases (in keeping with the Risk principle) (Phillips et al, 2016, p. 189). Practitioners may be disappointed by the limited impact that evaluation can have in this respect. This study found that practitioners perceived evaluation as a means of providing the foundations for managerial direction and support (which was seen to be lacking in the TR environment of the women's project) but that political and organisational instability reduced managerial capabilities to engage practitioners in moving forward with evaluation recommendations.

The full impact of PbR on probation outcomes is not yet clear. Findings regarding reoffending are mixed, incomplete and inconclusive (Raynor, 2017), though the latest figures from the Ministry of Justice (2018a) show that ten of the 21 CRCs are failing to hit targets regarding the proportion of service-users who reoffend (figures regarding the frequency rate of reoffending for the same cohort will not be available until January 2019). Far from the promise that PbR would incentivise probation practice, the performance of the CRC contracts to date follows a pattern familiar to critics of the privatisation of public services:

...when essential services and statutory duties are outsourced in arrangements supported by long contracts with private sector providers, the contractual arrangements seldom provide enough leverage to manage performance in a service which cannot be allowed to fail, so further public costs are incurred to keep providers afloat. (Raynor, 2017 p. 2).

On examining the accounts for government contracts for CRCs, the House of Commons Committee of Public Accounts released a report that was a damning indictment of what Andrew Neilson, director of campaigns at the Howard League, refers to as “a service where companies are failing and being paid to fail” (Savage, 2018).

More than three years into its seven year contracts with Community Rehabilitation Companies (CRCs), the Ministry is a long way from delivering the ‘rehabilitation revolution’ it had promised. In 2017, the Ministry was forced to adjust its contracts with CRCs because it had pushed through its reforms too quickly and failed to anticipate foreseeable consequences. The volumes of work paid for under the contracts has dramatically reduced, meaning that CRCs have not invested in probation services. The quality of rehabilitation services has suffered as a result and is undermining the objectives of the reforms.

The Ministry accepts that the CRC contracts were plainly not working as intended and has agreed to pay them up to £342 million more of taxpayers’ money. But the Ministry could not explain what it is getting back for this extra commitment. Despite this bailout, 14 out of 21 CRCs are still forecasting losses....

The future income of CRCs will increasingly depend on their ability to reduce stubbornly high reoffending rates. While there has been a modest decrease in the number of reoffenders since CRCs were introduced, 19 CRCs have not met their targets for reducing the frequency of reoffending. (House of Commons Committee of Public Accounts, 2018, p. 3).

Neilson goes on to conclude that the rhetoric of privatisation as delivering reductions in reoffending was little more than “a smokescreen for something that was about cutting costs and taking money out of the system” (Savage, 2018). It is certainly the case that evaluation commissioners seeking to use evidence of success to secure their position in a TR climate may well be disappointed. For example, an evaluation of the London Probation Trust’s engagement worker role, involving the recruitment of ex-offenders, found that the role was valued by staff and service-users, met the need for support workers to undertake several time-intensive tasks with service-users and had the potential to enhance service-user engagement. Regardless, the role was discontinued when the London CRC reviewed resourcing levels due to a financial shortfall under PbR (Hosking & Rico, 2018).

In the TR environment, where everything is a market, it is difficult to see how useful a commodity evidence might be. Based on the good health of probation research in other jurisdictions, Raynor (2018c) suggests that the political devolution of criminal justice to Wales (a motion for which was backed by the National Assembly for Wales in March

2018) may provide opportunities to “manage evidence-based reconstruction of probation services” through, eg, restored judicial involvement, a focus on practice skills, the evaluation of “innovations designed to be evaluated” and, crucially, the reversal of privatisation. Welsh policy has taken a positive step in this direction, with the announcement in June 2018 that the supervision of all offenders would be returned to a single National Probation Service (Office of the Secretary of State for Wales, 2018). However, plans for England are for the continued implementation of TR, albeit subject to a number of revisions including the early termination of CRC contracts, the alignment of CRC and NPS areas, and the launch of a new procurement exercise with an emphasis on the involvement of the voluntary sector in the delivery of services (Gauke, 2018). Meanwhile and in any case, the political climate may have changed (and, as the antecedent history of probation research in Britain teaches us, will change again), but service-users remain in need and the public at risk. This demands a response from academic interests in evaluation and effectiveness.

Happily, for those involved in the pursuit of effective practice in probation, this study demonstrates that practitioners and agencies remain committed to the concept of evidence-based practice, even in a context where “reforms [are] largely about saving money and about cost-effectiveness [and run] counter to Corston’s proposals for a more individualised, intensive (if needed), needs-based approach more suited to women” (Annison et al, 2015, p. 23). Testament to the resilience of the probation values of vocationalism and professionalism (Mawby & Worrall, 2013), staff in the women’s project sought evidence-based evaluation to provide a framework upon which to build effective and justifiable services for women. *This is the market for research to which evaluation must now respond – the practice-level desire for evidence.*

This Study’s Contribution to Context

This study was produced at a time when TR in England and Wales has renewed pressure on agencies to deliver effective services in the open market under PbR. Because context is the central feature of the work, the intention was not to validate the CPAI-2010’s scoring (using comparison with recidivism figures, as has been the case in

most research focused on its applicability across settings), but to examine the relationship between culture and context, and experience of evaluation and relevance of findings. Inter-jurisdictional application exposed turbulent political environments, such as TR, which had far-reaching implications for the potential of evaluation. What can be surmised from the findings of this study is that adopting motivational language (notably in categorisation), placing a service on a continuum and timing evaluation at an appropriate frequency are all factors that can reflect and ameliorate problems within turbulent political environments. In any event, this study emphasised that evaluation of organisational context is key to meaningful engagement with the core issues of a service, indicating the inadequacy of approaches that focus solely on practice and/or outcomes.

This study has consolidated arguments for “future-oriented, optimistic and approach goal focused” language to create an environment of positive growth and participation (Ward & Maruna, 2007, p. 127), since language that was perceived as stigmatising or judgmental emerged as a barrier at every level from commission of CPAI evaluation to dissemination of the report. Reference to the field of educational psychology assists in using language to inculcate a growth mindset (Dweck, 2017). Transitioning the language towards a more positive register also has the potential to unify ostensibly conflicted approaches, notably RNR and desistance – services that tend to be more ideologically aligned with strengths-based approaches still need to be able to articulate work in terms of effective practice in the criminal justice environment of the British Isles, and the value of a “three-legged” research approach cannot be underestimated.

As discussed in Chapter Two, notions of what constitutes effective practice, or what evidence of effectiveness is meaningful, are nevertheless hotly contested (pp. 67-69). This thesis joins the call (made by, amongst others, Herzog-Evans, 2018, and Maruna, 2015) for the integration of all the available evidence to develop and support practices that adhere to the RNR principles to reduce reoffending whilst supporting desistance through social reintegration. In this respect, the limitations of the CPAI were most clear in the women’s project, where its narrow focus on RNR meant that it was unable adequately to capture promising desistance and gender-responsive practices such as

work to divert women from criminalisation (Chapter Four, pp. 125-126; 136-137; 142). This inadequacy has particular ramifications in the TR environment whose neoliberal philosophy (core to which is the responsabilisation of individuals; Raynor, 2018d, p. 341) provides infertile ground for services wishing to address the social contexts of service-users. As JPACS's exceptional achievements suggest, there is considerable potential for services to develop effective practices through direct engagement with evidence, independent of policy developments. According to Heath and Raynor "[it] could be argued that the ponderous rate of legislative change in Jersey, has led to JPACS being a service which looks for practice based solutions to problems rather than assuming that legislation will provide the answer" (2014, p. 7). Process evaluation could, and should, offer the crucial opportunity to do this, but evaluation of the women's project in Wales using the CPAI failed to speak to the strengths of desistance-related work and so offered little direction in the directionless milieu of TR as well as potentially suppressing innovative practice with notions of failure. The need to develop a bespoke instrument for women's projects is clear.

Innovation in Evaluation

The innovation of this study is the development of an evidence-based evaluation approach that meets the needs of services, specifically women's projects, in the British Isles.

Gender-responsive services encounter particular obstacles when seeking to evaluate effectiveness (eg, low-risk service-user groups and the integration of a broad range of providers in the delivery of holistic services, etc), and the inadequacies of blunt measures and inflexible evaluation processes are clear (see, eg, Jolliffe et al, 2011). However, the need for demonstrably effective community services to divert women from harmful incarceration remains of vital importance while women are proportionately more likely to serve short sentences (ie, of less than 12 months) which increase their risk of reoffending and cause other, unmeasurable, damages (Ministry of Justice, 2017b; Prison Reform Trust, 2017; Baldwin & Epstein, 2017). Recently-released Ministry of Justice figures have revealed that one in four women sentenced to

prison in 2016 served sentences of 30 days or less, largely for non-violent offences such as non-payment of council tax or television licence (Merrick, 2017). Raynor (2017) points out that a decline in the use of community sentences by the Courts under TR is likely due, at least in part, to “a lack of confidence in, or a lack of knowledge about what CRCs are providing” (p. 2) and, reflecting a prevalent concern amongst the judiciary, a District Judge participating in HMIP inspection observed:

My concern is neither the NPS staff nor I have enough information about what the CRC are doing with offenders when they have been sentenced. I am not confident that the right work or intervention is being delivered or that it is being done swiftly after sentence. (HM Inspectorate of Probation, 2017, p. 48).

Perhaps without intention, Chairman of the Magistrates Association, John Bache, alludes to the fundamental need for evaluation: “...the problem that we have is if there isn’t a suitable alternative...we have to send them to custody.... We would like to see better provision of services in the community, particularly for women, so that we can avoid custodial sentences...” (Pollock & Glyn, 2018).

As Van Voorhis points out, women offenders have been overlooked by the majority of criminal justice research, resulting in a dearth of evidence to which policy makers, magistrates and practitioners alike can refer, though she cautions that women’s relative “invisibility to science” should not be used to ignore or deny the evidence that *has* been accrued (Van Voorhis, 2012, p. 132). The SSEI-W aims to provide evidence accessible to all parties, in addition to contributing to the evidence base itself.

Furthermore, in the context of Britain and TR, there is a clear and pressing need for the SSEI-W, which it is hoped will prove a powerful instrument in an increasingly sophisticated culture of evaluation, as called for by Sutherland, to address the deficit of evidence for women’s services:

In the context of the Transforming Rehabilitation agenda and ‘Payment by Results’, community organizations are increasingly at the forefront of criminal justice. The shift in recent years – driven by economic circumstances and a more discerning audience of commissioners – is the need to provide evidence of effectiveness. Some have embraced this and proactively sought to work towards an evidence base. Others have been pushed into *ex post* scrabbling for ‘proof’ that what they have been doing – often for decades without challenge – ‘works’ in some way. The mismatch between the ability of organizations to

provide evidence of success and the exacting standards imposed on them by government, researchers and a host of self-styled ‘evidence champions’ will continue to cause difficulties unless we can build a ‘culture of evaluation’ underpinned by a nuanced understanding of programme limitations and likely impacts. The risk of not doing so is that the evidence vacuum is filled with conjecture, opinion and unsubstantiated claims of effectiveness. (Sutherland, 2017, p.12).

The innovation of the SSEI-W is its nuanced response to the needs of women’s projects, their potential and limitations. Unlike its antecedents²⁰, it has been developed *from the ground up* to respond to the culture of women's services in the British Isles and evidence from the RNR, desistance and gender-responsive literatures.

Limitations of this Study

The British Isles is the target criminal justice “jurisdiction” explored here, though it is important to add that this does not actually constitute a single jurisdiction. At the time of the study, the sample services were operating in Wales and Jersey, these being very different contexts. The Transforming Rehabilitation (TR) agenda for England and Wales provided the somewhat hostile environment for the women’s project in Wales, whilst Jersey’s social-work-oriented Probation and After Care Service (JPACS) was operating in the self-governing British Crown dependency, relatively autonomously of political developments.

This study has not included probation services in other jurisdictions of the British Isles which, as Robinson and McNeill (2010) point out, are quite different to each other in a number of ways, with the social work orientations of Scotland and Northern Ireland providing especially stark contrast to the operation of probation in England and Wales as “law enforcement” (pp. 743-747). Within each of these jurisdictions, too, there are different probation contexts and cultures such as, for example, youth justice and women’s services, which, as the current study demonstrates, are likely to have their own

²⁰ The innovation of this study is not the use of RNR evaluation in a women’s project, nor the development of a gender-responsive evaluation instrument: both the CPAI and the Correctional Program Checklist (CPC – adapted to retain items of the CPAI with strongest correlation to effective practices) have been used in women’s services internationally (Latessa, 2012); the Gender-Responsive Program Assessment Tool (GRPA) developed by Covington and Bloom has been used alongside the CPAI to investigate adherence to the principles both of gender-responsivity and effective RNR practice in women’s services in Canada (Booth, 2012); and Patricia Van Voorhis is currently developing the Gender-Responsive Correctional Program Assessment Inventory (Salisbury, 2015). Though Van Voorhis’s study bears obvious comparison to the development of the SSEI-W, the very fact that her inventory’s title retains the language of “correctional” practices indicates its jurisdictional difference.

distinct evaluation needs. In this way, the study is clearly limited by its relatively small sample.

Although both participant services appear broadly representative of probation in England and Wales, as conceptualised by Mawby and Worrall (2013) and discussed in Chapter Three (pp. 97-102), and the findings of this study are supportive of the literature, suggesting that one can draw some reliable conclusions from the data, the usefulness and limitations of the CPAI have not been measured across all the different jurisdictions of the British Isles, nor their constituent probation services. In particular, the participant services' focus on the welfare and social inclusion of their users may be peculiar to their social-work-based (JPACS) or gender-responsive (women's project) organisational contexts, and in contrast to punitive practices in generic (ie, gender-neutral and, thus, predominantly male) services across England and Wales (Robinson & McNeill, 2010, p. 752). Nevertheless, the study does have something to contribute in terms of technology transfer from the US and Canada to the British Isles (perceived as a particularly significant route for policy transfer; Jones & Newburn, 2007). Its thesis indicates generalised features of evaluation responsiveness such as the components of reflective evaluation (see Fig 7.1), the importance of a "growth mindset" in evaluation, and the incorporation of organisational context as a focus for evaluation.

It was not within the scope of this study quantitatively to validate use of the CPAI-2010 in the British Isles due to the small sample size and lack of complete data relating to recidivism in the women's project. It is reasonable to say that, with only two services evaluated and a lack of recidivism data, *comparison* is the weakest "leg" of this study. Whilst the SSEI-W has been developed to respond to its cultural context, it, too, has not yet been validated. Given the difficulties with access and the administration of the CPAI-2010 under TR resulting in a significantly extended evaluation period and obstacles to dissemination, the possibilities for evaluation research were restricted. Factors disrupting the evaluation process included issues with role designation and referrals as well as complete dissolution of one third-sector agency with key responsibilities. Hough's 2016 paper demonstrates the sometimes-insurmountable obstacles that TR can present – in that case, a lack of referrals resulted in no evaluation

participants at all. It is therefore a considerable strength of this study that the evaluation process was flexible enough to accommodate the various issues that arose in the women's project. Whether it would be feasible to replicate this in different services, however, is not answerable within the confines of this study.

Furthermore, the exclusion of service-users from the participant group for Phase Two is a significant limitation. The potential for evaluation and dissemination to be "co-produced" (p. 192) remains unexplored. More fundamentally, attempts to engage with investigation of issues around how and why people stop offending lack true *understanding* without the perspectives of service-users themselves. Staff identified service-users as equal, if not the most important, stakeholders in evaluation, and their exclusion from Phase Two can only be remedied by further research.

Future Work

In future pilots of the SSEI-W and CPAI-2010 in the British Isles, service-users should have the opportunity to engage in the investigation of their experiences of evaluation and their preferences regarding how findings are presented. Since the literature around service-user involvement recommends the co-production of evaluation (Clinks 2016; 2018), this may be a barrier to the use of expert-directed, evidence-based evaluation with instruments such as the CPAI and/or may provide direction for the development of evaluation approaches that better meet the needs of service-users *and* staff.

It would be beneficial to pilot changes in language in future evaluations using the CPAI-2010 in the British Isles. This would have to be done through consultation with the instruments' developers to overcome copyright issues. It would also present some other significant challenges. This study exposed how organisational instability under TR resulted in a turbulent research experience with difficulties in access, timing, participant motivation and dissemination, amongst other factors. In a real-world setting (outside of the comparative luxury of doctoral research), this would make the completion of evaluation within budgetary and time limits extremely difficult, if not insurmountable. Additionally, the appetite for evaluative research is likely to be reduced in the TR

environment, bound as it is by PbR and fraught with competing priorities. As discussed above, the experience of RNR research in the British Isles has been tainted by, amongst other things, poor implementation, and so it is unclear whether even the non-TR environments in Scotland and Ireland, etc, would welcome use of the CPAI, particularly given their social-work orientations and the lack of attention given to desistance by the instrument. Although JPACS was receptive and enthusiastic about future opportunities for the CPAI, further consideration needs to be given as to whether its use is a realistic proposition for other adult services in Britain, especially given the associated costs.

The study also identified some interesting avenues for enquiry outside of the evaluation remit. Firstly, it appears that there is a signal deficit in practitioner use of role-play to support service-users' skill-building. Further research to identify the way to address this may be fruitful – reference to arts practitioners/academics as well as the fields of psychology and motivational training would enrich this investigation.

The development of the SSEI-W also introduces further opportunities for research to fulfill its potential:

- To define features of effective practice with women, this study has relied on the terminology of gender-responsive literature, though this has exposed definitional limitations, especially in the contemporary context of transgender rights, where questions should be raised as to what constitutes a “safe space” for women (generally “women-only”). Reference to gender studies to define a “safe space” would involve exploring what constitutes “safety” and what that means for women and transgender users.
- Reference to the field of linguistics could examine whether incorporating the language of two contrasting approaches (RNR and desistance) in a single tool can be functionally unifying.
- There is potential for the SSEI-W to be developed as software to enhance its usability. Further research will need to be done with agencies and practitioners

to assess the desirability of this venture and reference to software technology would determine practical considerations.

Thus, it is anticipated that this research will benefit in the future from contributions by, at least, software technology, arts practitioners/academics, psychologists, motivational trainers, gender studies and linguistics.

The most significant piece of work arising from this study is the need to pilot the newly-developed SSEI-W in women's services, with feedback sought on participant experience and usefulness, and validation of the tool through comparison of SSEI-W scores and recidivism data for assessed services. Preliminary pilots of the SSEI-W are planned for women's projects in three Welsh probation areas in 2019. Given the dissolution of TR, there is now the opportunity for evidence-based restructuring of criminal justice, with use of the SSEI-W having the potential to shape the development both of women's services and evaluation approaches in Wales.

Concluding Thoughts

This study has revealed the weaknesses and considerable strengths of the CPAI-2010 when applied in the British Isles. Hypotheses that problems might arise around language difference have been borne out by the findings, as has a need for the inclusion of gender-responsive items for the evaluation of women's services. However, the instrument's holistic approach and use of an unparalleled evidence base for effective practice were shown to be crucial components of meaningful evaluation.

This study has produced a bespoke, gender-responsive instrument for women's projects in the British Isles that addresses issues around cultural context at the same time incorporating RNR, desistance and gender-responsive literatures.

Evaluation under TR may not have the ability to achieve all that commissioners *want* it to (eg, the assurance of improved political status for services based on success) and it may not be able to achieve the highest ideals for public criminology in terms of influence at a policy-level and on public understanding, but this study suggests that

working *with* services and their users to develop sustainable (ie, relevant, palatable, inexpensive and reflexive) evaluation is the only way in which to survive the onslaught against evidence in “post-truth” Britain. On the other hand, Wales is, at the time of writing, positioning itself to repair the damage of TR – and for the restructuring of probation services to be meaningful and effective, evidence must be at its heart. The potential for evaluation research in this context is clear and crucial, if not irrefutable.

It is expected that the SSEI-W will become an eminent addition to the suite of evaluation instruments available in practice, especially at this time, when short-term imprisonment is overused (and disproportionately harmful) for women, and services need to be able to demonstrate their effectiveness to provide a genuine alternative.

INDEX TO APPENDICES

Appendix 1	CPAI-2010 Staff Member Survey
Appendix 2	CPAI-2010 Request for Information Form
Appendix 3	CPAI-2010 Invitation to Practitioners
Appendix 4	CPAI-2010 Invitation to Service-Users
Appendix 5	Feedback Interview Schedule
Appendix 6	CPAI-2010 Pre-Visit Letter
Appendix 7	CPAI-2010 Checklist of Materials
Appendix 8	Information for Research Participants
Appendix 9	Research Participant Consent Form
Appendix 10	CPAI-2010 Report Conducted on JPACS
Appendix 11	CPAI-2010 Report Conducted on a Women's Project in Wales
Appendix 12	Research Participant Consent to Audio Recording
Appendix 13	Swansea Service Evaluation Inventory – Women's Projects V1.4

APPENDIX 1

CPAI-2010 Staff Member Survey

Staff Member Survey

Name:_____

Programme Name:_____

Location:_____

Position/Title:_____

How long have you worked for this programme?_____

Educational Degree(s):_____
(e.g., B.A., M.Sc etc. & where gained)

Area of Degree(s):_____
(e.g., Criminal Justice, Social Work, etc.)

Certifications:_____
(e.g., Chemical Dependency, Licensed Social Worker, etc.)

Before coming to this programme have you worked for other
treatment programmes with offenders? Yes_____ No_____

If yes, please provide the following information for the three most
recent positions:

Name of the programme:_____

Position:_____

Years/Months on job:_____

Name of the programme:_____

Position:_____

Years/Months on job:_____

Name of the programme:_____

Position:_____

Years/Months on job:_____

APPENDIX 2

CPAI-2010 Request for Information Form

CPAI-2010 REQUEST FOR INFORMATION

It would be helpful for the CPAI evaluator to have the following information before the on-site visit.

1. Name of programme

Name specific to programme

2. Name of two contact persons

#1 Name _____
Title _____
Address _____
Phone _____
Fax _____
Email _____

#2 Name _____
Title _____
Address _____
Phone _____
Fax _____
Email _____

3. Programme setting

Is the programme located in a:

- a) Community non-residential/parole/probation setting
- b) Community residential care
- c) Institutional setting
 - i. Minimum security
 - ii. Medium security
 - iii. Maximum security

Describe the general structure and physical set-up of the programme setting.

4. Programme sponsor

Record the type of agency/organisation that sponsors the programme (e.g., criminal justice, mental health or other public agency, private for profit/non-profit, university-based research group, etc.)

5. Funding source

How is funding allocated to the programme?

6. Number of offenders

	Current	Capacity	%
#Juvenile	<hr/>	<hr/>	<hr/>
#Adult	<hr/>	<hr/>	<hr/>
% Male/Female	<hr/>		

7. Number of staff

Full time staff members are those who work at least 32 hours per week unless otherwise specified.

	<i>Number</i>	<i>%Male/Female</i>
<i>Full-time</i>	_____	_____
<i>Part-time</i>	_____	_____

8. Programme budget

What is the programme budget? _____

9. Referral source

Please describe how the participants are referred to the programme.

10. Presenting problem

When participants initially begin the programme, what types of problems do you most often encounter (e.g. substance misuse, sexual offending, history of violence, anger problems, anti-social values, etc.)?

Presenting problems	%
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

APPENDIX 3

CPAI-2010 Invitation to Practitioners

Information to staff pre-visit – invitation to participate

Throughout the week commencing _____ a comprehensive and structured evaluation of _____ will be carried out using the Correctional Programme Assessment Inventory-2010 (CPAI-2010). The CPAI-2010 was designed to assess intervention programmes in relation to the research evidence on effective practice. The goal of this evaluation is to provide the organisation with more knowledge on a variety of areas, including how to assist staff with interventions that will address the needs of probationers, address incidents as they happen, prevent incidents during service engagement, prevent future incidence of crime once discharged, provide staff training to meet those needs etc.

The CPAI-2010 is widely used in Canada and the US and is currently being piloted in the British Isles by the Swansea Service Evaluation Team (SSET). The evaluation of _____ will be led by Bridget Kerr who is currently undertaking her PhD research at Swansea University as part of SSET. The team is seeking the input of practitioners and probationers regarding the appropriateness and usefulness of service evaluation here and will be collecting data relating to your experience of evaluation to assist them in developing or adapting the CPAI-2010.

The evaluation process includes interviews with probationers, practitioners and the management team as well as observations of intervention sessions and case management meetings. During this process a number of documents will be reviewed such as a sample of open and closed files, manuals, and any documentation that will help Bridget to understand how staff and probationers interact, intervene, provide services, etc.

Staff interviews will take 30 minutes to one hour. If lengthier interviews are necessary, staff/managers will be notified in advance as much as possible.

Once the evaluation is complete Bridget will submit a full report to _____ that details the evaluation findings and gives proposals for development. You will then be given the opportunity to participate in the second phase in which you will be asked for your experience of evaluation, opinions regarding the tool and its implementation and ideas for improvement.

I have identified you as a potential participant for this study but your participation is entirely voluntary. Bridget will be onsite and available to give more information on the evaluation and research study, answer any questions that you may have, discuss consent, observe practice and interview you regarding your role in the week commencing _____. If you would like any further information before then, you can contact her at Bridget.Kerr@swansea.ac.uk

APPENDIX 4

CPAI-2010 Invitation to Service-Users

Message to Probationers Regarding CPAI-2010

Instructions: Staff should read this to Probationers prior to evaluation.

We have a visitor to our service over the next week, Bridget Kerr. She's here to find out more about the supervision that we provide and how we provide it to you in order to help us plan for future work.

Bridget may ask to join some of your sessions. She will sit in the back of the room and take notes. She will be writing notes *about the session, not* about you.

Bridget may also ask to interview you about your experiences of supervision.

Your participation in this evaluation is voluntary. If you choose to participate in the interview, she will not use your name or personal information.

She is interested in your ideas and feedback to improve services.

APPENDIX 5

Feedback Interview Schedule

INTERVIEW SCHEDULE

Suggested Questions for Semi-Structured Interview with Staff Participants

- How do you see your role in the evaluation process?
- How did you find the evaluation process overall?
- Was there anything that wasn't looked at that you feel should have been/anything that was looked at that you feel shouldn't have been?
- What might have made the evaluation process better/worse?
- Is the report useful? If so, how will it be used? If not, why not?
- Is there anything that you would like to change about the report?
- What are the best/worst things about the report?
- Do you think evaluation is necessary/important? Why/why not?
- What do you think is most important when evaluating intervention services?
- Is there anything else you would like to add?

APPENDIX 6

CPAI-2010 Pre-Visit Letter

Dear _____,

In preparation for the forthcoming on-site visit to evaluate _____, I have attached some important information for you to review. The first document is a survey that I would like to be completed by all your staff members who regularly deliver treatment services to probationers. Please do not include security, clerical or maintenance staff members. The second document is a checklist of the materials that I will need to review. The third document is a suggested memo to be sent to staff outlining the evaluation and inviting them to participate. The fourth document is to be read with probationers to prepare them for my visit.

I will be undertaking the CPAI-2010 evaluation during the week commencing _____. The following will be required on site:

Accommodation

- I require a work space from which I can easily come and go, leave personal belongings, review files etc.
- Private space is needed in which to conduct interviews with staff and probationers – if appropriate, this can be my work space.

Interviews

- Interviews with the senior management team (_____) can be scheduled as soon as possible. These may take approximately one hour.
- Interviews need to be arranged with a sample of practitioners and these will last approximately 30 minutes to 1 hour. I have included a timetable to be completed by your team. Please allocate an hour to interview slots.
- Probationers will also be interviewed where appropriate and this can either be pre-scheduled or arranged onsite.
- Please note that participant consent will be sought and agreed to prior to every interview and that this may lengthen the interview time slightly and/or impact on the numbers involved in evaluation.

Observation

- I will observe the most current _____ videos and will need access to these and appropriate AV facilities on-site.
- I will observe a sample of live _____ sessions and/or any other treatment services (eg, group programmes) provided under the _____ remit. Please schedule these allowing adequate time for participant consent to be sought and agreed to prior to observation taking place.

Files

- I will require access to a random sample of active and inactive electronic and paper files relating to probationers.

Feedback

- I will collect initial feedback about practitioners' experiences of the evaluation at the end of the week. This can be arranged onsite.

Please ensure that the staff member surveys are distributed, completed and returned to me on _____, when I will commence my onsite work.

My onsite hours during the evaluation period will generally be _____, though I will be available to undertake tasks outside of these hours whenever necessary (for example, observation of group-work sessions). It would be good if I could have a copy of the preliminary timetable at least a week before I arrive on-site.

In the meantime, please don't hesitate to contact me with any questions. My details are below.

Yours sincerely,

Bridget

Bridget Kerr
Centre for Criminal Justice and Criminology
School of Law
Swansea University
Singleton Park
Swansea SA2 8PP

Tel: _____

Email: Bridget.Kerr@swansea.ac.uk

APPENDIX 7

CPAI-2010 Checklist of Materials

CHECKLIST OF MATERIALS FOR THE CPAI-2010 VISIT

In preparation for the upcoming CPAI-2010 site visit, please have available the following information for review:

- ✓ Description of the programme and the treatment it provides
- ✓ Any assessment instruments being used by the programme (e.g., risk and need assessments, drug and alcohol assessments, psycho-social structured interviews, etc.) and scoring guides for each
- ✓ Any validation studies conducted within the past five years on the assessment instruments currently in use
- ✓ Written documentation of the ethical guidelines followed by staff
- ✓ Daily schedule for offenders
- ✓ All forms used to evaluate staff performance
- ✓ All forms used to evaluate offender performance in the programme
- ✓ Client satisfaction forms
- ✓ Any data collection instruments that have been used for evaluation purposes
- ✓ Copies of any evaluation studies conducted on the programme within the last five years
- ✓ Several active and inactive case files (selected at random)
- ✓ List of reinforcers and punishers used by the programme
- ✓ Any documentation assessing the need for the programme

APPENDIX 8

Information for Research Participants

Information sheet for participants

This PhD study is part of a wider project to pilot the use of an intervention service evaluation tool, the Correctional Program Assessment Inventory [CPAI-2010]¹, in criminal justice agencies in the British Isles.

There is a great deal of evidence regarding which methods and practices are successful in reducing re-offending² and service evaluation measures the extent to which practice adheres to these evidence-based principles. Most importantly, service evaluation gives agencies valuable guidance towards ensuring that all aspects of their service are effective.

The CPAI-2010 was developed in North America and is used widely there, where a number of studies have found that intervention services with the highest (best) CPAI scores have greater measured effect in terms of reducing re-offending.

Each intervention service is scored in areas relating to the agency's practices as well as its organisation and management, based on data from multiple sources:

- Interviews with management, probation practitioners and client participants.
- Documents including risk assessment tools, case records, programme manuals and policy and procedure documents. .
- Direct observation of staff and clients in programme and supervision sessions.
- A survey completed by staff members.

Based at Swansea University, Wales, the Swansea Service Evaluation Team (SSET) is committed to developing evaluation instruments and approaches for the British Isles and is undertaking a number of evaluations of different intervention services in probation and youth offending agencies. A report detailing the outcome of each evaluation will be provided to the agency to assist them in refining and further developing their interventions.

For the purposes of this study, feedback will also be sought from management, probation practitioners and service-user participants to draw out issues related to their experience of the evaluation which will guide the re-design and development of the instruments and approaches used by SSET.

¹ Gendreau, P., Andrews, D. A. & Thériault, Y. (2010). *Correctional Program Assessment Inventory – 2010 CPAI-2010*. Ottawa.

² Andrews, D. A. & Bonta, J. (2010). *The Psychology of Criminal Conduct, 5th edition*. New Providence, NJ: Anderson.

Use of Data

Data will be collected through a variety of methods – interview, document-reads, survey and observation. Evaluation findings will initially be written in note form and/or audio-recorded where permission is given, and will then be typed up into a report that will be supplied to the agency. Key practitioners (for example _____ etc.) will be easily identifiable from their role descriptions and so will be subject to limited anonymity in the report. However, no individuals will be named in the report and, for the majority of practitioners and all service-user participants, characteristics and opinions will be generalised and not attributed to any one individual. Some aspects of the evaluation report, eg, the scoring, will be included in Bridget Kerr's PhD study, dissertation and related SSET articles and may be subject to analysis and comparison with evaluations from other intervention programmes and reconviction figures.

The second phase of field-work involves interviews in which participants will have the opportunity to reflect on their thoughts and feelings regarding the evaluation tool, process and report findings and for their concerns and ideas to contribute to future development of the tool. Again, data will be hand-written in note form during interview and/or audio-recorded with permission, before being typed up at a later date. This data will be used by SSET to guide any re-design of the tool and relevant findings will also be incorporated into the PhD study. Participants may be quoted in the PhD dissertation and related articles, though they will not be named at any point and efforts will be made to anonymise the area in which they work to reduce the risk of their being identified. Quotes will only be used if participants give their express permission for this.

Manual files (e.g. hand written notes) and handheld recording devices will be kept in a locked filing cabinet and carried by the researcher at all times whilst on agency premises or in transit. Computer (including MP3 audio) files will be encrypted in a data vault on a password protected laptop. No personal details (names, dates of birth etc.) will be included in the CPAI report, PhD dissertation or any report, presentation or other aspect of the study at any time. All case records and documents containing personal details belonging to _____ will remain on _____ premises and subject to _____ data protection procedures at all times.

Findings from this study will be presented both nationally and internationally at conferences and in journal articles and may be used to inform policy making at a local and governmental level.

Whilst every effort will be made to respect the confidentiality of information supplied by research participants, action will be taken to manage risk through _____ procedures where researchers are made party to information indicating a risk to others or to the participant themselves.

APPENDIX 9

Research Participant Consent Form



Research Participant Consent Form

I have read and/or had explained to me the information sheet relating to this study regarding intervention service evaluation. I have had the opportunity to ask questions relating to the study and use of data and have had these answered satisfactorily. I agree to participate in the study which may involve all or some of the following:

My participation in interview.

Observation of me engaging in activities and practice relevant to this study.

My completion of a survey.

I understand that my participation is entirely voluntary, I have the right to say no and I can choose not to answer specific questions or to withdraw from the study or stop participating at any time. This will not impact on my work with

_____.

My signature below means that I voluntarily agree to participate in this research study.

Signature

Date

Direct quotations taken from my interview responses may be used in documents, reports, articles and/or presentations related to this study. I understand that quotations will be used anonymously and every effort made to reduce the risk of individuals being identified. I understand that I can participate in the study without giving my consent to use of direct quotations.

I do/do not give my consent to the use of direct quotations (delete as applicable).

Signature

Date

APPENDIX 10

CPAI-2010 Report Conducted on JPACS



**CORRECTIONAL PROGRAM ASSESSMENT INVENTORY – 2010
(CPAI-2010)©**

**Conducted on Jersey Probation and Aftercare Service's Community
Supervision service**

**Report prepared by CPAI-2010 accredited evaluator Bridget Kerr,
Swansea Service Evaluation Team (SSET)
Bridget.Kerr@Swansea.ac.uk**

TABLE OF CONTENTS

Executive Summary.....	p.3
Introduction.....	p.4
Description of the Correctional Program Assessment Inventory – 2010 (CPAI-2010).....	p.4
Methodology.....	p.6
Results.....	p.8
A. Programme demographic.....	p.8
B. Organisational culture.....	p.10
C. Programme implementation/maintenance.....	p.12
D. Management/staff characteristics.....	p.14
E. Client risk/need practices.....	p.16
F. Programme characteristics.....	p.18
G. Core correctional practice.....	p.22
H. Inter-agency communication.....	p.24
I. Evaluation.....	p.25
Scoring Summary.....	p.27
Conclusions.....	p.27
References.....	p.29

EXECUTIVE SUMMARY

In September 2016 Jersey Probation and Aftercare Service (JPACS) was the subject of an evaluation to assess the extent to which its community supervision services adhere to principles of effective practice. The aims of this study were congruent with the Service's intentions to reduce re-offending and produce reports to assist with policy formation. The CPAI-2010 (Gendreau, Andrews & Thériault, 2010) was used to assess nine domains of service delivery (eight of which are scored). The evaluation identified JPACS as having a number of strengths, achieving the highest possible rating of very satisfactory in every domain relating to effective practice and 100% scores in domains relating to the characteristics of management and staff, inter-agency communication and evaluation practices. This indicates that JPACS is currently well placed to deliver effective intervention services that reduce the reoffending of clients and, as a strong, professionalised service, to continue to develop these in response to the recommendations outlined in this report.

INTRODUCTION

This evaluation was undertaken as part of Swansea Service Evaluation Team's (SSET) project piloting use of the CPAI-2010 in the British Isles. Jersey represents a very different jurisdictional context from the services in Wales that have thus far participated in evaluation as part of this pilot. Autonomous from the legal systems of England and Wales and France, the States of Jersey legislate for the island and JPACS is under the governance of a Probation Board consisting of five Jurats (lay elected judges). Probation Officers, or Deleagués, have a number of responsibilities across the Courts and States of Jersey, including the management of Probation Orders (a form of provisional release comparable to the former Probation Orders of England and Wales).

Independent from strategic and political developments in probation in England and Wales, JPACS has developed longstanding relationships with academics, researchers and professionals in the criminological field, which has shaped the development of the Service. This includes the delivery of community supervision for offenders sentenced to Probation Orders, as well as offenders released from custody. It is this aspect of service that was evaluated using the CPAI-2010.

DESCRIPTION OF THE CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI-2010)

The CPAI-2010 (Gendreau, Andrews & Thériault, 2010) was designed to measure how closely offender intervention services adhere to the principles of effective

practice that have been developed based on meta-analytical research (Andrews & Bonta, 2010). The CPAI-2010 has been used extensively in the United States and Canada where studies have found that services with the highest degree of integrity have greater measured effect in terms of reducing re-offending (Andrews & Dowden, 2005, Gray, 1997, Holsinger, 1999, Nesovic, 2003 & Lowenkamp, 2004).

Using interviews with practitioners and offenders, review of agency documents (including case files) and direct observation of agency activity and core correctional practice in supervision and programme sessions, CPAI-trained evaluators undertake a detailed assessment of intervention services. The nine domains of the CPAI-2010 include 143 items, 133 of which are scored 'yes', 'no' or 'not applicable' according to whether different aspects of practice adhere to the principles of effective practice (Gendreau, French and Taylor, 2002):

A. Programme demographics	10 items (none scored)
B. Organisational culture	9 items
C. Programme implementation/maintenance	10 items
D. Management/staff characteristics	18 items
E. Client risk/need practices	13 items
F. Programme characteristics	25 items
G. Core correctional practice	45 items
H. Inter-agency communication	5 items
I. Evaluation	8 items

To enhance reliability and reduce subjective scoring, there is a confidence rating for each scored item that requires the scorer to indicate the reliability of their assessment on a five-point scale according to the evidence available to support the score. An overall score is calculated by collating the scores in each domain and dividing the sum of these scores by the total number of applicable items. The overall score is attributed a classification of programme integrity. These range from 'very

satisfactory' (70% and above) to 'unsatisfactory' (below 50%). The average confidence scores for each section are used to ascertain the reliability and validity of the evaluation's conclusions. The results of the CPAI-2010 evaluation are given in the form of an in-depth report, which includes the scoring summary, overall classification, qualitative evidence and recommendations. This enables services to articulate the work they do in terms of evidence based practice, identify the strengths and weaknesses of programmes and develop strategies to improve their service.

METHODOLOGY

An initial orientation meeting regarding this evaluation was held on 10th May 2016 to provide an overview of the CPAI-2010 to JPACS's Chief Probation Officer, Assistant Chief Probation Officer and Probation Team Leader. A timescale for the evaluation was agreed following this meeting and on-site data collection took place during the week commencing 12th September 2016, though key interviews with the Chief Probation Officer and Probation Team Leader were undertaken over the telephone on 2nd August and 23rd August 2016 to accommodate conflicting schedules. On-site data collection included observation of group sessions of the Adapt Domestic Abuse Prevention Training (ADAPT) and Emotional Coping Skills programmes as well as live and video-recorded supervision sessions. Interviews were also conducted with the Assistant Chief Probation Officer; five probation officers; research and information officer; practitioners in substance misuse, restorative justice, etc; six service-users and one magistrate.

MATERIALS REVIEWED

- Adapt Domestic Abuse Prevention Training programme manual
- Emotional Coping Skills programme delivery materials
- Aggression Control Training manual
- Core programme manual
- Focus on People – Effect Change training materials
- Jersey Supervision Interview Checklist
- Observing Interview Skills: a manual for users of the Jersey Supervision Interview Checklist
- Journal articles including The Jersey Supervision Skills Study: outcomes and reconvictions (Raynor et al, 2012); Skills and Training in British Probation: A tale of neglect and possible revival (Raynor & Ugwudike, 2013); The Impact of Skills in Probation Work: A reconviction study (Raynor et al, 2014) and Moving Away from Social Work and Half Way Back Again: New research on skills in probation (Raynor & Vanstone, 2015).
- Jersey Probation and After-Care Service Annual Report for 2015 and Business Plan for 2016
- Community Sentences and their Outcomes in Jersey: the fourth report
- Summary of Three-quarter Year Stats, September 2015
- Probation Information leaflet
- Online resources at probation.je
- Probation Client Feedback leaflets
- Probation Client Feedback Questionnaire Results 2015
- Level of Service Inventory-Revised
- Risk Matrix 2000
- Spousal Abuse Risk Assessment
- HCR20
- Policy Book
- Code of Practice
- Probation Standards document
- Referral forms
- Psychometric Testing reports for the Sex Offender Treatment Programme
- Performance Review and Appraisal documents

RESULTS

A. PROGRAMME DEMOGRAPHICS

This section is descriptive only and does not contribute to the scoring of the CPAI-2010.

JPACS is co-located with the Magistrates' Court in a purpose-built building, central to the town of St Helier, and provides a number of services across the courts and States of Jersey including representation in Parish Enquiries, provision of a community service scheme, probation, restorative justice, prisoner through-care and the preparation of Social Enquiry Reports. Under the auspices of the courts, JPACS's remit extends to the Jersey Family Court Advisory Service (JFCAS), which has the stated intention of looking after the interests of children involved in family court proceedings. Most recently, Independent Domestic Violence Advisors (IDVAs) have moved from their previous location in Police HQ to the JPACS building. The differing needs, risks and vulnerabilities of clients in these different services are addressed through careful design including separate, secure entrances and rooms for JFCAS, viewing windows in doors and location of IDVA interview rooms in the Magistrates' Court building, etc.

This evaluation focuses exclusively on probation services offered by JPACS, and specifically the provision of community supervision in the management of Probation Orders and prisoner aftercare. These services are largely delivered in the JPACS building, which is easily accessible from the city centre, though elements, such as group programmes, are sometimes delivered in other local venues. Accommodation in

the JPACS building is arranged over three floors. The reception and waiting area are on the ground floor with access by elevator and stairs to the first and second floors, where practitioners' offices are used for supervision sessions with the majority of clients. There are also meeting rooms which are used for group-work sessions and staff meetings. All rooms are comfortable, light and clean and the JFCAS facilities can be used should clients have particular vulnerabilities or needs. These include child-friendly rooms with toys.

JPACS is funded by the States of Jersey on an annual basis and the current financial settlement has been agreed until 2019. Some elements of service are delivered in partnership with community and third-sector agencies through arrangements such as the Building a Safer Society strategy (BASS). The ADAPT programme, for example, is delivered by the Jersey Domestic Violence Forum under BASS. JPACS also works closely with Guernsey Probation Service to reduce costs and maximise efficiency relating to the development of ICT systems, accessing and providing training, etc.

At the time of this evaluation, the number of offenders served by JPACS totalled 390 adults and 26 youths, 91% of whom were male. The predominant presenting problems amongst those receiving community supervision were alcohol and substance misuse, aggression and emotional instability, relationship issues and, to a lesser extent, financial issues, deficits in employment, negative peer groups and health-related issues.

B. ORGANISATIONAL CULTURE

Strengths:

JPACS's goals and mandate are clearly laid out in the Annual Report and Business Plan. The report is reviewed by the Probation Board, ensuring accountability, and is also made available to the public through JPACS's website, reflecting an embedded organisational commitment to transparency. The website, probation.je, is accessible, comprehensive and offers a number of resources including a summary of the Service's remit and the specific standards and practice issues relating to probation supervision.

There is a strong managerial commitment to the development of initiatives to improve services and, where opportunities have been identified, JPACS has responded within a reasonable period of time. For example, interventions were put in place to address the emergent issue of internet sexual offences in consultation with a specialist forensic psychologist and the Jersey Supervision Interview Checklist has been embedded into the clinical supervision of probation officers following recommendations from the Jersey Supervision Skills Study (JS3).

There are clear procedures for responding to problems within the team, which are routinely dealt with using a non-confrontational, non-crisis approach. In general, problems are dealt with through discussion and mediation, enabled by a supportive and hands-on managerial approach by all members of senior staff, whose offices share corridors with practitioners, operating an open-door policy. In interview, all staff were aware of the formal procedures for grievance and disciplinaries, etc, and relevant documents are readily available, though at the time of the evaluation there had been

no formal grievance for a number of years. There is a collegiate atmosphere in staff meetings and day-to-day practices. A number of practitioners identified the support, knowledge and experience of their colleagues as crucial resources in their daily work. They pointed out that, whilst staff may have philosophical differences, they are united by their common objective to, in the words of one, “do our best rehabilitating people back into the community”. To facilitate constructive communication, information is frequently shared throughout JPACS in meetings and through the use of email, workshops, reports and the Service’s electronic policy book.

High levels of organisational harmony are reflected in extremely high levels of staff retention, with 100% of staff surveyed having remained with JPACS for more than two years. In fact, over 70% of staff members involved in the evaluation had worked in the Service for over ten years.

In its self-evaluation and research involvement practices, Jersey has gained international recognition as an exemplar of learning organisations. This evaluation found evidence to support this reputation. The team takes a conscientious and methodical approach to the collection and recording of data, enabling the Service to undertake internal quantitative assessments of different areas of service delivery, largely facilitated by the research and information officer working in collaboration with external researchers. Additionally, JPACS makes use of a number of eminent external consultants to advise on several aspects of service delivery.

Areas for improvement:

Those members of staff who are registered social workers are governed by the Health & Care Professions Council (HCPC) standards of conduct, performance and ethics. However, the standards relating to probation supervision make no provision for ethical conduct and there is no documented code of ethics specific to JPACS, but for a brief statement in the Annual Report that “In all its work the Probation and After Care Service promotes respect and dignity for all.” Whilst this statement is undoubtedly worthy, it lacks the detail and description of responsibilities that would make it functional as policy.

Recommendations:

1. A document detailing the ethics of intervention would consolidate the shared vision expressed by practitioners and management and would provide guidance and accountability for all staff working within, or in partnership with, JPACS in the delivery of services. Staff should be trained to ensure that the ethical code is inculcated in every aspect of services.

Rating: 8/9 (89%) Very Satisfactory CR: 4.6

C. PROGRAMME IMPLEMENTATION/MAINTENANCE**Strengths:**

JPACS’s service is congruent with the values and practices of the courts. Whilst they describe conflicting perspectives at times regarding individual cases or practice issues, interviews with both the magistrate and JPACS managers and practitioners revealed mutual respect and a shared philosophy regarding the rehabilitation of

offenders. The benefits of co-location in enabling a high level of responsiveness to the needs and wishes of the Court and meeting the demands of case management during court procedures were clear throughout the evaluation period, as practitioners and magistrates are able to cross to each other's buildings and discuss issues *ad hoc*. JPACS's governance by a Probation Board consisting of five Jurats ensures that the interests of the courts are considered in all aspects of service delivery and that the Service is accountable to its stakeholders in terms of its fiscal responsibility. The fact that JPACS is governed by its "principle customer" was repeatedly identified in interview as responsible for the Service's success.

Stakeholder investment and stability are key to providing effective services. The current Chief Probation Officer has been instrumental in establishing JPACS, which, due to its autonomy, has developed a culturally-responsive range of services based on consultation with community agencies and statutory bodies in Jersey as well as on review of the international literature relating to effective practices. The staffing levels of probation officers have been prioritised and made resilient to recent funding cuts and the current annual budget has been agreed until 2019. The professional credentials of staff are also maintained and enhanced through continuing professional development practices.

Areas for improvement:

Perhaps the greatest issue currently facing JPACS is a proposal that the Service should be brought under the jurisdiction of Jersey's Community and Constitutional Affairs Department. Evidence shows that it is detrimental to the effectiveness of services when they are faced by far-reaching, contentious issues, and it is clear that

JPACS's development and administration of services is inextricably linked to its governance by the courts. The proposed change would, at the very least, result in upheaval and disruption for an extended period. It is difficult to predict what features of its services JPACS may retain or lose in a move away from the courts.

To ensure that JPACS continues to develop services that meet the needs of the community and to evidence this need for its services, it may benefit from taking a more structured approach to identifying gaps in service provision and/or emerging issues through the use of, for example, focus groups and/or surveys.

Recommendations:

1. The use of focus groups and/or surveys would provide JPACS with on-going documentation of the need for its services.
2. It is a recommendation of this evaluation that JPACS remains under its current governance structure to maintain the organisational stability necessary to develop and deliver effective services.

Rating: 8/10 (80%) Very Satisfactory CR: 4.3

D. MANAGEMENT/STAFF CHARACTERISTICS

Strengths:

JPACS has a committed, professionalised, highly-qualified and experienced staff team. All staff members delivering interventions have relevant undergraduate degrees and years of experience in probation and related fields. A large percentage of staff, including the Chief and Assistant Chief Probation Officers, have relevant advanced

postgraduate qualification. All staff are Disclosure and Barring Service checked and are selected for employment based on assessment of their relationship skills and skills specific to evidence-based practice (eg, pro-social modelling) using vignettes.

JPACS provides a comprehensive programme of training to its staff, with training in the theory and practice of effective interventions being delivered two-three times annually by a consultant forensic psychologist. Additional training is sourced through JPACS's research relationships with criminologists in Swansea University and the Cognitive Centre Foundation as well as links with eminent academics who provide specialist training in specific skills. Training in pro-social modelling and problem-solving is delivered to all members of staff, including the administration team and all members of senior management, demonstrating JPACS's considerable commitment to evidence-based practice throughout the Service. There is an annual Performance Review and Appraisal process and monthly supervision of probation officers. In recent months, JPACS has implemented use of the Jersey Interview Skills Checklist to undertake a structured assessment of practitioner skills in video-recorded intervention sessions as part of the clinical supervision process.

The Chief and Assistant Chief Probation Officers and Team Leader all have an intimate knowledge of the Service and staff and are involved in the supervision of practitioners and direct service delivery, for example, covering office duty on occasion. The managerial style is supportive and reciprocal rather than top-down. Though practitioners described their complex workloads and the high expectations of the organisation as stressful at times, they all expressed confidence in their skills, knowledge, resources and ability to run services effectively. Staff have the

opportunity to identify issues with, and propose modifications to, the Service through discussion with managers or during staff 'away-days'. There is a high level of staff investment and a shared belief in rehabilitation is evident at every level and in every aspect of practice.

Areas for improvement:

The Service requires a high level of skill and commitment and it may be beneficial for practitioners' morale to formalise recognition of positive practice in team meetings over and above the current level of support. Additionally, some practitioners feel that a more transparent case allocation process (ie, discussing the reasons for each allocation with probation officers) may help to prevent them feeling overburdened.

Recommendations:

1. JPACS should continue seeking to recognise, acknowledge, develop and support the skills and commitment of its staff, and enhance the areas of case allocation and positive reinforcement of good practice.

Rating: 18/18 (100%) Very Satisfactory CR: 4.7

E. CLIENT RISK/NEED PRACTICES

Strengths:

There are clear and appropriate selection criteria for the various offender programmes and services that JPACS delivers (eg, Offending Is Not The Only Choice, Self Management And Rational Thinking, Sex Offender Treatment Programme, etc) and

the exclusion criteria for each are entirely relevant, relating to risk levels, criminogenic needs and ability to engage.

All clients are assessed using a valid instrument (LSI-R), which summarises clients' personal characteristics as to the level of risk they present. Additional risk assessment tools are used to assess specific risks of sexual offending, domestic abuse and violence where appropriate (eg, Risk Matrix 2000; SARA; HCR-20). Assessment of dynamic needs is undertaken periodically and the LSI-R is re-administered following completion of interventions and at end of Order, with clients re-classified as appropriate. Clients are also screened to identify any literacy or numeracy needs and language issues that may impact upon their engagement with the Service.

Areas for improvement:

Whilst some responsivity assessment tools are available to assess factors such as depression through referral to specialist agencies, JPACS does not routinely conduct structured assessments of the responsivity of offenders to different styles and modes of service. Staff members are informed instead by their experience and perception of a client's engagement with interventions when determining the way in which they are delivered. Practitioners are assigned to service activities and clients to staff according to caseload capacity and availability; or using experience, intuition and perception of skill and need rather than any formal assessment of staff skills and matching process.

Recommendations:

1. Actuarial measures used to assess responsivity factors including cognitive ability, psychopathy, anxiety and depression (eg, General Aptitude Test

Battery; PCL-R factor 1; Spielberger-State-Trait Anxiety; Beck Depression Inventory) should be available to JPACS staff to inform their assessments.

2. Where possible, staff should routinely be assigned to activities best matching their skills (as assessed through staff selection and supervision processes) and should also be assigned to clients with whom they can work effectively (based on assessment of participant and staff characteristics – eg, staff would be assigned to clients with a similar conceptual level). The mode and style of service delivery should also be adjusted in response to key offender characteristics (eg, low conceptual offenders respond better to higher levels of programme structure).
3. The LSI-R should be re-administered at least every 6 months as well as at end of engagement with interventions.

Rating: 10/13 (77%) Very Satisfactory CR: 4.2

F. PROGRAMME CHARACTERISTICS

Strengths:

JPACS primarily targets criminogenic needs such as addressing pro-criminal attitudes, reducing negative peer associations, enhancing constructive use of leisure time, reducing substance misuse, promoting family affection, communication and problem solving and increasing Education, Training and Employment (ETE)-related performance. Non-criminogenic targets, such as decreasing anxiety and increasing self-esteem, are given less attention, though are accounted for as important

responsivity issues, specific to each case and catered to by interventions such as the 'opt-in' Emotional Coping Skills group.

To address offending behaviour, the Service adopts social-learning and cognitive-behavioural treatment strategies through the delivery of evidence-based manualised programmes and the use of external consultants. Through their engagement with manualised work in group programmes or 1:1 supervision sessions, clients are trained to observe and anticipate problem situations, plan and rehearse alternative pro-social responses, practice new pro-social behaviours in increasingly difficult situations and are trained to monitor and cope with high risk situations, developing relapse prevention plans. Positive feedback is routinely used to reinforce clients' pro-social behaviours. There is also evidence of clients' family and friends being trained to provide support through family problem-solving sessions for young people and home visits/supervision sessions including significant others for adults. Clients are also given the opportunity to engage in further work with JPACS to relearn/reinforce pro-social behaviour, following completion of their statutory engagement.

Practitioners monitor closely the whereabouts and peer associations of clients through information-sharing and recording of contacts. There is an expectation that clients will spend a large proportion of their week in pro-social tasks. Where a client is unemployed, their job-searching activities are monitored as part of their contact with practitioners and interventions are suggested for pro-social use of leisure time, including provision of an Active Card to some clients, enabling their use of local gym facilities.

In line with the evidence regarding effective practice, JPACS predominantly targets higher-risk offenders, with approximately 80% of clients being assessed as presenting a moderate to high risk of re-offending. Programmes with the highest intensity and duration are reserved for higher-risk offenders. For example, the Self Management and Rational Thinking programme targeting high-risk offenders runs for thirty-five sessions, whilst the Offending Is Not The Only Choice programme for medium-risk offenders comprises twenty. Where a client subject to a Probation Order makes good progress, completes interventions successfully and their risk of reoffending is reduced, there is an opportunity for the Order to be returned to the Court for early discharge with the support of JPACS.

Areas for improvement:

Whilst the treatment dosage of programmes delivered to offenders varies by risk, with higher-risk offenders receiving the highest intensity or duration of service, practitioners appear resistant to using risk classification to prioritise resources such as their time and level of investment in case management. A number of practitioners described struggling to manage workloads of chaotic and “needy” cases. It may be beneficial for practitioners to re-frame their prioritisation of cases around risk and to deal with non-criminogenic needs through referral and delegation where possible.

JPACS may also benefit from the use of an operant programme (eg, a token economy) to motivate clients’ compliance through the use of tangible reinforcers and punishers. Appropriate reinforcers must be meaningful to clients and administered immediately, consistently, and made contingent upon performance of pro-social behaviours. Appropriate punishers should be individualised to ensure that they are meaningful to

clients in order to suppress behaviour and should be brief, immediate, consistent and high-intensity.

Whilst clients have some input into the delivery of interventions at the sentence-planning stage and on an informal basis throughout their engagement with JPACS, the only formal mechanism for them to modify the structure and rules of the services provided is through their completion of a feedback questionnaire at the end of their engagement.

Recommendations:

1. Clients should spend at least 40% of their time (or 50 hrs per week) in pro-social tasks. Development and maintenance of a personal weekly timetable would enable practitioners to assist clients in monitoring and achieving this.
2. Treatment dosage of supervision as well as programmes should vary by risk and practitioners should receive periodic training, support and guidance to ensure that their management of cases is commensurate with the level of risk. The evidence suggests that engagement in interventions should total 100 hours for moderate-risk and at least double that for higher-risk offenders.
3. To motivate behaviour-change effectively JPACS could develop a menu of reinforcers (eg, leisure activities, recreational equipment, social reinforcers such as praise, etc) and punishers (eg, withdrawal of privileges, expressions of disapproval, etc). Clients should be given the opportunity to nominate and approve reinforcers and punishers to ensure that they are

meaningful to them individually. Reinforcers must outnumber punishers by a ratio of at least 4:1. Staff also need to be trained to assess whether the punishment is appropriate or produces negative reactions.

4. An additional formal mechanism for clients to provide input to the Service, such as a service-user focus group, would be beneficial in developing client investment in JPACS and improving services.

Rating: 20/24 (83%) Very Satisfactory CR: 4

G. CORE CORRECTIONAL PRACTICE

Strengths:

The evaluator observed a number of positive strategies employed by staff members, particularly in their use of authority and relationship practices. They focused on behaviour rather than the individual, gave specific directions, used their normal (not raised) voices, informed clients of the choices and attendant consequences regarding compliance/non-compliance, gave encouraging messages, supported their words with action, were ‘firm but fair’, and were generally positive regarding a client’s engagement, providing respectful guidance towards, and praising, compliance. They were open, warm, respectful, non-blaming, genuine, flexible, enthusiastic, optimistic, solution-focused and motivational in their work with clients, avoiding argumentation, developing discrepancy gently and supporting self-efficacy. In view of the foregoing, it is unsurprising that all clients identified JPACS as a supportive environment, with the majority describing “having someone to talk to” as the most important element of the Service.

Practitioners positively reinforced pro-social behaviour, explaining why they liked the behaviour exhibited by the client and encouraging clients to consider the benefits of continued use of this behaviour. They also engaged clients in some problem-solving techniques including focusing on the antecedents, behaviour and consequences of their offending, clarifying goals that could resolve or prevent future negative consequences and generating alternative solutions. The evaluator also observed a number of practitioners, including the ADAPT programme facilitators, using elements of cognitive restructuring by developing descriptions of problematic situations, generating descriptions of thoughts and feelings, generating discussions and identifying risky thinking, establishing alternatives and setting up additional practice opportunities.

Areas for improvement:

There was some disparity in delivery, with some practitioners utilising more components of effective treatment interventions than others. There was inconsistency, for example, in the disapproval given to clients regarding anti-social behaviour, which should be immediate and detailed in order to provide the foundation for anti-criminal modelling. Whilst the skill-building procedures of modelling, role-play and practice were referred to in programme manuals and interviews with practitioners and clients, the evaluator observed no practitioner using skill-building techniques with clients and it does appear that this element of effective practice may be somewhat underused. Clients recalling their involvement in role-play did not understand it as a skill-building technique but as a way of illustrating or ‘playing-out’ offending behaviour to increase understanding. It is possible that practitioners are under-confident in using these techniques.

Recommendations:

1. All staff delivering interventions would benefit from continued training in the core correctional practices with a focus on effective disapproval and structured learning procedures for skills-building.
2. It is encouraging that JPACS has recently initiated use of video-recording of intervention sessions and use of the Jersey Supervision Interview Checklist in clinical supervision. These should be used to provide feedback to staff to improve their therapeutic skills.

Rating: 35/45 (78%) Very Satisfactory CR: 3.5

H. INTER-AGENCY COMMUNICATION**Strengths:**

JPACS works in a multi-agency context, including involvement in JMAPPA and RAMAS processes for the multi-agency management of risk and Child Protection processes alongside Social Services. Additionally, JPACS is formally linked to a number of agencies that offer services relevant to the needs of offenders, some of which (eg, victim support, mental health, substance misuse, etc) are represented amongst the staff team at JPACS. Staff routinely advocate with outside agencies on behalf of their clients, and the evaluator observed several examples of this as well as evidence of effective inter-agency communication, including telephone conversations, email correspondence and meetings with staff from other organisations. Leaflets and posters in the waiting room and practitioners' offices provide clients with links to other organisations and the community, and referral documents in JPACS's policy

book are easily accessible to staff members. Staff also have the opportunity to engage in multi-agency training across the various organisations (eg, Child Protection training delivered by Social Services), which enables effective co-ordination between JPACS and other agencies in regards to shared clients.

Areas for improvement:

One client mentioned the embarrassment associated with being seen to pick up certain leaflets or read certain posters (eg, substance misuse support, sexual health, etc) in the communal waiting area.

Recommendations:

1. For greater confidentiality, it might be useful for JPACS to maintain a comprehensive, regularly updated handbook or folder of links to community resources, which could be made available in the waiting room for ease of reference.

Rating: 5/5 (100%) Very Satisfactory CR: 4.8

I. EVALUATION

Strengths:

JPACS takes a conscientious and rigorous approach to evaluation, involving: quarterly checks reviewing files and monitoring treatment progress; video-tapes of programme sessions being sent to the Cognitive Centre Foundation to ensure adherence to the manual; use of the Jersey Supervision Interview Checklist to evaluate the integrity of probation supervision sessions; the distribution of client

feedback questionnaires, with results published annually; the standardised re-assessment of clients on target behaviours; and the gathering of follow-up reconviction data.

The fourth of a series of reports evaluating the effectiveness of JPACS's community supervision services was published in November 2015 and the Service has also been the subject of a number of peer-reviewed journal articles regarding the Jersey Supervision Skills Study (JS3).

Recommendations:

1. The addition of structured measures such as psychometric assessments of attitudes and skills may be of benefit in providing a more in-depth and accurate measure of progress than analysis of changes in the LSI-R scores alone.

Rating: 8/8 (100%) Very Satisfactory CR: 5

SCORING SUMMARY

Category	Points	# of points	N/A	% Score	CR
A. Programme demographics	no scoring				
B. Organisational culture	09	08	0	89%	4.6
C. Programme implementation/maintenance	10	08	0	80%	4.3
D. Management/staff characteristics	18	18	0	100%	4.7
E. Client risk/need practices	13	10	0	77%	4.2
F. Programme characteristics	25	20	1	83%	4
G. Core correctional practice	45	35	0	78%	3.5
H. Inter-agency communication	05	05	0	100%	4.8
I. Evaluation	08	08	0	100%	5
Treatment sub-total (E and F)	38	30	1	81%	4.1
Total	133	112	1	85%	4.4

Very satisfactory	70%+
Satisfactory	50-69%
Unsatisfactory	below 50%

CONCLUSIONS

Overall, Jersey Probation and Aftercare Service received a very satisfactory rating on the CPAI-2010 for its community supervision services. The total score of 85% compares extremely favourably with other published CPAI-2010 scores. It was evident in interviews for this assessment that JPACS's staff are committed, enthusiastic, knowledgeable and skilled.

In the scoring of the CPAI-2010, particular attention is paid to section E (client risk/need practices) and section F (programme characteristics), because the evidence base suggests that both items indicate the strongest validity for intervention success. The combined score of E and F for JPACS is 81% and rated very satisfactory, reflecting the Service's considerable investment in the adoption of evidence-based

approaches to risk assessment and intervention practices, though the evaluation did identify some areas for improvement here, including the need for a more structured approach to assessing and accommodating responsivity issues; and the opportunity to increase effectiveness of services through the use of an operant programme, such as a token economy. It is hoped that these and other recommendations made in this report will contribute to JPACS's ongoing efforts to improve and develop services in line with the evidence base.

REFERENCES

The following reference list contains articles cited in the text of this report as well as significant works on offender rehabilitation and key documents influencing the development of the CPAI. The report author recommends two texts in particular as a starting point for a review of the literature. *The Psychology of Criminal Conduct*, 5th ed. (Andrews & Bonta, 2010 – a sixth edition is forthcoming at the time of writing) gives a comprehensive overview of the extensive research and evidence base for rehabilitative endeavours. *Contemporary Behaviour Therapy*, 5th ed. (Spiegler & Guevremont, 2009) details effective treatment modalities for programme delivery and is an excellent resource for programme development and staff training.

* Cited in the text

Andrews, D. A. (2006). Ensuring Adherence to Risk-Need Responsivity: Making Quality a Matter of Policy. *Criminology and Public Policy*, 5, 595-602.

Andrews, D. A., & Bonta, J. (2010). Rehabilitating Criminal Justice Policy and Practice. *Psychology, Public Policy, and Law*, 16, 39-55.

*Andrews, D. A. & Bonta, J. (2010) *The Psychology of Criminal Conduct*, 5th ed. Cincinnati, OH: Anderson.

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

Andrews, D. A. & Carvell, C. (1998). *Core Correctional Treatment: Core Correctional Supervision and Counseling: Theory, Research, Assessment and Practice*. Carleton University, Ottawa, Ontario.

* Andrews, D. A. And Dowden, C. (2005). Managing Correctional Treatment for Reduced Recidivism: A Meta-Analytic Review of Programme Integrity. , 10, 173–187.

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

Bernfield, G., Farrington, P. & Leschied, (Eds.) (2001). *Offender Rehabilitation in Practice*. Chichester, Wiley.

Bonta, J. (2001). Offender Assessment: General Issues and Considerations. In L. L. Motiuk, & R. C. Serin (Eds.), *Compendium 2000 on Effective Correctional Programming* (pp. 22-29). Ottawa, Canada: Correctional Service of Canada.

Bonta, J., Bourgon, G. & Rugge, T. (2011). *Strategic Training Initiative in Community Supervision (STICS)*. Public Safety of Canada.

Chapman, T. & Hough, M. (1998) *Evidence Based Practice: A Guide to Effective Practice* London: HMIP

Cullen, F. T. (2002) Rehabilitation and Treatment Programs in Wilon, J. Q. & Petersilia, J. (Eds.) *Crime and Public Policy*, 2nd ed. Pp. 253-289 San Francisco, CA: ICS Press.

Cullen, F. T., & Gendreau, P. (2000). Assessing Correctional Rehabilitation: Policy, Practice, and Prospects. In J. Horney (Ed.), *National Institute of Justice Criminal Justice 2000: Changes in Decision Making and Discretion in the Criminal Justice System*. (pp. 109-175). Washington, DC: Department of Justice, National Institute of Justice.

Cullen, F. T. & Gilbert, K. E. (1982). *Reaffirming Rehabilitation*. Cincinnati, OH: Anderson Publishing Company.

French, S. A. & Gendreau, P. (2006) Reducing Prison Misconducts: What Works! *Criminal Justice and Behaviour*, 33, 185-218 DOI: 10.1177/0093854805284406

Gendreau, P. (1996). The Rehabilitation of Offenders: What We Know and What Has To Be Done. *Criminal Justice and Behaviour*, 23, 144-161.

Gendreau, P. (1996). The Principles of Effective Intervention with Offenders. In A.T. Harland (Ed.), *Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply* (pp. 117-130). Thousand Oaks, CA: Sage.

Gendreau, P. & Andrews, D. A. (1990). Tertiary prevention: what the meta-analyses of the offender treatment literature tell us about 'what works'. *Canadian Journal of Criminology*, 32, 173-84.

*Gendreau, P., Andrews, D. A. & Thériault, Y. (2010). *Correctional Program Assessment Inventory – 2010. (CPAI-2010)*. Ottawa, Ontario, Canada.

*Gendreau, P., French, S., & Taylor, A. (2002). *What Works (What Doesn't Work)-Revised 2002: The Principles of Effective Correctional Treatment*. University of New Brunswick at Saint John: Unpublished Manuscript.

Gendreau, P., Goggin, C., Cullen, F. T. & Paparozzi, M. (2002). The common-sense revolution and correctional policy. In J. McGuire (Ed.) *Offender Rehabilitation and*

Treatment: Effective Programmes and Policies to Reduce Re-Offending. Chichester: Wiley.

Gendreau, P., Goggin, C., French, S., & Smith, P. (2006). Practicing Psychology in Correctional Settings. In I. B. Weiner & A. K. Hess (Eds.), *The Handbook of Forensic Psychology*, 3rd ed. (pp. 722-750). Hoboken, NJ: Wiley & Sons. (Gendreau & Goggin's contribution to the 4th edition is currently in press).

Gendreau, P., Goggin, C. & Fulton, B. (2000). Intensive Supervision in Probation and Parole Settings. In C. R. Hollin (Ed.) *Handbook of Offender Assessment and Treatment* (pp. 195-204). Chichester: Wiley.

Gendreau, P., Goggin, C. & Smith, P. (2000). Implementation Guidelines for Correctional Programs in the "Real World". In G. Bernfield, D. P. Farrington & A. W. Leschied (Eds.) *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programmes* (pp. 247-268). London: John Wiley & Sons.

Gendreau, P. & Ross, R. R. (1979). Effective Correctional Treatment: Bibliotherapy for Cynics. *Crime & Delinquency*, 25, 463-489.

Gendreau, P. & Smith, P. (2012). Assessment and Treatment Strategies for Correctional Institutions. In J. Dvoskin, J. Skeem, R. Novaco, & K. Douglas (Eds.). *Using Social Science to Reduce Violent Offending* (pp. 157-177). NY: Oxford Press.

Gendreau, P., Smith, P., & French, S. (2006). The Theory of Effective Correctional Intervention: Empirical Status and Future Direction. In F. Cullen, J. Wright, & M. Coleman (Eds.) *Taking Stock: The Status of Criminological Theory* (pp. 419-446). Piscataway, NJ: Transaction Press.

Gendreau, P., Smith, P., & Theriault, Y. L. (2009). Chaos Theory and Correctional Treatment: Common Sense, Correctional Quackery and the Law of Fartcatchers. *Journal of Contemporary Criminal Justice*, 25, 384-396.

Goggin, C., & Gendreau, P. (2006). The Implementation and Maintenance of Quality Services in Offender Rehabilitation Programs. In C. R. Hollin, & E. J. Palmer (Eds.) *Offending Behaviour Programmes: Development, Application, and Controversies* (pp.209-246). Chichester: John Wiley.

Goldblatt, P. & Lewis, C. (1998) *Reducing Offending: An Assessment of Research Evidence on Ways of Dealing with Offending Behaviour* Home Office Research Study 187. London: Home Office.

*Gray, G. A. (1997). *Does coercion play a significant role in community treatment programs that reduce offender recidivism?* University of New Brunswick. Unpublished Master Thesis.

*Holsinger, A. M. (1999). *Opening the 'black box': Assessing the relationship between program integrity and recidivism*. University of Cincinnati. Unpublished Ph.D. dissertation.

- Latessa, E. J. (1999). What Works in Correctional Intervention. *Southern Illinois University Law Journal*, 23, 415-425.
- Latessa, E. J., Cullen, F. T. & Gendreau, P. (2002). Beyond Correctional Quackery – Professionalism and the Possibility of Effective Treatment. *Federal Probation*, 66, 43-49.
- Latessa, E. J. & Holsinger, A. (1998). The Importance of Evaluating Correctional Programs: Assessing Outcome and Quality. *Corrections Management Quarterly*, 2, 22-29.
- *Lowenkamp, C. T. (2004) *Correctional Program Integrity and Treatment Effectiveness: A Multi-site, Program-level Analysis*. University of Cincinnati. Unpublished Ph.D. dissertation.
- Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does Correctional Program Quality Really Matter? The Impact of Adhering to the Principles of Effective Intervention. *Criminology and Public Policy*, 5, 575-594.
- Matthews, B., Hubbard, D. J. & Latessa, E. (2001). Making the Next Step: Using Evaluability Assessment to Improve Correctional Programming. *The Prison Journal*, 81, 454-472.
- McGuire, J. (1995) *What Works: Reducing Re-offending* Chichester: Wiley Press.
- McGuire, J. (2000) *Cognitive-behavioural Approaches: An Introduction to the Theory and Research* London: Home Office.
- McGuire, J. (2002) Integrated Findings from Research Reviews in McGuire, J. (Ed.) *Offender Rehabilitation and Treatment: Effective Programmes and Policies to Reduce Re-offending* pp. 3-38. Chichester: John Wiley & Sons.
- McGuire, J. (2004). *Understanding Psychology and Crime: Perspectives on Theory and Action*. Maidenhead: Open University Press.
- McGuire, J. (2007) Programmes for Probationers in McIvor, G. & Raynor, P. (Eds.) *Developments in Social Work with Offenders* London: Jessica Kingsley.
- *Nesovic, A. (2003). *Psychometric Evaluation of the Correctional Program Assessment Inventory (CPAI)*. Carleton University, Ottawa, Ontario. Unpublished Ph.D. dissertation.
- Posavac, E. J., & Carey, R. G. (2010). *Program Evaluation: Methods and Case Studies*, 8th ed. Upper Saddle River, N.J.: Prentice Hall.
- Raynor, P. & Vanstone, M. (1996) Reasoning and Rehabilitation in Britain: The Results of the Straight Thinking on Probation (STOP) Programme. *International Journal of Offender Therapy and Comparative Criminology*, 40, 272-284 DOI: 10.1177/0306624X96404003

Smith, P., Gendreau, P., & Swartz, K., (2009). Validating the Principles of Effective Intervention: A Systematic Review of the Contributions of Meta-analysis in the Field of Corrections. *Victims and Offenders*, 4, 1-22.

*Spiegler, M. D. & Guevremont, D. C. (2009). *Contemporary Behaviour Therapy*, 5th ed. Belmont, CA: Wadsworth.

APPENDIX 11

CPAI-2010 Report Conducted on a Women's Project in Wales



**CORRECTIONAL PROGRAM ASSESSMENT INVENTORY – 2010
(CPAI-2010)©**

Conducted on a Women's Project in Wales

**Report prepared by CPAI-2010 accredited evaluator Bridget Kerr,
Swansea Service Evaluation Team (SSET)
375332@swansea.ac.uk**

TABLE OF CONTENTS

Executive Summary.....	p.3
Introduction.....	p.4
Description of the Correctional Program Assessment Inventory – 2010 (CPAI-2010).....	p.5
Methodology.....	p.6
Results.....	p.8
A. Programme demographic.....	p.8
B. Organisational culture.....	p.12
C. Programme implementation/maintenance.....	p.14
D. Management/staff characteristics.....	p.16
E. Client risk/need practices.....	p.18
F. Programme characteristics.....	p.21
G. Core correctional practice.....	p.25
H. Inter-agency communication.....	p.27
I. Evaluation.....	p.28
Scoring Summary.....	p.30
Conclusions.....	p.30
References.....	p.32

EXECUTIVE SUMMARY

In 2014 the XXX was the subject of an evaluation to assess the extent to which services adhere to principles of effective practice. The aims of this study were congruent with the National Offender Management Service's (NOMS) commissioning intention to “ensure that appropriate provision is available to enable women to complete their sentences successfully and that their risks of reoffending are addressed” (2012, p.26). The CPAI-2010 (Gendreau, Andrews & Thériault, 2010) was used to assess nine domains of service delivery (eight of which are scored). The evaluation identified the XXX as having particular strengths in the areas of inter-agency communication and evaluation. Overall, however, the programme achieved an unsatisfactory rating, indicating that it would benefit from substantial improvements in some areas (particularly programme characteristics) in order to achieve its aims.

This evaluation finds that the holistic and needs-responsive approach taken by XXX is worthwhile and strengthened by assiduous evaluation practices. However, it also highlights that effective treatment modalities and risk assessment practices must be central to work that aims to reduce recidivism amongst every offender population, including women.

INTRODUCTION

Established in XXX, XXX is being developed in response to a number of key issues, notably, the Reducing Reoffending Strategy for 2014-2016, which identifies women as a key priority group; Baroness Corston's 2007 *Review of Women with Particular Vulnerabilities in the Criminal Justice System* and subsequent inquiry findings which argue for "a distinct, radically different, visibly-led, strategic, proportionate, holistic, woman-centred, integrated approach" (Home Office, 2007, p.79); and NOMS guidance to taking a "distinct approach" to working with women offenders (2012).

XXX

DESCRIPTION OF THE CORRECTIONAL PROGRAM ASSESSMENT INVENTORY (CPAI-2010)

The CPAI-2010 (Gendreau, Andrews & Thériault, 2010) was designed to measure how closely correctional programmes adhere to the principles of effective practice that have been developed based on meta-analytical research (Andrews & Bonta, 2010). The CPAI-2010 has been used extensively in the United States and Canada where studies have found that programmes with the highest degree of integrity have greater measured effect in terms of reducing re-offending (Andrews & Dowden, 2005, Gray, 1997, Holsinger, 1999, Nesovic, 2003 & Lowenkamp, 2004).

Using interviews with practitioners and offenders, review of agency documents (including case files) and direct observation of agency activity and core correctional practice in programme sessions, CPAI-trained evaluators undertake a detailed assessment of intervention programmes. The nine domains of the CPAI-2010 include 143 items, 133 of which are scored ‘yes’, ‘no’ or ‘not applicable’ according to whether different aspects of practice adhere to the principles of effective practice (Gendreau, French and Taylor, 2002):

A. Programme demographics	10 items (none scored)
B. Organisational culture	9 items
C. Programme implementation/maintenance	10 items
D. Management/staff characteristics	18 items
E. Client risk/need practices	13 items
F. Programme characteristics	25 items
G. Core correctional practice	45 items
H. Inter-agency communication	5 items
I. Evaluation	8 items

To enhance reliability and reduce subjective scoring, there is a confidence rating for each scored item that requires the scorer to indicate the reliability of their assessment

on a five-point scale according to the evidence available to support the score. An overall score is calculated by collating the scores in each domain and dividing the sum of these scores by the total number of applicable items. The overall score is attributed a classification of programme integrity. These range from ‘very satisfactory’ (70% and above) to unsatisfactory (below 50%). The average confidence scores for each section are used to ascertain the reliability and validity of the evaluation’s conclusions. The results of the CPAI-2010 evaluation are given in the form of an in-depth report, which includes the scoring summary, overall classification, qualitative evidence and recommendations. This enables practitioners to articulate the work they do in terms of evidence based practice, identify the strengths and weaknesses of programmes and develop strategies to improve their service.

METHODOLOGY

Initial orientation meetings regarding this evaluation were held on 3rd and 28th April 2014. These were used to provide an overview of the CPAI-2010 to XXX. A timescale for the evaluation was agreed at this time, though it was subject to a number of changes and data collection subsequently spanned the period 23rd June – 12th December 2014.

This report recognises that ideally the CPAI-2010 should be administered over a short period of time – preferably one week or less – in order to ensure that the data collected is representative of day-to-day practice and provides a ‘snapshot’ of the service for evaluation. A number of factors prevented data collection being completed to these timescales in this case. To some extent, this was due to complexities of XXX

itself – ie, the need to identify and collect data from several different agencies at a number of different stages. The evaluation also took place during a period of substantial upheaval, at a time when Wales Probation Trust was dissolved and replaced by the Wales Community Rehabilitation Company (CRC) and National Probation Service (NPS). A number of related and unrelated factors including staffing issues, bidding issues (leading to the dissolution of one partner service), delays in the roll out of the pilot, teething issues, access issues, false starts, communication lapses, etc, all resulted in the lengthy data collection period. Whilst this clearly is not ideal with regard to the integrity of the CPAI-2010, following discussions with the Project Manager, the decision was made to extend the timescale as far as was needed to complete the full evaluation. Swansea Service Evaluation Team (SSET) recognises the need to be responsive to the agency under evaluation and in this way hopes to have provided the opportunity to account for the many changing pressures and achievements peculiar to XXX.

MATERIALS REVIEWED

XXX

RESULTS

A. PROGRAMME DEMOGRAPHICS

This section is descriptive only and does not contribute to the scoring of the CPAI-2010.

XXX is broad in its remit, aiming to engage all women from first point of contact with the CJS in XXX (usually arrest) and to encompass all women engaged in statutory CJS involvement in the community as well as introducing voluntary engagement through diversion to services as an alternative to charge, during bail and pre-sentence and during resettlement from custody.

The programme is delivered in partnership with a number of agencies – The National Probation Service (NPS - responsible for managing “high risk” statutory cases), Wales Community Rehabilitation Company (CRC - responsible for “low-medium risk” statutory cases) and a number of statutory and third sector partners who deliver services to women offenders and women “at risk” of offending. For the purposes of evaluation, practitioners from the NPS, Wales CRC, XXX, XXX and XXX were interviewed whilst multi-agency practices involving a number of other agencies – eg, XXX, XXX, etc – were also reviewed and observed.

Whilst XXX probation office operates as an administrative centre for XXX, there are two premises operating as venues for XXX – XXX and XXX. Both are easily accessible from the city centre, are women-only environments and provide child-friendly facilities. XXX has been purpose built for the provision of services to

vulnerable and volatile women. Its accommodation is arranged over one floor and consists of an open-plan staff office, a large reception with buzzer entry and comfortable waiting area, easily accessible individual interview rooms which are booked out and monitored from reception, a large group training room and kitchen and toilet facilities.

XXX is based in a Victorian three-storey building (converted town house) in need of repair in some places but generally comfortable, with efforts having been made to provide suitable furniture, children's play areas and baby-change facilities, etc. Some adaptations have been made to accommodate XXX, such as the installation of viewing windows in the upstairs interview room doors, and a reception area on the ground floor, though staffing this has been problematic due to an issue over funding. The reception area is a small, two-doored room situated in the centre of the ground floor, separating the toilet, back door and kitchen from the front entrance, group-work room and stairs leading to interview rooms and staff offices. It is therefore a thoroughfare for staff as well as XXX and XXX service users, resulting in a chaotic atmosphere at times. During the evaluation period a member of CRC staff was assaulted by a service user in the reception area, which highlighted the risks associated with the space. A lack of alarms or procedures for such incidents resulted in staff in other areas of the building being unaware that the incident was occurring.

Notwithstanding this concern, XXX service-user feedback is generally positive about the XXX premises as well as XXX, with women identifying both venues as comfortable, accessible, friendly and safe in contrast to their previous experiences of

XXX probation office, which they describe as institutional, uncaring and unsafe in some cases, due to the presence of male probationers.

XXX central team of Project Director and Research and Evaluation Officer is funded via the Integrated Offender Management (IOM) Cymru board. Funding has been agreed for these and one additional support officer post until June 2015 and, whilst there is an expectation of future funding of XXX, the exact nature of source, amount and scope is not confirmed at present. The additional post of Accommodation Pathway Officer has been funded by the National Offender Management Service (NOMS) Cymru and Welsh Government until March 2016. Applying IOM principles regarding the reconfiguration of existing resources, funding for XXX services comes from a variety of sources. The involvement of some agencies is reliant on their successful applications for funds from charitable and statutory sources whilst XXX - dedicated offender managers, link workers and management staff are provided through the existing funding arrangements of the NPS and Wales CRC. At the time of this evaluation, two full time NPS offender managers and three full time CRC offender managers were dedicated to XXX.

All women in the CJS are in the scope of XXX and so there is no prescribed maximum capacity for the project, though individual offender manager caseloads of statutory cases are limited by the NPS and Wales CRC and, at last point of data collection for evaluation, were at around 90-99% in some cases. Practitioners expressed concern that the limit is appropriate to a mixed-gender caseload and that they are unable to meet the complex and multiple needs of women-only caseloads at this capacity.

Additional to CJS referrals as part of the Diversion Scheme, pre-court, when sentenced in the community and on release from custody, women can self-refer to XXX if they have previously been involved with the CJS or identify themselves as being at risk of committing offences. The majority of CJS involvement with women service users is related to violent and/or acquisitive offending behaviour, whilst prevailing presenting problems are deficits in thinking and behaviour, lifestyle and associates, substance misuse, accommodation, mental health issues and victimisation issues, these latter two most often not being directly related to offences but significant as responsivity issues when engaging women with services.

B. ORGANISATIONAL CULTURE

Strengths:

XXX goals and mandate are clearly laid out in XXX documents and a brief overview given in leaflets for stakeholders and participants. This provides a single strategic direction that unites the staff, agencies and service users that comprise XXX. The pilot has been initiated within a reasonable period of time, with changes being made to accommodate new services and/or processes as necessary, facilitated through generally excellent information sharing practices including stakeholder meetings, emails and newsletters. A particular strength of XXX is the use of external consultants to advise on programming and the establishment of designated internal evaluation and research services. Continued and broadened use of these to inform all areas of service delivery would increase programme effectiveness as well as to contribute to the evidence base in respect of women offenders.

Areas for improvement:

XXX lacks a documented code of ethics. This has particular importance in the multi-agency context of XXX where disparate agencies and staff members must be united under a shared value base.

There is no clear, coherent process for dealing with problems, resulting in organisational disharmony when difficulties arise. This is due, in part, to a complicated and unwieldy management structure, which will be explored further in Section D. At present the programme is at risk of low morale amongst practitioners and high staff turnover, which would reduce the effectiveness of services.

Recommendations:

1. A document detailing ethics of intervention would consolidate the shared vision outlined in XXX documents. Staff should be trained to ensure that the ethical code inculcates every aspect of services.
2. Extended use of external consultants would be particularly helpful in, for example, developing gender-informed manuals for effective programme delivery.
3. To stabilise the programme, ensure organisational harmony and consolidate processes, the management structure must be cohesive and clearly defined, with appropriate policies and procedures in place to respond to problems in a non-confrontational, crisis-free fashion. These must be easily accessible to all staff (and service users) in XXX and must be routinely followed.

Rating: 5/9 (56%) Satisfactory CR: 4

C. PROGRAMME IMPLEMENTATION/MAINTENANCE

Strengths:

XXX is being developed in part as a response to national evidence regarding the high number of low-risk female offenders in custody and the need to take “a distinct approach” to women offenders (NOMS 2012), though is also informed by local needs established through consultation with existing agencies in the field - women’s groups, police etc. The Project Director has been instrumental in designing, establishing and maintaining XXX through ongoing consultation with stakeholders. The creation of a steering group ensures that XXX is congruent with existing stakeholder values and practices, whilst planned monthly multi-agency training events will orientate practitioners and management staff to services offered across XXX.

A particular strength of XXX is its commitment to piloting and evaluation, which provides the opportunity to attend to logistics and content prior to full implementation.

Areas for improvement:

XXX has been established at a time of significant upheaval in the field of corrections, due to the division of probation services into the NPS and Wales CRC. Though this is not within the Project Director’s sphere of influence, evidence suggests that far-reaching contentious issues such as marked changes in policy, staffing and management structure do have a negative impact on programme quality. It is therefore especially important to proof the service against this pressure as far as possible by providing a stable and supportive working environment. The current organisational

context of XXX, whereby CRC and NPS staff have been relocated to hub premises without on-site supervision, managerial or administrative support, is likely to add to tensions. This will be further explored in Section D.

The quality of the programme is also dependent on the stability of stakeholder agencies, and they are operating in a context of funding cuts and increased competition. Funding for these services, and indeed for XXX itself, is not guaranteed beyond the short term and one partner service actually dissolved during the evaluation period. It is crucial that XXX is responsive to the differing pressures on third sector agencies in particular, and that their presence in a competitive environment is strengthened by their participation.

The XXX literature review completed by the Research and Evaluation Officer is focused on the demographics of the female offender population. The development of interventions and group work programmes should also be informed by a full review of the literature in respect of effective treatment practices.

Recommendations:

1. The potential for collaborative bidding processes should be explored with Stakeholders, whilst evaluation and research into XXX should be used to inform commissioners.
2. A comprehensive search of literature should be undertaken to identify programme materials directly related to effective treatment using cognitive, behavioural or social learning modalities. Reviewed literature should consist of criminological and psychological journals and texts and this process should

be ongoing to inform programme development. The reference list for this evaluation provides a starting point for such a search.

Rating: 5/8 (63%) Satisfactory CR: 4.4

D. MANAGEMENT/STAFF CHARACTERISTICS

Strengths:

There is a high level of staff investment in XXX across all services and a shared belief in rehabilitation is evident at all levels. The Project Director has previous full time direct experience of delivering interventions to offenders and an intimate knowledge of the agencies and staff involved in XXX as well as the support of senior officials and line staff, meaning that she is well placed to develop services. The majority of XXX staff have relevant qualifications, training and experience and are confident of their skills, knowledge and ability to run services effectively. It is encouraging that the Project Director is committed to enabling staff to propose modifications to the programme, and processes to facilitate this should remain in place throughout implementation and maintenance of XXX.

Areas for improvement:

Individual agencies are responsible for their own hiring protocol and at present there is no process in place to ensure that all staff working under the aegis of XXX are DBS checked.

With regard specifically to statutory practices, the NPS and CRC staff are currently line managed independently by staff unrelated to XXX and based at XXX probation office, where all support staff (admin, IT, finance, etc) are also based. This lack of on-site management, supervision and administrative support is contributing to XXX team's isolation from the service as a whole, a lack of oversight (regarding hours worked, working practices, programme logistics, etc), limited managerial support, the absence of a clearly delineated structure for addressing problems and low morale amongst staff members.

Additionally, staff selection processes do not currently involve a structured assessment or appraisal of skills relevant to reducing recidivism, there is no schedule in place to provide regular training in effective practice, and there is no clinical supervision of practice. Staff identified their previous experience of the Skills for Effective Engagement, Development and Supervision (SEEDS) programme as positive, key to which is the development of practitioner skills and practice in working with offenders as well as ongoing supervision of practice, though this was discontinued during the restructuring of services.

Recommendations:

1. There should be centralised responsibility for ensuring that DBS checks are current for all members of staff operating under the aegis of XXX across all agencies.

Items 2-5 are specific to NPS & CRC management/staffing practices:

2. There is a need for an on-site member of staff with a managerial role in the selection and training of staff, conducting some aspects of intervention

delivery, and providing direct supervision of staff. XXX practitioners must be managed as a team, which the current management structure cannot achieve.

3. It is recommended that processes are put in place to ensure that staff are selected according to a structured assessment of relationship skills (eg, enthusiasm, reflectiveness, non-blaming, etc) using, for example, psychometric evaluation or vignettes, as well as skill factors specific to evidence based practice (ie, directive, solution-focused, structured, contingency based, cognitive restructuring, pro-social modelling, effective reinforcement, disapproval and problem solving).
4. Staff should receive regular training in the theory and practice of interventions and the skill factors needed to deliver services effectively.
5. Effective practice must include formal supervision of staff which should involve at least an annual assessment of clinical skills related to service delivery, whilst regular clinical supervision of staff members' work with women should be held at least bi-monthly.

Rating: 8/18 (44%) Unsatisfactory CR: 4.4

E. CLIENT RISK/NEED PRACTICES

Strengths:

A particular strength of XXX is its commitment to diverting low-risk offenders from the CJS. This reflects the evidence base which indicates that successful offending

behaviour programmes are targeted at participants presenting a medium-high risk of re-offending, whilst also taking a humanitarian and holistic approach to ensuring that services are in place to respond to the needs of women outside of the CJS, ie, not “diversion to nothing”.

There are clear and appropriate selection criteria for admittance into XXX and the exclusion criteria are entirely relevant, relating to gender and location.

Statutory cases are assessed using a valid instrument (OASys), which summarises offenders’ personal characteristics as to the level of risk that they present. This is re-administered periodically, and offenders re-classified as appropriate.

XXX intention to generate recidivism norms on its own population reflects the service’s dedication to strengthening its risk/need practices.

Areas for improvement:

No assessment of risk of re-offending is undertaken, except that required by statutory case management. Non-statutory cases are referred to services based on an assessment of their eligibility for diversion (which does not necessarily preclude those presenting a higher-risk of recidivism), and areas of need related to re-offending.

Whilst some responsivity assessment tools are available to assess factors such as depression through referral to specialist agencies, XXX does not routinely assess the responsivity of offenders to different styles and modes of service. Staff members are instead informed by their experience and perception of a woman’s engagement with

interventions when determining the way in which they are delivered. Staff are assigned to programme activities and women to staff according to caseload capacity and availability or using experience, intuition and perception of skill and need rather than any formal assessment of staff skills and matching process.

Recommendations:

1. A validated risk assessment tool such as the Level of Service Inventory – Revised - Screening Version (LSI-R-SV) should be used on all admissions to XXX to summarise the level of risk of re-offending presented by cases, which should then be used to inform the level of interventions delivered. Cases presenting a low risk of re-offending are not suitable for probation type intervention such as group programmes focused on offending behaviour.
2. XXX should routinely assess how the characteristics of women (eg, ethnicity, motivation and psychopathy) may interact with the style and mode of service delivery.
3. Actuarial measures should be used to assess responsivity factors such as cognitive ability (General Aptitude Test Battery), psychopathy (PCL-R factor 1), anxiety (Spielberger-State-Trait Anxiety) and depression (Beck Depression Inventory).
4. Where possible, staff should routinely be assigned to activities best matching their skills (as assessed through staff selection and supervision processes) and should also be assigned to women with whom they can work effectively (based on assessment of participant and staff characteristics – eg, staff would be assigned to women with a similar

conceptual level). The mode and style of service delivery would also be adjusted in response to key offender characteristics (eg, low conceptual offenders respond better to higher levels of programme structure).

Rating: 6/13 (46%) Unsatisfactory CR: 4.6

F. PROGRAMME CHARACTERISTICS

This section has been scored based on the characteristics of services within XXX that aim to reduce the recidivism of statutory participants. It is not applicable to diversion services, since exposing women who present a low risk of re-offending to the probation type interventions itemised in this section of the evaluation is ineffective practice and it is not possible to delineate voluntary participants according to their risk level due to the absence of assessment and categorisation in this respect.

Strengths:

XXX primarily targets criminogenic needs such as addressing pro-criminal attitudes, reducing negative peer associations, enhancing constructive use of leisure time, reducing substance misuse, promoting family affection, communication and problem solving and increasing Education, Training and Employment (ETE) related performance. Positively (in statutory services) non-criminogenic targets, such as decreasing anxiety and increasing self-esteem, are given less attention, though are accounted for as important responsivity issues, specific to each case. XXX is particularly effective in enabling staff from all services to monitor closely the

whereabouts and peer associations of women and to share information regarding this. Evidence suggests that these behaviours are especially indicative of risk of re-offending and so close monitoring enables effective risk assessment and management throughout women's engagement with XXX. Additionally, XXX has a strong participatory ethos and this is reflected in practitioners' commitment to a "mutual and collaborative approach". Processes whereby women can have input into the programme structures (with the approval of senior staff) should be made explicit.

Areas for improvement:

There are a number of services and group work programmes available across XXX that aim to reduce reoffending, and to which statutory cases can be referred (and in some cases required to complete) as part of their Order. However, there are no discernible differences in criteria for inclusion in, for example, the Women's Stand Alone Requirement (SAR) delivered by the NPS and CRC Wales, XXX programme and services delivered by XXX. As a result, it appears relatively arbitrary which services women are referred to. Additionally, these services do not routinely employ behavioural, social learning or cognitive behavioural strategies as treatment modalities, instead employing techniques based on restorative practices, focused discussion, outreach, mentoring, practical activities and (in the case of the XXX programme that XXX hopes to introduce) corporate development strategies. Whilst NPS and CRC practitioners did mention some techniques associated with effective practice, no manual is routinely used for the statutory supervision of XXX offenders, and the emphasis is instead on client-centered supervision. The observed exception to this was the accredited Enhanced Thinking Skills (ETS) programme, which employs effective behavioural strategies through manualised groupwork. However, the two

participants of ETS at the time of observation both presented a low risk of re-offending according to their Offender Group Reconviction Scale (OGRS) score. At present, the level of interventions delivered across XXX is not determined according to the level of risk of reoffending and this limits the effectiveness of services.

Whilst a concerted effort is made to provide, and engage women in, pro-social leisure and ETE related activities, there is no dedicated record of women's weekly timetables making effective monitoring in this area more difficult.

Recommendations:

1. Clearer criteria for service referral and/or programme inclusion would reduce duplication across services.
2. Group work and 1:1 supervision programmes should be developed to utilise effective behavioural, social learning and cognitive-behavioural strategies through the provision of manuals detailing these types of treatment and the activities to be delivered.
3. The formulation of relapse prevention strategies should be built in to delivery of XXX. Effective strategies involve training women to observe and anticipate problem situations, plan and rehearse alternative pro-social responses and practice these in increasingly difficult situations. Improved competencies should be rewarded to reinforce changed behaviours and family members and/or friends should be trained to provide support. The effectiveness of the programme would be further strengthened through the addition of booster sessions to consolidate progress after completion.

4. Statutory interventions should target medium-high risk offenders and treatment dosage should vary by risk, with higher-risk offenders receiving the highest intensity or duration of service. Therefore, intervention completion criteria should be based on the acquisition of pro-social behaviours and beliefs and an associated reduction in risk of re-offending.
5. Women should spend at least 40% of their time (or 50 hrs per week) in pro-social tasks. Development and maintenance of a personal weekly timetable would enable practitioners to assist women in monitoring and achieving this.
6. To motivate behaviour change effectively XXX should develop a menu of reinforcers across services (eg, leisure activities, recreational equipment, social reinforcers such as praise, etc) and punishers (eg, withdrawal of privileges, expressions of disapproval, etc). Programme participants should be given the opportunity to nominate and approve reinforcers and punishers to ensure that they are meaningful to them individually. Reinforcers must outnumber punishers by a ratio of at least 4:1. Staff also need to be trained to assess whether the punishment is appropriate or produces negative reactions.

Rating: 6/25 (24%) Unsatisfactory CR: 4.6

G. CORE CORRECTIONAL PRACTICE

As above, this section applies to services delivered to statutory cases. The scoring of items in this section is based on actual observations of the therapeutic relationship, in this instance one-to-one supervision sessions delivered by NPS and CRC practitioners and group work sessions of the Women's SAR and ETS programmes.

Strengths:

The evaluator observed a number of positive strategies employed by staff members, particularly in their use of authority, where they focused on behaviour rather than the individual, gave specific directions, used their normal (not raised) voices, informed women of the choices and attendant consequences regarding compliance/non-compliance, gave encouraging messages, supported their words with action, were "firm but fair", and were generally positive regarding a woman's engagement, providing respectful guidance towards, and praising, compliance.

Practitioners also demonstrated positive relationship practices. They were generally open, warm, respectful, non-blaming, genuine, flexible, enthusiastic and optimistic in their engagement with women.

In most sessions, women were reinforced for demonstrating desired behaviours and practitioners were rarely punishing or negative, though did express disapproval of anti-social thinking and/or behaviour when appropriate.

Areas for improvement:

Practitioner approach and level of therapeutic skills were generally of a similar standard, with no session demonstrating all of the necessary components of effective treatment intervention. The employment of incomplete and ineffective strategies resulted partly from the lack of standardised use of an evidence-based therapeutic manual throughout. The ETS programme was a notable exception. The observed session drew on an evidence-based, accredited manual and the employed treatment strategies were in line with effective core correctional practice.

Recommendations:

1. All staff delivering statutory interventions would benefit from training in the core correctional practices of anti-criminal modelling, effective reinforcement and disapproval, problem-solving techniques, structured learning procedures for skill building and elements of cognitive restructuring. Training in aspects of relationship practices, particularly motivational interviewing techniques (eg, rolling with resistance, developing discrepancy and supporting self-efficacy), would also be of use. This would ensure that all frontline staff are equipped to provide therapeutic interventions and would broaden existing practice to encompass all aspects of effective treatment.
2. Structured, manualised programmes should also encompass all of these core correctional practices to maximise efficacy.
3. Clinical supervision should be used to provide feedback to staff to improve their therapeutic skills.

Rating: 18/45 (40%) Unsatisfactory CR: 4.4

H. INTER-AGENCY COMMUNICATION

Strengths:

XXX has the potential to be an exemplar of a multi-agency service. Cases are managed co-operatively and referrals discussed in weekly multi-agency case meetings, ensuring that services are delivered across agencies according to the presenting needs of women. Additionally, XXX staff regularly have contact with relevant agencies outside of XXX (eg, schools, hospitals, housing providers, etc) and advocate for service users where appropriate. The planned monthly multi-agency training events will add to effective co-ordination of treatment for service users, as will use of a shared database to record contacts.

Areas for improvement:

At the time of evaluation, a handbook providing links with other organisations and the community had not been placed in the reception areas of XXX premises as planned. This would be a valuable addition to the literature already available to service users. Staff have noted that some agency representatives are regularly absent from case meetings and have also identified an over-reliance on email for communicating information related to XXX. These issues present significant but easily surmountable obstacles to effective inter-agency practice.

Recommendations:

1. Processes should be put in place to ensure that regular non-attendance in case meetings is followed up with the relevant agency.

2. XXX staff team meetings for core practitioners would be a useful forum for disseminating information, addressing issues related to practice, consolidating working relationships, providing training, reducing isolation and improving morale.
3. The introduction of a shared database for recording contacts across agencies will enable more effective coordination of services.
4. A handbook providing links to other organisations and services should be easily available to service users.

Rating: 5/5 (100%) Very Satisfactory CR: 4

I. EVALUATION

With regard to this section, please note that the low confidence score reflects the lack of evidence available for practices that could only be undertaken once XXX is established such as, for example, publication of an article detailing its effectiveness. This section has been scored according to the processes being put in place for future practice.

Strengths:

XXX will be subject to ongoing evaluation of both statutory and diversion services. Programme checks will monitor treatment progress, women will be surveyed as to their satisfaction with the service, assessment scoring regarding target behaviours will be analysed, and (re-)arrest, (re-)conviction and/or (re-)incarceration data will be

gathered for a substantial amount of time following programme completion and compared with a risk-control comparison group for analysis. This dedicated approach to evaluation will enable the production of a formally-written document detailing the effectiveness of the programme that can be used to inform commissioners, policy makers and service providers.

Areas for improvement:

There is no current plan to contribute to the academic literature in respect of interventions for women, which an evaluation of XXX would be well placed to do.

Recommendations:

1. Following full evaluation, a document containing introduction, method, results and discussion that details the effectiveness of XXX should be published in a peer-reviewed journal. This would enhance the credibility of XXX as well as being a useful addition to the academic literature.
2. Programme sessions should be observed regularly to ensure fidelity to the manual.
3. The addition of structured measures such as psychometric assessments of attitudes and skills would be of benefit in providing a more accurate measure of progress than analysis of changes in the (OASys) scores for target behaviours alone.

Rating: 6/8 (75%) Very Satisfactory CR: 3.5

SCORING SUMMARY

Category	Points	# of points	N/A	% Score	CR
A. Programme demographics	no scoring				
B. Organisational culture	09	05	0	56%	4
C. Programme implementation/maintenance	10	05	2	63%	4.4
D. Management/staff characteristics	18	08	0	44%	4.4
E. Client risk/need practices	13	06	0	46%	4.6
F. Programme characteristics	25	06	0	24%	4.6
G. Core correctional practice	45	18	0	40%	4.4
H. Inter-agency communication	05	05	0	100%	4
I. Evaluation	08	06	0	75%	3.5
Treatment sub-total (E and F)	38	12	0	32%	4.6
Total	133	59	2	45%	4.2

Very satisfactory	70%+
Satisfactory	50-69%
Unsatisfactory	below 50%

CONCLUSIONS

Overall, XXX received an unsatisfactory rating on the CPAI-2010, indicating that improvements need to be made in order for the programme to meet its aims, though the evaluation also found a number of strengths, particularly in the areas of inter-agency communication and evaluation.

In the scoring of the CPAI-2010, particular attention is paid to section E (client risk/need practices) and section F (programme characteristics), because the evidence base suggests that both items indicate the strongest validity for programme success. The combined score of E and F for XXX is 32% and rated unsatisfactory. XXX focus on criminogenic needs is to be commended, though the essential components of effective risk assessment and treatment strategies are under-used, limiting the programme's ability to address these.

Additionally, deficits in the management structure have the potential to irreparably damage the cohesiveness of the organisation and possibly result in further serious incidents. In this regard, particular attention should be paid to addressing the relevant deficits reported in sections A (programme demographics), B (organisational culture), C (programme implementation/maintenance), and D (management/staff characteristics).

Since the CPAI-2010 has been formulated to evaluate programmes that primarily aim to reduce the recidivism of proven offender populations, some items are not relevant to some aspects of XXX Diversion Scheme. These elements have been highlighted throughout this report, which aims to give an overview of XXX practice as a whole, whilst more closely investigating the use of evidence based strategies that reduce reoffending. The statutory practices of CRC Wales and NPS have therefore come under particular scrutiny by this evaluation. It is important and encouraging that IOM Cymru are commissioning a separate, independent evaluation of the Diversion Scheme, which should properly address its diverse and holistic aims.

REFERENCES

The following reference list contains articles cited in the text of this report as well as significant works on offender rehabilitation and key documents influencing the development of the CPAI. The report author recommends three texts in particular as a starting point for a review of the literature. *The Psychology of Criminal Conduct*, 5th ed. (Andrews & Bonta, 2010) gives a comprehensive overview of the extensive research and evidence base for rehabilitative endeavours. *The Assessment and Treatment of Women Offenders* (Blanchette & Brown, 2006) applies this specifically to women. *Contemporary Behaviour Therapy*, 5th ed. (Spiegler & Guevremont, 2009) details effective treatment modalities for programme delivery and is an excellent resource for programme development and staff training.

* Cited in the text

Andrews, D. A. (2006). Ensuring Adherence to Risk-Need Responsivity: Making Quality a Matter of Policy. *Criminology and Public Policy*, 5, 595-602.

Andrews, D. A., & Bonta, J. (2010). Rehabilitating Criminal Justice Policy and Practice. *Psychology, Public Policy, and Law*, 16, 39-55.

* Andrews, D. A. & Bonta, J. (2010) *The Psychology of Criminal Conduct*, 5th ed. Cincinnati, OH: Anderson.

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

Andrews, D. A. & Carvell, C. (1998). *Core Correctional Treatment: Core Correctional Supervision and Counseling: Theory, Research, Assessment and Practice*. Carleton University, Ottawa, Ontario.

Andrews, D.A. & Dowden, C. (1999). A Meta-analytic Investigation into Effective Correctional Intervention for Female Offenders. *Forum on Corrections Research*, 11, 18-21.

* Andrews, D. A. And Dowden, C. (2005). Managing Correctional Treatment for Reduced Recidivism: A Meta-Analytic Review of Programme Integrity. , 10, 173–187.

- Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P., et al. (1990). Does correctional treatment work? A clinically relevant and psychologically informed meta-analysis. *Criminology*, 28, 369-404.
- Bernfield, G., Farrington, P. & Leschied, (Eds.) (2001). *Offender Rehabilitation in Practice*. Chichester, Wiley.
- Blanchette, K. & Brown, S. L. (2006). *The Assessment and Treatment of Women Offenders*. Chichester: Wiley.
- Bonta, J. (2001). Offender Assessment: General Issues and Considerations. In L. L. Motiuk, & R. C. Serin (Eds.), *Compendium 2000 on Effective Correctional Programming* (pp. 22-29). Ottawa, Canada: Correctional Service of Canada.
- Bonta, J., Bourgon, G. & Rugge, T. (2011). *Strategic Training Initiative in Community Supervision (STICS)*. Public Safety of Canada.
- Chapman, T. & Hough, M. (1998) *Evidence Based Practice: A Guide to Effective Practice* London: HMIP
- Cullen, F. T. (2002) Rehabilitation and Treatment Programs in Wilon, J. Q. & Petersilia, J. (Eds.) *Crime and Public Policy*, 2nd ed. Pp. 253-289 San Francisco, CA: ICS Press.
- Cullen, F. T., & Gendreau, P. (2000). Assessing Correctional Rehabilitation: Policy, Practice, and Prospects. In J. Horney (Ed.), *National Institute of Justice Criminal Justice 2000: Changes in Decision Making and Discretion in the Criminal Justice System*. (pp. 109-175). Washington, DC: Department of Justice, National Institute of Justice.
- Cullen, F. T. & Gilbert, K. E. (1982). *Reaffirming Rehabilitation*. Cincinnati, OH: Anderson Publishing Company.
- Dowden, C. & Andrews, D. A. (1999). What Works for Female Offenders: A Meta-analytic Review. *Crime and Delinquency*, 45, 438-452.
- French, S. A. & Gendreau, P. (2006) Reducing Prison Misconducts: What Works! *Criminal Justice and Behaviour*, 33, 185-218 DOI: 10.1177/0093854805284406
- Gendreau, P. (1996). The Rehabilitation of Offenders: What We Know and What Has To Be Done. *Criminal Justice and Behaviour*, 23, 144-161.
- Gendreau, P. (1996). The Principles of Effective Intervention with Offenders. In A.T. Harland (Ed.), *Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply* (pp. 117-130). Thousand Oaks, CA: Sage.
- Gendreau, P. & Andrews, D. A. (1990). Tertiary prevention: what the meta-analyses of the offender treatment literature tell us about 'what works'. *Canadian Journal of Criminology*, 32, 173-84.

*Gendreau, P., Andrews, D. A. & Thériault, Y. (2010). *Correctional Program Assessment Inventory – 2010. (CPAI-2010)*. Ottawa, Ontario, Canada.

*Gendreau, P., French, S., & Taylor, A. (2002). *What Works (What Doesn't Work)-Revised 2002: The Principles of Effective Correctional Treatment*. University of New Brunswick at Saint John: Unpublished Manuscript.

Gendreau, P., Goggin, C., Cullen, F. T. & Paparozzi, M. (2002). The common-sense revolution and correctional policy. In J. McGuire (Ed.) *Offender Rehabilitation and Treatment: Effective Programmes and Policies to Reduce Re-Offending*. Chichester: Wiley.

Gendreau, P., Goggin, C., French, S., & Smith, P. (2006). Practicing Psychology in Correctional Settings. In I. B. Weiner & A. K. Hess (Eds.), *The Handbook of Forensic Psychology, 3rd ed.* (pp. 722-750). Hoboken, NJ: Wiley & Sons. (Gendreau & Goggin's contribution to the 4th edition is currently in press).

Gendreau, P., Goggin, C. & Fulton, B. (2000). Intensive Supervision in Probation and Parole Settings. In C. R. Hollin (Ed.) *Handbook of Offender Assessment and Treatment* (pp. 195-204). Chichester: Wiley.

Gendreau, P., Goggin, C. & Smith, P. (2000). Implementation Guidelines for Correctional Programs in the "Real World". In G. Bernfield, D. P. Farrington & A. W. Leschied (Eds.) *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programmes* (pp. 247-268). London: John Wiley & Sons.

Gendreau, P. & Ross, R. R. (1979). Effective Correctional Treatment: Bibliotherapy for Cynics. *Crime & Delinquency*, 25, 463-489.

Gendreau, P. & Smith, P. (2012). Assessment and Treatment Strategies for Correctional Institutions. In J. Dvoskin, J. Skeem, R. Novaco, & K. Douglas (Eds.) *Using Social Science to Reduce Violent Offending* (pp. 157-177). NY: Oxford Press.

Gendreau, P., Smith, P., & French, S. (2006). The Theory of Effective Correctional Intervention: Empirical Status and Future Direction. In F. Cullen, J. Wright, & M. Coleman (Eds.) *Taking Stock: The Status of Criminological Theory* (pp. 419-446). Piscataway, NJ: Transaction Press.

Gendreau, P., Smith, P., & Theriault, Y. L. (2009). Chaos Theory and Correctional Treatment: Common Sense, Correctional Quackery and the Law of Fartcatchers. *Journal of Contemporary Criminal Justice*, 25, 384-396.

Goggin, C., & Gendreau, P. (2006). The Implementation and Maintenance of Quality Services in Offender Rehabilitation Programs. In C. R. Hollin, & E. J. Palmer (Eds.) *Offending Behaviour Programmes: Development, Application, and Controversies* (pp.209-246). Chichester: John Wiley.

Goldblatt, P. & Lewis, C. (1998) *Reducing Offending: An Assessment of Research Evidence on Ways of Dealing with Offending Behaviour* Home Office Research Study 187. London: Home Office.

*Gray, G. A. (1997). *Does coercion play a significant role in community treatment programs that reduce offender recidivism?* University of New Brunswick. Unpublished Master Thesis.

*Holsinger, A. M. (1999). *Opening the 'black box': Assessing the relationship between program integrity and recidivism.* University of Cincinnati. Unpublished Ph.D. dissertation.

* Home Office (2007). *The Corston Report: A Report by Baroness Corston of a Review of Women With Particular Vulnerabilities in the Criminal Justice System.* London: Home Office.

Latessa, E. J. (1999). What Works in Correctional Intervention. *Southern Illinois University Law Journal*, 23, 415-425.

Latessa, E. J., Cullen, F. T. & Gendreau, P. (2002). Beyond Correctional Quackery – Professionalism and the Possibility of Effective Treatment. *Federal Probation*, 66, 43-49.

Latessa, E. J. & Holsinger, A. (1998). The Importance of Evaluating Correctional Programs: Assessing Outcome and Quality. *Corrections Management Quarterly*, 2, 22-29.

*Lowenkamp, C. T. (2004) *Correctional Program Integrity and Treatment Effectiveness: A Multi-site, Program-level Analysis.* University of Cincinnati. Unpublished Ph.D. dissertation.

Lowenkamp, C. T., Latessa, E. J., & Smith, P. (2006). Does Correctional Program Quality Really Matter? The Impact of Adhering to the Principles of Effective Intervention. *Criminology and Public Policy*, 5, 575-594.

Matthews, B., Hubbard, D. J. & Latessa, E. (2001). Making the Next Step: Using Evaluability Assessment to Improve Correctional Programming. *The Prison Journal*, 81, 454-472.

McGuire, J. (1995) *What Works: Reducing Re-offending* Chichester: Wiley Press.

McGuire, J. (2000) *Cognitive-behavioural Approaches: An Introduction to the Theory and Research* London: Home Office.

McGuire, J. (2002) Integrated Findings from Research Reviews in McGuire, J. (Ed.) *Offender Rehabilitation and Treatment: Effective Programmes and Policies to Reduce Re-offending* pp. 3-38. Chichester: John Wiley & Sons.

McGuire, J. (2004). *Understanding Psychology and Crime: Perspectives on Theory and Action.* Maidenhead: Open University Press.

McGuire, J. (2007) Programmes for Probationers in McIvor, G. & Raynor, P. (Eds.) *Developments in Social Work with Offenders* London: Jessica Kingsley.

* National Offender Management Service (2012). *Discussion Document: NOMS Commissioning Intentions 2012-13, Version 2*. London: Ministry of Justice, NOMS.

*National Offender Management Service Women and Equalities Group (2012). *A Distinct Approach: A Guide to Working With Women Offenders*. London: Ministry of Justice, NOMS.

*Nesovic, A. (2003). *Psychometric Evaluation of the Correctional Program Assessment Inventory (CPAI)*. Carleton University, Ottawa, Ontario. Unpublished Ph.D. dissertation.

Posavac, E. J., & Carey, R. G. (2010). *Program Evaluation: Methods and Case Studies, 8th ed.* Upper Saddle River, N.J.: Prentice Hall.

Raynor, P. & Vanstone, M. (1996) Reasoning and Rehabilitation in Britain: The Results of the Straight Thinking on Probation (STOP) Programme. *International Journal of Offender Therapy and Comparative Criminology*, 40, 272-284 DOI: 10.1177/0306624X96404003

Smith, P., Gendreau, P., & Swartz, K., (2009). Validating the Principles of Effective Intervention: A Systematic Review of the Contributions of Meta-analysis in the Field of Corrections. *Victims and Offenders*, 4, 1-22.

*Spiegler, M. D. & Guevremont, D. C. (2009). *Contemporary Behaviour Therapy, 5th ed.* Belmont, CA: Wadsworth.

APPENDIX 12

Research Participant Consent to Audio Recording



Research Participant Consent to Audio Recording – intervention sessions

Swansea Service Evaluation Team (SSET) will be recording sessions onto MP3 files to check that the CPAI-2010 evaluation is being conducted properly. The recordings may also be listened to by the CPAI training team.

I understand that audio files will be stored on password protected computers and that my real name will not be used by SSET at any time.

I understand that I can still be involved in the study even if I don't give my consent to the recording.

I do/do not give my consent to the session on _____ being audio-recorded.

Signature

Date

I understand that SSET may want to keep recordings on file, for future research and teaching. If I don't agree to this, I understand that SSET will delete the recording of this session when the study ends on 28th June 2017.

I do/do not give my consent to the audio-recording being kept for future research and teaching.

Signature

Date



Research Participant Consent to Audio Recording – feedback interviews

Swansea Service Evaluation Team (SSET) will be recording feedback interviews onto MP3 files in order to fully transcribe responses. The recordings may also be listened to by the CPAI training team.

I understand that audio files will be stored on password protected computers and that my real name will not be used by SSET at any time.

I understand that I can still be involved in the study even if I don't give my consent to the recording.

I do/do not give my consent to the interview on _____ being audio-recorded.

Signature

Date

I understand that SSET wants to keep recordings on file, for auditing purposes. If I don't agree to this, I understand that SSET will delete the recording of this session when the study ends on 28th June 2017.

I do/do not give my consent to the audio-recording being kept for auditing purposes.

Signature

Date

APPENDIX 13

Swansea Service Evaluation Inventory – Women’s Projects V1.4

SWANSEA SERVICE EVALUATION INVENTORY - WOMEN'S PROJECTS

**Developed by Swansea Service Evaluation Team (SSET)
2016**

Designed by Bridget Kerr, Peter Raynor and Pamela Ugwudike

Dimension A:	Description of Service
Dimension B:	Leadership & Innovation
Dimension C:	Quality Assurance
Dimension D:	Staff Attributes & Development
Dimension E:	Assessment Practices & Impact
Dimension F:	Service Characteristics
Dimension G:	Collaborative Working
Dimension H:	Practice Skills

KEY



Interview with **organisational lead**



Interview with **managerial/supervisory staff**



Interview with **practitioner**



Interview with **service-user**



Refer to **documents**



Observe practice

VS = Verification Score

Score on a scale of 1-5, according to the evidence available to verify the scoring for each item.

Dimension A: Description of Service

A draft of Dimension A should be completed based upon pre-evaluation information supplied by the service. This draft should then be shared with all interview participants to clarify that the information included is accurate and representative.

A1. Name of service.

A2. Funding source.

A3. Description of service, eg:

**Community/custodial
Single/multi-agency
Intervention/agency**

A4. Description of setting, eg:

Stand-alone/shared premises

Physical location

Physical set-up

General appearance

--

A5. Referral process (into the service).

--

A6. Staff numbers.

	Male	Female	Total
Full-time			
Part-time			
Total			

A7. Service-user numbers.

Male	
Female	
Total	

A8. Most common needs upon referral.

--

Dimension B: Leadership & Innovation

B1. The service's goals include reducing re-offending, reflect the evidence base for effective work with women (eg, diversion from custody, the provision of social support, family reunification, protection from abusive relationships, recovery from victimisation, etc) and are clearly documented.

☐ Yes☐ No☐ N/A

VS

**DOC**

Mission statement

B2. The service is bound by ethical considerations which are clearly documented.

☐ Yes☐ No☐ N/A

VS

**DOC**

Ethical code

B3. There is evidence of pressing local need for the service.

☐ Yes☐ No☐ N/A

VS

**OL**

How was the need for this service identified?

B4. The service is valued by the system within which it operates (eg, agency, parent organisation, larger system).

☐ Yes ☐ No ☐ N/A

VS



OL

How is the service viewed by, eg, the parent organisation, governing agency, etc?

B5. The service is valued by its stakeholders (organisational lead, managerial/supervisory staff, core practitioners, service-users and external practitioners).

☐ Yes ☐ No ☐ N/A

VS



OL

M/S

PRAC

SU

How do you view the usefulness or importance of this service?

B6. The benefits outweigh the costs of implementing the service for all stakeholders.

☐ Yes ☐ No ☐ N/A

VS



OL

M/S

PRAC

SU

Are there any direct benefits from your involvement with the service? Are there any costs?

B7. The service is sustainable (consider budget, resources, staffing levels, political environment, etc).

☐ Yes ☐ No ☐ N/A

VS



OL

Are there any issues with the budget for the service? (eg, short-term period of funding, budget too small, administration issues, etc). Are there any policy changes or operational issues affecting the future of the service?

B8. There is a clear line of authority.

☐ Yes ☐ No ☐ N/A

VS



OL

M/S

Is there ever any confusion over who has what authority when making policy or organisational decisions? Is there always someone with the authority to make final decisions without the need for extended debate?

PRAC

B9. All major changes/innovations undergo consultation and are implemented incrementally (eg, the service should have been piloted before wider implementation).

☐ Yes ☐ No ☐ N/A

VS



OL

What was the process for setting up the service initially, and what processes are in place for introducing major change to the service?

B10. Staff turnover is low.

☐ Yes ☐ No ☐ N/A

VS



OL

Have the majority of staff been with the organisation for at least two years?

DOC

Staff survey

B11. The organisational lead is committed to the service's mandate and ethical policy.

☐ Yes ☐ No ☐ N/A

VS



OL

Do the service's goals and ethics reflect your own ideas and values?

B12. The organisational lead is engaged with developments in gender-responsive work.

☐ Yes ☐ No ☐ N/A

VS



OL

How do developments in gender-responsive practice and theory influence your leadership of the service?

B13. A working relationship exists between the organisational lead and practitioners.

☐ Yes

☐ No

☐ N/A

VS



OL

What procedures are in place for your regular communication with managerial/supervisory staff and practitioners?

B14. The organisational lead is able to effect change efficiently.

☐ Yes

☐ No

☐ N/A

VS



OL

Can you give me an example of an improvement you've made to the service and how long it took to put in place?

B15. The organisational lead and managerial/supervisory staff refrain from coercion, threat or other authoritarian approaches and instead negotiate with staff to initiate change.

☐ Yes

☐ No

☐ N/A

VS



OL

How do you get staff to implement changes?

M/S

B16. Managerial/supervisory staff are supportive of the organisational lead's vision and use their ideas and values to develop the service.

☐ Yes ☐ No ☐ N/A

VS



M/S

Do you feel that you share the vision of the organisation lead? Does their influence have a direct impact on how you develop the service? Please give examples.

B17. Practitioners are invested in the service.

☐ Yes ☐ No ☐ N/A

VS



M/S

Do you involve practitioners in service development and how?

B18. Practitioners have regular supervision with the service's managerial/supervisory staff.

☐ Yes ☐ No ☐ N/A

VS



PRAC

What procedures are in place for your regular contact with your manager/supervisor?

B19. Stakeholders' (including service-users') interests are actively embraced by the service.

☐ Yes ☐ No ☐ N/A

VS



OL

What input do the various stakeholders have in terms of service development?

B20. The organisation is effective at addressing and resolving problems (eg, staff conflict, communication issues, structural inadequacies, service-user difficulties, etc) in a calm, empathetic and practical way.

☐ Yes ☐ No ☐ N/A

VS



OL

M/S

PRAC

SU

What happens when a problem arises?

B21. Responsibilities for case management are clearly identified and/or assigned and not duplicated by the service.

☐ Yes ☐ No ☐ N/A

VS



OL

M/S

What procedures are in place for effective case management? How do you ensure that responsibilities are not duplicated?

B22. Information is shared across the service through organised channels of communication (eg, memos, conferences, regular reports, etc).

☐ Yes ☐ No ☐ N/A

VS



OL

M/S

How is information communicated across the service?

PRAC

B23. The service is promoted on an ongoing basis through organised channels to interested parties (eg, partners and potential partners, the courts, referring bodies/agencies and potential referrers, community groups, etc).

☐ Yes ☐ No ☐ N/A

VS



OL

How do you ensure that the service is and remains known about, supported and used?

Dimension C: Quality Assurance

C1. Service-users' progress is monitored and recorded.

☐ Yes ☐ No ☐ N/A

VS



OL

How is service-users' progress reviewed?

DOC

Progress monitoring documents

C2. Service-users' feedback is sought, recorded and used to develop the service.

☐ Yes ☐ No ☐ N/A

VS



OL

Do you seek service-user feedback about the quality of the service? How do you do this and how is it used?

DOC

Feedback documents

C3. Data relating to changes in targeted behaviours and human/social needs (eg, acquisition of ETE related skills, change in level of risk, improvements in accommodation, etc) are collated and analysed.

☐ Yes ☐ No ☐ N/A

VS



OL

How do you evaluate changes in targeted behaviours and needs?

DOC

Monitoring documents

C4. Data relating to measures of interest (eg, whether service-users have experienced domestic abuse, whether service-users are diagnosed as having mental health issues, etc) are recorded in a standardised way, to facilitate external and internal evaluator access.

☐ Yes ☐ No ☐ N/A

VS



OL

How do you ensure that evaluators can get the information that they need from your records?

DOC

Monitoring documents

C5. Recidivism figures for service-users are collected twelve months or more after completion of engagement with the service.

☐ Yes ☐ No ☐ N/A

VS



OL

How and at what point do you measure the recidivism rates of service users?

C6. Service effectiveness is measured using a risk-control comparison group at least once every five years.

☐ Yes ☐ No ☐ N/A

VS



OL

How do you use service-user recidivism rates to evaluate effectiveness of the service? How often is this done?

DOC

Evaluation documents

C7. Evaluations are documented, methodologically sound and reports are kept on file.

☐ Yes

☐ No

☐ N/A

VS



DOC

Evaluation reports

C8. Practice is observed regularly to ensure that manual requirements are addressed. (N/A if no aspect of service is manualised).

☐ Yes

☐ No

☐ N/A

VS



OL

How do you ensure that manuals are being properly used?

Dimension D: Staff Attributes & Development

D1. Core service staff (ie, those with case management and supervisory responsibilities) are female.

☐ Yes ☐ No ☐ N/A

VS



DOC

Staff survey

D2. All staff members are Disclosure and Barring Service (DBS) checked.

☐ Yes ☐ No ☐ N/A

VS



OL

What procedures are in place for criminal records checks of service staff?

DOC

Staff records

D3. The organisational lead has post-graduate qualifications in criminal justice or a related field (eg, psychology, social work).

☐ Yes ☐ No ☐ N/A

VS



DOC

Staff survey

D4. The organisational lead has at least three years' frontline experience delivering services such as facilitating group work, supervising service-users, etc (not necessarily women).

☐ Yes

☐ No

☐ N/A

VS



DOC

Staff survey

D5. The organisational lead influences commission of services, hiring of core staff and the provision of training.

☐ Yes

☐ No

☐ N/A

VS



OL

What input do you have in commissioning services, hiring staff and professional development?

D6. The majority of core service staff members have, or are working towards, undergraduate or equivalent qualifications in criminal justice or related fields (eg, social work, psychology, education, substance misuse, etc).

☐ Yes

☐ No

☐ N/A

VS



DOC

Staff survey

D7. The majority of core service staff members have at least two years' experience delivering similar services (not necessarily to women).

☐ Yes ☐ No ☐ N/A

VS



DOC

Staff survey

D8. Hiring and job selection processes involve consideration of relationship skills (empathy, warmth, flexibility, respectfulness, etc), structuring skills (pro-social modelling, cognitive restructuring, motivational interviewing, etc) and enthusiasm for work with women in particular.

☐ Yes ☐ No ☐ N/A

VS



OL

What qualities do you look for in candidates for roles within the service?

D9. All core service staff are conversant in, and committed to, the evidence base for work with women. They must acknowledge the importance of at least 7 of the following:

☐ Yes ☐ No ☐ N/A

VS



- * Safe, women-only provisions
- * Empowerment and development of self-efficacy
- * Reciprocal relationships (staff should be aware of women's experiences of oppressive, abusive and exploitative relationships)
- * Services that address women's pathways (accommodation, ETE, mental and physical health, substance misuse, finance, children & families, experience of domestic abuse, sexual exploitation and sex work)
- * Diversion from custody
- * Holistic and practical work to address social problems
- * Multi-agency work to provide services in health, debt advice and counselling
- * Strengthening pro-social networks
- * Providing practical help with transport and childcare

OL**M/S**

In terms of the practical support offered, ways of working, women's particular needs and the focus for services, what do you think are the important features of work with women?

PRAC

D10. Practitioners are regularly observed (at least once a year) and their skills assessed. Regular supervision (at least bi-monthly) is used to identify any training needs.

☐ Yes☐ No☐ N/A

VS

**M/S**

How is practitioner performance assessed and how are training needs identified?

PRAC**DOC**

Staff records

D11. Staff are engaged with regular, comprehensive training in effective practice (theory and skills) as well as training in the delivery of gender-responsive interventions.

☐ Yes☐ No☐ N/A

VS

**OL****M/S**

What training do staff members receive?

PRAC**DOC**

Staff records

D12. Staff across the service work to provide a supportive, empowering environment for women and clearly believe in the possibility for change.

☐ Yes

☐ No

☐ N/A

VS



OBS

Day-to-day practice

D13. Staff have confidence in the effectiveness of their service.

☐ Yes

☐ No

☐ N/A

VS



M/S

Do you have everything you need (eg, skills, resources, etc) to deliver an effective service?

PRAC

Dimension E: Assessment Practices & Impact

E1. Service-users' risk of re-offending is routinely measured using an assessment tool evidenced to predict the recidivism of women (eg, the Level of Service Inventory - Revised/LSI-R). Service-users are classified into risk categories based on this assessment.

☐ Yes☐ No☐ N/A

VS

**DOC**

Case files/assessment documents

E2. Risk assessment and classification include consideration of dynamic factors in a service-user's life, with a focus on: behaviour; personality pattern; cognition; associates; family/ marital relationships; ETE; leisure and substance use.

☐ Yes☐ No☐ N/A

VS

**DOC**

Case files/assessment documents

E3. Service-users' capabilities/protective factors are assessed as predictors of the reduced likelihood of reoffending or of desistance.

☐ Yes☐ No☐ N/A

VS

**DOC**

Case files/assessment documents

E4. Assessment includes consideration of needs particularly relevant to women (eg, substance misuse, victimisation, poverty, mental health, parenting, etc).

☐ Yes ☐ No ☐ N/A

VS



DOC

Case files/assessment documents

E5. Service-users' risk of re-offending, needs and strengths are regularly re-assessed throughout their engagement with the service.

☐ Yes ☐ No ☐ N/A

VS



DOC

Case files/assessment documents

E6. Service-users' responsivity factors (eg, mental health, childcare, language, culture, literacy level, cognitive ability, etc) are routinely screened using appropriate measures (eg, General Aptitude Test Battery, Beck Depression Inventory, etc).

☐ Yes ☐ No ☐ N/A

VS



DOC

Case files/assessment documents

E7. The assessment and case-planning process is appropriate for the service-user's language, culture, literacy level and cognitive functioning (informed by responsivity assessment).

☐ Yes ☐ No ☐ N/A

VS | 

PRAC

How do you adapt the assessment and case-planning process to fit the service-user?

OBS

Assessment/case-planning session

E8. Service-users are engaged in assessment and case-planning through reciprocal, respectful and empathic practitioner interactions.

☐ Yes ☐ No ☐ N/A

VS | 

PRAC

Can you describe your relationship with service-users during the assessment and case-planning process?

OBS

Assessment/case-planning session

E9. Case-planning is informed by assessment/re-assessment and classification/re-classification to ensure that service-users are referred to/provided services appropriate and useful to them and that they meet intake protocol.

☐ Yes ☐ No ☐ N/A

VS | 

DOC

Case files/referrals/intake documents

E10. The level of *criminal justice* intervention (eg, offence-focused work) provided is congruent with service-users' risk of re-offending classification. High-risk service-users receive the highest level of criminal justice service (targeting of resources, intensity, duration, etc). Low-risk service-users are diverted from criminal justice interventions to supportive generic services (eg, substance misuse agency, trauma recovery programme, housing, health, etc).

☐ Yes

☐ No

☐ N/A

VS



DOC

Case files/referrals/intake documents/process documents

E11. Consideration is given to characteristics that may facilitate or impede an effective working relationship (eg, personal regard, learning/facilitation style, conceptual level, language, culture, etc) when allocating practitioners to service-users.

☐ Yes

☐ No

☐ N/A

VS



M/S

How do you allocate practitioners to service-users?

E12. Service delivery is modified in response to the characteristics of service-users (eg, different learning styles are accommodated, intrinsically motivated service-users are engaged in tasks they enjoy, extrinsically motivated service-users are rewarded for their participation, etc).

☐ Yes

☐ No

☐ N/A

VS



PRAC

How do you respond to differences in the learning styles and characteristics of service-users?

OBS

Service delivery

E13. Inclusion and exclusion criteria are clear, appropriate and enforced.

☐ Yes

☐ No

☐ N/A

VS



OL

What criteria are there for inclusion or exclusion from the service?

M/S

DOC

Process documents

Dimension F: Service Characteristics

F1. The service provides a women-only, safe space.

☐ Yes ☐ No ☐ N/A

VS



M/S

PRAC

Do any men have access to the service (including staff, service-users, relatives, police, maintenance, etc)? How is this managed in terms of maintaining a safe space for women?

DOC

Policy documents

F2. The service facilitates engagement through the provision of childcare and transport.

☐ Yes ☐ No ☐ N/A

VS



M/S

PRAC

Does the service make arrangements for childcare and transport if needed? Please give examples.

SU

F3. Criminal justice services *predominantly* target criminogenic needs, ie, building and/or reinforcing pro-social strengths in service-users': behaviour; personality pattern; cognition; associates; family/marital relationships; ETE; leisure and substance use, *where deficits in these areas have been assessed as linked to risk of re-offending.*

☐ Yes

☐ No

☐ N/A

VS


DOC

Case files/intervention plans/manuals for service-delivery

F4-F6. The service uses behavioural, social learning and cognitive-behavioural strategies where service-users are engaged in interventions focused on transforming behaviour.

M/S
PRAC

What strategies does the service use to encourage and empower service-users to transform their behaviour?

DOC

Service delivery manual

F4. The service uses behavioural strategies (eg, reinforcement systems) to reinforce positive safe, healthy behaviours and decelerate negative, unsafe behaviours. Reinforcers may be tangible (eg, material) and/or social (eg, praise) and/or activities and are individualised where possible. Where relevant, there is a clear written protocol for the administration of tangible or activity reinforcers.

☐ Yes

☐ No

☐ N/A

VS



F5. The service uses social learning strategies (eg, pro-social modelling) to empower service users through teaching, prompting, motivating and confidence-building to enable changed behaviours.

☐ Yes

☐ No

☐ N/A

VS



F6. The service uses cognitive-behavioural strategies (ie, cognitive restructuring and developing cognitive-behavioural coping skills) to guide service-users through changing unhelpful self-talk, problem-solving, skills-building and practicing new behaviours.

☐ Yes ☐ No ☐ N/A

VS



F7. The service uses manuals to provide practitioners with structured guidance and support in delivering interventions.

☐ Yes ☐ No ☐ N/A

VS



DOC

Service delivery manual(s)

F8. Practitioners are interested in and aware of service-users' social contexts, including peer networks, activities, etc.

☐ Yes ☐ No ☐ N/A

VS



PRAC

How much do you know about the daily lives of your service-users and how do you monitor this? Please give examples.

F9. The service refers to and/or provides substance misuse programmes that integrate trauma/victimisation and mental health.

☐ Yes ☐ No ☐ N/A

VS



DOC

Substance misuse programme referral literature, intake protocol, manual, case notes, etc.

F10. The service provides service-users with opportunities to improve their socio-economic conditions through education, training and employment.

☐ Yes ☐ No ☐ N/A

VS | 

PRAC

Are service-users given any support with education, training and employment?

F11. The service promotes and supports healthy relationships with children, families, significant others, and the community.

☐ Yes ☐ No ☐ N/A

VS | 

PRAC

Does the service work with service-users towards their healthy relationships with children, families, significant others, and the community? Please give examples.

F12. There is the opportunity for criminalised service-users to access generic (women-only) services (eg, substance misuse, health, support groups, etc) together with non-criminalised women to normalise their experiences.

☐ Yes ☐ No ☐ N/A

VS | 

PRAC

Are there any opportunities for criminalised service-users to engage with any services alongside women who haven't been criminalised?

F13. Service-users' empowerment and attachment to the service are achieved through their active participation in decision-making relating to components of its structure and delivery.

☐ Yes ☐ No ☐ N/A

VS



M/S

PRAC

SU

Is there a mechanism for service-users to participate in decision-making about the structure and delivery of the service? Has this had an impact on the service? Can you give examples?

F14. Disruptive and/or unsafe behaviours are immediately and consistently dealt with using supportive measures that have been influenced and approved by service-users, are proportionate to the behaviour and appropriate to the individual service-user (eg, verbal warning, coaching, temporary privilege restrictions etc).

☐ Yes ☐ No ☐ N/A

VS



PRAC

How are disruptive or unsafe behaviours dealt with?

SU

OBS

Service delivery

F15. Service-users' success in the service is measurable and visible using criteria such as reduction in risk category, engagement with community services, employment, child reunification, managed substance misuse, etc.

☐ Yes ☐ No ☐ N/A

VS



PRAC

How is the success of service-users measured?

SU

DOC

Case files

F16. The service builds long-term community support networks for service-users

☐ Yes ☐ No ☐ N/A

VS



PRAC

SU

Are there any arrangements made for the ongoing support of service-users after they've left the service?

F17. Relapse prevention strategies involve working with service-users and their support networks (including family and friends) to identify risky situations and triggers and to plan and practice alternative responses.

☐ Yes ☐ No ☐ N/A

VS



PRAC

How are service-users enabled to avoid problematic situations and repeating negative behaviour patterns in future?

F18. Service-users can return to the service to access further support as and when needed.

☐ Yes ☐ No ☐ N/A

VS



M/S

PRAC

Can service-users come back for ongoing support, even when their formal engagement with the service may have ended?

Dimension G: Collaborative Working

G1. The service has established links with a range of relevant projects, agencies and programmes including, eg:

☐ Yes

☐ No

☐ N/A

VS



- * Mental health systems
- * Alcohol and other drug agencies
- * Family service agencies
- * Programmes for survivors of family and sexual violence
- * Emergency accommodation, food and financial assistance programmes
- * Educational organisations
- * Vocation and employment services
- * Healthcare
- * Social services, childcare, etc
- * Transportation
- * Self help groups
- * Consumer advocacy groups
- * Organisations providing leisure and recreational activities
- * Faith-based organisations
- * Women's groups

DOC

Description of service pre-visit document

G2. An holistic and culturally sensitive plan for each service-user draws on a coordinated range of services within her community. Referral processes are clear and documented.

☐ Yes

☐ No

☐ N/A

VS


DOC

Case files, referral documents, etc.

G3. Service staff advocate for individual service-users where necessary. For example, with the Local Authority, housing association, school, etc.

☐ Yes ☐ No ☐ N/A

VS



PRAC

Do you represent service-users in their dealings with authorities and officials, etc?
Please give details

G4. Service staff are in regular communication with external providers.

☐ Yes ☐ No ☐ N/A

VS



M/S

How do you maintain relationships with external providers?

PRAC

G5. Service-users' progress in each collaborating agency is monitored, recorded and reviewed through coordinated case management.

☐ Yes ☐ No ☐ N/A

VS



PRAC

How do you track the progress of service-users across their engagement with different providers?

DOC

Case files (referral outcomes, case notes, etc)

G6. Service-users are signposted to other organisations through readily available culturally and language appropriate literature, posters, etc.

☐ Yes

☐ No

☐ N/A

VS



OBS

Service premises

Dimension H: Practice Skills

Adapted from The Jersey Supervision Interview Checklist
Version 7C May 2009 © The Jersey Crime and Society Project
Designed by Peter Raynor, Pamela Ugwudike & Maurice Vanstone

OBS

Dimension H is scored based on observation of service sessions.

H1-H4. Set up of session.

* Observe whether seating arrangements are such that both/all parties appear relaxed with appropriate distance maintained to ensure freedom of movement.

* Please consider that the nature of interactions may affect seating arrangements. For instance, where the practitioner uses a PC for illustrations during the session, this may alter seating arrangements and levels of eye contact.

H1. Privacy assured to enhance disclosure.

☐ Yes ☐ No ☐ N/A

VS | 

H2. Confidentiality - assured.

☐ Yes ☐ No ☐ N/A

VS | 

H3. Seating - approximate proximity - service-user not crowded or uneasy.

☐ Yes ☐ No ☐ N/A

VS | 

H4. No distractions (or minimal distraction and the practitioner apologises).

☐ Yes ☐ No ☐ N/A

VS | 

H5 - H9. Quality of non-verbal communication.

* Observe the practitioner's body language and apply the SOLER criteria:

The practitioner *squarely* faces the service-user to indicate involvement, maintains an *open posture* by ensuring arms and legs are uncrossed, slightly *leaning forward* to indicate involvement, intermittent *eye contact* is maintained and practitioner appears to be *relaxed*, natural not tense.

* Scoring in this section should reflect the degree to which the practitioner's body language demonstrates attentiveness, interest and the desire to stimulate effective dialogue with the service-user.

H5. Facing the service-user.
☐ Yes ☐ No ☐ N/A

VS

**H6. Open posture / arms legs uncrossed.**
☐ Yes ☐ No ☐ N/A

VS

**H7. Attentive to service-user.**
☐ Yes ☐ No ☐ N/A

VS

**H8. Adequate eye contact**
☐ Yes ☐ No ☐ N/A

VS

**H9. Appears relaxed.**
☐ Yes ☐ No ☐ N/A

VS



H10 - H19. Quality of verbal communication.

When observing the use and style of questions in the session, bear in mind the following:

- * Open questions encourage dialogue and disclosure.
- * Closed questions invite monosyllabic responses; most closed questions attract yes/no responses.
- * Leading questions pressure the service-user to provide a suggested or specific response.
- * Count the number of closed questions, open questions and leading questions and score accordingly, BUT:
 - * Where the closed questions are appropriate, they should not be counted. For example, information-seeking questions and questions to check the service-user's understanding are typically closed questions. These are appropriate and should not be counted.
 - * Where the practitioner appears to be adhering to a service delivery/programme manual, count only questions not contained in the manual.
 - * Observe the service-user's response as well as the behaviour of the practitioner.

H10. Mostly open questions.
☐ Yes ☐ No ☐ N/A

VS

**H11. No leading questions.**
☐ Yes ☐ No ☐ N/A

VS

**H12. Practitioner shows understanding.**
☐ Yes ☐ No ☐ N/A

VS

**H13. Displays warmth (not stiff/cold/formal).**
☐ Yes ☐ No ☐ N/A

VS





**H14. Enthusiastic dialogue.**
☐ Yes ☐ No ☐ N/A

VS

**H15. Practitioner is polite/respectful (eg, not sarcastic, rude, dismissive).**
☐ Yes ☐ No ☐ N/A

VS



H16. Promotes flexible dialogue (eg, does not dominate or interrupt).	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS	
H17. Uses humour to engage.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS	
H18. Optimistic about possibility of change.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS	
H19. There appears to be mutual liking.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS	

H20 - H24. Effective/legitimate use of authority.

Legitimate use of authority requires fairness and a willingness to give the other party the chance to state their case.

Effective use of authority suggests that where the service-user becomes defensive/resistant, the practitioner could:

- * Consider how the service-user's defensiveness/resistance may stem from the probationer's approach/actions/body language.
- * Avoid argument.
- * Avoid an authoritarian stance - warning, criticising, blaming, lecturing, talking/forcing people into change, etc.
- * Aim to develop rapport and empathy by using reflective listening and by using reflective and open questions to change focus. Eg, the question: *"you said...tell me how you feel about that"* may discourage resistance by redirecting focus/changing track and encouraging the service-user to re-evaluate the discrepancies in their line of reasoning.
- * Highlight the service-user's control over their actions/disclosure.
- * Ensure that the service-user's problems, views, concerns are taken into account during decision making.

Scores in this section should also reflect:

- * The practitioner's response to defensiveness/resistance (consider the practitioner's verbal and non-verbal behaviour).
- * The practitioner's decision-making approach (is it collaborative?).
- * The practitioner's clarification of roles and responsibilities.

H20. Does not argue / 'changes track' with reflective question.

☐ Yes ☐ No ☐ N/A

VS | 

H21. Encourages collaboration during decision making.

☐ Yes ☐ No ☐ N/A

VS | 

H22. Positive comments outweigh negative.

☐ Yes ☐ No ☐ N/A

VS | 

H23. Firm but fair.

☐ Yes ☐ No ☐ N/A

VS | 

H24. Clarifies roles/responsibilities.

☐ Yes ☐ No ☐ N/A

VS | 


H25 - H33. Motivational Interviewing.

In completing this section, the observer should consider the following components identified in the Motivational Interviewing literature:

- * *Showing empathy* - involves using comments that demonstrate genuine understanding. Comments such as "...that must have been difficult for you" demonstrate empathy. Reflective listening and attending to the service-user also demonstrate empathy.
- * *Developing discrepancies* - involves highlighting the difference between the service-user's 'current state' and 'desired state'.

* *Rolling with resistance* - entails avoiding arguments by using reflections and open questions.
 * *Developing self-efficacy* - involves reassuring service-users of their ability to repeat past successes.

Observe whether the practitioner uses Motivational Interviewing skills to identify the service-user's location on cycle of change, overcome resistance and stimulate the change process.

H25. Paraphrases, nods, maintains eye contact.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H26. Makes empathic comments.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H27. Avoids argument / rolls with resistance.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H28. Uses reflections / develops discrepancies.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H29. Uses reflections to counter resistance or improve understanding.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H30. Promotes self-efficacy.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H31. Adapts approach to the service-user's location on cycle of change.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H32. Elicits self-motivating comments.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 

H33. Service-user becomes less resistant as interview progresses.

☐ Yes ☐ No ☐ N/A

VS



H34 - H38. Pro-social modelling.

Pro-social modelling skills can be used in a structured manner in order to guide the service-user towards replacing anti-social behaviours/attitudes/thinking with pro-social alternatives.

Effective modelling occurs where the officer 'concretely or vividly demonstrates the pro-social behaviour'.

Effective praise/affirmation should refer to specific behaviour/attitudes/thinking.

Effective challenging involves subtle but firm disapproval of anti-social attitudes/behaviours, examples:

- * Highlighting risky behaviour and its consequences.
- * Discouraging rationalisations/refusing to collude with the service-user.
- * Effective challenge should also refer to the specific behaviour and should not be entirely negative.
- * Observe whether the practitioner includes positive feedback, highlights reasons for disapproval and provides an invitation to the service-user to consider the inappropriateness of the anti-social behaviour (encouraging 'self-challenge').

H34. Several examples of modelling.

☐ Yes ☐ No ☐ N/A

VS






H35. Several examples of praise.

☐ Yes ☐ No ☐ N/A

VS





H36. Praise refers to specific behaviour or thinking.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A	VS 
H37. Challenges anti-social behaviour or thinking in a positive way (eg, emphasises strengths). Not confrontational or over-critical.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A	VS 
H38. Service-user is encouraged to practice more pro-social behaviour/thinking.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A	VS 









H39 - H48. Problem solving.

- * If assessment session: Is an adequate assessment instrument in use?
- * If compliance meeting: Particular attention should be paid to evidence of effective use of authority/pro-social modelling/Motivational Interviewing.
- * If programme session: Is the service-user actively involved/participating?

Observe whether the focus is on criminogenic needs, for example:

- * Accommodation, employment/education, substance misuse, attitude, family relationship, financial, emotional stability/mental health, anti-social peers, recreation/social, health, crime-prone personality traits, anti-social behaviour/attitudes/thinking, anti-social associates, family/marital problems.

H39. Practitioner identifies evident need(s).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A	VS 
H40. Focus is on service-user's assessment of problem(s).	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A	VS 

H41. Focus is on criminogenic needs.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H42. Plans/goals/actions/options discussed, evaluated and agreed.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H43. Target(s) set.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H44. Solution-focused.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H45. Optimistic about possibility of change.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H46. Acts as advocate/makes referral where appropriate.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H47. Provides details of access to referral agency where appropriate referral has been made.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H48. Discusses benefit of referral where appropriate referral has been made.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 

H49 - H55. Cognitive restructuring.

In this section, observe how the practitioner engages with the service-user's attitudes and thinking - for example, is the service-user encouraged/enabled to report thinking, attitudes, feelings in relation to a particular problem? Are alternative, less risky or offence-prone ways of thinking identified and practiced? Learning new ways of thinking to some extent resembles learning other skills, and the basic steps of skill acquisition are likely to be relevant, for example:

- * Defining the skill,
- * Modelling or demonstrating the skill,
- * Providing the service-user the opportunity to practice the skill - eg, in role play,
- * Evaluating performance and providing feedback,
- * Repetition.

H49. Practitioner identifies anti-social thinking.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H50. Suggests alternatives to anti-social thinking.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H51. Models alternative thinking.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H52. Encourages service-user to practice alternative thinking.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H53. Service-user has opportunity to practice alternative thinking.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H54. Discusses costs of anti-social thinking.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 
H55. Discusses the benefits of the alternative thinking.	<input type="radio"/> Yes	<input type="radio"/> No	<input type="radio"/> N/A	VS 

H56 - H63. Overall structure of session.

The observer should consider whether the session is structured appropriately to ensure the effective involvement or participation of the service-user, bearing in mind the nature and purpose of the session.

H56. Summary of previous work provided.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A VS 
H57. Focus of session is identifiable.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A VS 
H58. Identifiable beginning, middle and end.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A VS 
H59. Service-user engaged in proceedings.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A VS 
H60. Practitioner sums up/provides feedback.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A VS 
H61. Arrangements made for next session.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A VS 
H62. Tasks given for the interim.	<input type="radio"/> Yes <input type="radio"/> No <input type="radio"/> N/A VS 

H63. Good quality overall relationship.

☐ Yes

☐ No

☐ N/A

VS



BIBLIOGRAPHY

Andrews, D. A. & Bonta, J. (2010). *The Psychology of Criminal Conduct*, 5th edition. New Providence, NJ: Anderson.

Andrews, D. A., Bonta, J. & Hoge, R. (1990b). Classification for Effective Rehabilitation. *Criminal Justice and Behavior*, 17, 19–52.

Andrews, D. A., Bonta, J. & Wormith, J. S. (2011). The Risk-Need-Responsivity (RNR) Model: Does Adding the Good Lives Model Contribute to Effective Crime Prevention? *Criminal Justice and Behaviour*, 38, 735-755.

Andrews, D. A. & Dowden, G. (2006). Risk Principle of Case Classification in Correctional Treatment: A Meta-Analytic Investigation. *International Journal of Offender Therapy and Comparative Criminology*, 50, 88-100.

Andrews, D. A., Guzzo, L., Raynor, P., Rowe, R. C., Rettinger, J., Brews, A. & Wormith, S. J. (2012). Are the Major Risk/Need Factors Predictive of Both Female and Male Reoffending? A Test With the Eight Domains of the Level of Service/Case Management Inventory. *International Journal of Offender Therapy and Comparative Criminology*, 56, 113-133.

Andrews, D. A., & Kiessling, J. J. (1980). Program Structure and Effective Correctional Practices: A Summary of the CaVIC Research. In R.R. Ross & P. Gendreau (Eds.), *Effective Correctional Treatment* (pp. 441-463). Toronto, Canada: Butterworth.

Andrews, D. A. & Wormith, J. S. (1989). Personality and Crime: Knowledge Destruction and Construction in Criminology. *Justice Quarterly*, 6, 289-309.

Andrews, D. A., Zinger, I., Hoge, R. D., Bonta, J., Gendreau, P. & Cullen, F. T. (1990a). Does Correctional Treatment Work? A Clinically Relevant and Psychologically Informed Meta-Analysis. *Criminology*, 28, 369-404.

Annison, J., Brayford, J. & Deering, J. (2015). *Women and Criminal Justice: From the Corston Report to Transforming Rehabilitation*. Bristol: Policy Press.

Armstrong, R., Gelsthorpe, L. & Crewe, B. (2014). From Paper Ethics to Real-World Research: Supervising Ethical Reflexivity When Taking Risks in Research with ‘the Risky’. In K. Lumsden & A. Winter (Eds.) *Reflexivity in Criminological Research: Experiences with the Powerful and the Powerless* (pp. 207-219). Basingstoke: Palgrave Macmillan.

Ashley, O. S., Marsden, M. E. & Brady, T. M. (2003). Effectiveness of Substance Abuse Treatment Programming for Women: A Review. *American Journal of Drug and Alcohol Abuse*, 29, 19-53.

Austin, J., Bloom, B. & Donahue, T. (1992). *Female Offenders in the Community: An Analysis of Innovative Strategies and Programs*. Washington, DC: National Institute of Corrections.

Backer, T. E., David, S. L. & Saucy, G. (1995). *Reviewing the Behavioral Science Knowledge Base on Technology Transfer*. Rockville, MD: U.S. Department of Health and Human Services.

Baldwin, L. & Epstein, R. (2017). *Short But Not Sweet: A Study of the Impact of Custodial Sentences on Mothers and their Children*. Leicester: De Montfort University.

Bandura, A. (1994). Self-Efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of Human Behavior, Vol. 4* (pp. 71-81). New York: Academic Press.

Bandura, A. (1997). *Self Efficacy: The Exercise of Control*. New York: W. H. Freeman and Company.

Bardens, J. & Grimwood, G. G. (2013). *Introducing "Payment by Results" in Offender Rehabilitation and Other Reforms*. House of Commons briefing paper. London: House of Commons Library.

Bauman, L. J., Stein, R. E. K. & Reys, H. T. (1991). Reinventing Fidelity: The Transfer of Social Technology Among Settings. *American Journal of Community Psychology*, 19, 619-639.

Bauwens, A. (2010). The Use of Method Triangulation in Probation Research. *European Journal of Probation*, 2, 39-52.

Bergold, J. & Thomas, S. (2012). Participatory Research Methods: A Methodological Approach in Motion. *Forum: Qualitative Social Research*, 13. Available at: <http://www.qualitative-research.net/index.php/fqs/article/view/1801/3334>

Bernfeld, G. A., Farrington, D. P. & Leschied, A. W. (2001). *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs*. Chichester: Wiley.

Birkett, G. (2017). Transforming Women's Rehabilitation? An Early Assessment of Gender-Specific Provision in Three Community Rehabilitation Companies. Available at: <http://openaccess.city.ac.uk/18458/3/Transforming%20Women%2527s%20Rehabilitation.pdf>

Blanchette, K. & Brown, S. L. (2006). *The Assessment and Treatment of Women Offenders: An Integrative Perspective*. Chichester: John Wiley & Sons, Ltd.

Blanchette, K. & Eldjupovic-Guzina, G. (1998). *Results of a Pilot Study of the Peer Support Program for Women Offenders*. Ottawa, Ontario: Correctional Service Canada.

Bloom, B. (1999). Gender Responsive Programming for Women Offenders: Guiding Principles and Practices. *Forum on Corrections Research*, 11, 22-27.

Bloom, B. (2000). Beyond Recidivism: Perspectives on Evaluation of Programs for Female Offenders in Community Corrections. In M. McMahon (Ed.) *Assessment to Assistance: Programs for Women in Community Corrections* (pp. 107-138). Arlington: American Correctional Association.

Bloom, B. & Covington, S. (2000). *Gendered Justice: Programming for Women in Correctional Settings*. Paper presented at the 52nd Annual Meeting of the American Society of Criminology, San Francisco, CA. Available at: <http://www.stephaniecovington.com/assets/files/11.pdf>

Bloom, B. & Covington, S. (2001). *Effective Gender-Responsive Interventions in Juvenile Justice: Addressing the Lives of Delinquent Girls*. Paper presented at the Annual Meeting of the American Criminology Association, Atlanta, Georgia.

Bloom, B., Owen, B. & Covington, S. (2002). *Gender Responsive Strategies: Research, Practice and Guiding Principles for Women Offenders*. U.S. Department of Justice, National Institute of Corrections. Available at: <https://s3.amazonaws.com/static.nicic.gov/Library/018017.pdf>

Bloom, B., Owen, B. & Covington, S. S. (2003). *Gender-Responsive Strategies for Women Offenders: Research, Practice, and Guiding Principles for Women Offenders*. Available at: <http://www.nicic.org>.

Bloom, B., Owen, B. & Covington, S. (2005). A Summary of Research, Practice and Guiding Principles for Women Offenders. *Gender-Responsive Strategies for Women Offenders, Bulletin 1*. U.S. Department of Justice, National Institute of Corrections. Available at: www.nicic.org

Bonta, J. & Andrews, D. (2010). Viewing Offender Assessment and Rehabilitation Through the Lens of the Risk-Needs-Responsivity Model. In F. McNeill, P. Raynor & C. Trotter (Eds.) *Offender Supervision: New Directions in Theory, Research and Practice* (pp. 19-40). Cullompton: Willan.

Bonta, J. & Andrews, D. A. (2017). *The Psychology of Criminal Conduct*, 6th edition. New York: Routledge.

Bonta, J., Blais, J. & Wilson, H. A. (2014). A Theoretically Informed Meta-Analysis of the Risk for General and Violent Recidivism for Mentally Disordered Offenders. *Aggression and Violent Behavior*, 19, 278-287.

Bonta, J., Bourgon, G., Rugge, T., Gress, C. & Gutierrez, L. (2013). Taking the Leap: From Pilot Project to Wide-Scale Implementation of the Strategic Training Initiative in Community Supervision (STICS). *Justice Research and Policy*, 15, 17-35.

- Bonta, J., Bourgon, G., Rugge, T., Scott, T. Yessine, A. K., Gutierrez, L., & Li, J. (2010). *The Strategic Training Initiative in Community Supervision: Risk-Need-Responsivity in the Real World*. User Report 2010-01. Ottawa: Public Safety and Emergency Preparedness Canada.
- Bonta, J., Bourgon, G., Rugge, T., Scott, T. Yessine, A. K., Gutierrez, L., & Li, J. (2011). An Experimental Demonstration of Training Probation Officers in Evidence-Based Community Supervision. *Criminal Justice and Behaviour*, 38, 1127-1148.
- Bonta, J., Rugge, T., Scott, T., Bourgon, G. & Yessine, A. (2008). Exploring the Black Box of Community Supervision. *Journal of Offender Rehabilitation*, 47, 248-270.
- Booth, L. (2012). *Effective Correctional Programs for Women Offenders*. Public Safety Canada research report No. R-279. Available at:
<https://www.publicsafety.gc.ca/lbrr/archives/cn21543-eng.pdf>
- Bottoms, A. & McWilliams, W. (1979). A Non-Treatment Paradigm for Probation Practice. *British Journal of Social Work*, 9, 159-202
- Bourgon, G. & Bonta, J. (2014). Reconsidering the Responsivity Principle: A Way to Move Forward. *Federal Probation*, 78, 3-10.
- Bourgon, G., Bonta, J., Rugge, T. & Gutierrez, L. (2010). Technology Transfer: The Importance of Ongoing Clinical Supervision in Translating ‘What Works’ to Everyday Community Supervision. In F. McNeill, P. Raynor & C. Trotter (Eds.) *Offender Supervision: New Directions in Theory, Research and Practice* (pp. 91-112). Oxon: Willan Publishing.
- Brimblecombe, N., Hu, B. & Stevens, M. (2014). *Tales from Social Care*. London: Personal Social Services Research Unit, London School of Economics and Political Science.
- British Society of Criminology. (2015). *Statement of Ethics 2015*. Available at:
<http://www.britsoccrim.org/new/?q=node/22>
- Bryman, A. (2012). *Social Research Methods*, 4th edition. Oxford: Oxford University Press.
- Bryman, A. & Burgess, R.G. (Eds.) (1994). *Analyzing Qualitative Data*. London: Routledge.
- Burgess, R. (1984). *In the Field: An Introduction to Field Research*. London: Allen and Unwin.
- Burnett, R. and McNeill, F. (2005). The Place of the Officer–Offender Relationship in Assisting Offenders to Desist from Crime. *Probation Journal*, 52, 247–68.

- Bushe, G. R. (2007). Appreciative Inquiry Is Not (Just) About the Positive. *OD Practitioner*, 39, 30-35.
- Campbell, D. & Fiske, D. (1959). Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, 56, 81-105.
- Canadian Evaluation Society. (2010). *Competencies for Canadian Evaluation Practice*. Available at:
https://evaluationcanada.ca/txt/2_competencies_cdn_evaluation_practice.pdf
- Cann, J. (2006). *Cognitive Skills Programmes: Impact on Reducing Reconviction Among a Sample of Female Prisoners*. Home Office Research Findings No. 276. London: Home Office.
- Cann, J., Falshaw, L., Nugent, F. & Friendship, C. (2003). *Understanding What Works: Accredited Cognitive Skills Programmes for Adult Men and Young Offenders*. Home Office Research Findings No. 226. London: Home Office.
- Canton, R. (2018). Probation and the Philosophy of Punishment. *Probation Journal*, 65, 252-268.
- Carter, P. (2003). *Managing Offenders, Reducing Crime: A New Approach*. London: Home Office.
- Casper, J. & Brereton, D. (1984). Evaluating Criminal Justice Reforms. *Law and Society Review*, 18, 121-143.
- Chadwick, N., Dewolf, A. & Serin, R. (2015). Effectively Training Community Supervision Officers: A Meta-Analytic Review of the Impact on Offender Outcome. *Criminal Justice and Behavior*, 42, 977-989.
- Cherryholmes, C. H. (1992). Notes on Pragmatism and Scientific Realism. *Educational Researcher*, 21, 13-17.
- Chesney-Lind, M. (2000). What To Do About Girls? Thinking About Programs for Young Women. In M. McMahon (Ed.) *Assessment to Assistance: Programs for Women in Community Corrections* (pp. 139-170). Arlington: American Correctional Association.
- Clancy, A., Hudson, K., Maguire, M., Peake, R., Raynor, P., Vanstone, M. & Kynch, J. (2006). *Getting Out and Staying Out: Results of the Prisoner Resettlement Pathfinders*. Bristol: Policy Press.
- Clavert, S. J., Beutler, L. E. & Crago, M. (1988). Psychotherapy Outcome as a Function of Therapist-Patient Matching on Selected Variables. *Journal of Social and Clinical Psychology*, 6, 104-117.

- Clinks. (2016). *A Guide to Service User Involvement and Co-Production*. London: Clinks. Available at: <http://www.revolving-doors.org.uk/file/1857/download?token=cDW0pd6M>
- Clinks. (2018). *Clinks Response to the HM Inspectorate of Probation's 2018 Consultation on Inspection Framework and Programmes*. London: Clinks. Available at: https://www.clinks.org/sites/default/files/2018-10/clinks_response_hmip-inspection-framework.pdf
- Cohen, S. (1980). *Folk Devils and Moral Panics*, 2nd edition. Oxford: Martin Robertson.
- Cohen, S. (1985). *Visions of Social Control: Crime, Punishment and Classification*. Cambridge: Polity Press.
- Cohen, S. & Taylor, L. (1981). *Psychological Survival: The Experience of Long-Term Imprisonment*, 2nd edition. Harmondsworth: Penguin.
- Cohn, E. G. & Farrington, D. P. (1990). Differences Between British and American Criminology: An Analysis of Citations. *British Journal of Criminology*, 30, 467-482.
- Commission on Justice in Wales. (2018). *Commission on Justice in Wales: Call for Evidence*. Cardiff: Commission on Justice in Wales. Available at: <https://beta.gov.wales/sites/default/files/publications/2018-02/180227-commission-justice-wales-call-for-evidence.pdf>
- Cooperrider, D. L. & Srivastva, S. (1987). Appreciative Inquiry in Organizational Life. In W. A. Pasmore & R. W. Woodman (Eds.) *Research in Organizational Change and Development, Vol. 1* (pp. 129-169). Greenwich, CT: JAI Press.
- Cornwall, A. & Jewkes, R. (1995). What is Participatory Research? *Social Science and Medicine*, 41, 1667-1676
- Corston, J. (2007). *The Corston Report: A Report by Baroness Corston of a Review of Women with Particular Vulnerabilities in the Criminal Justice System*. London: Home Office.
- Covington, S. S. (2000). Helping Women Recover: Creating Gender-Specific Treatment for Substance-Abusing Women and Girls in Community Corrections. In M. McMahon (Ed.) *Assessment to Assistance: Programs for Women in Community Corrections* (pp. 171-233). Arlington: American Correctional Associations.
- Covington, S. (2007). *Women and Addiction: A Gender-Responsive Approach*. Center City, MN: Hazelton.
- Covington, S. (2018). *The TIER System: A Trauma Informed Effective Reinforcement System for Residential Facilities*. Available at: <https://www.stephaniecovington.com/tier-system.php>

- Covington, S. S. & Bloom, B. E. (1999). *Gender-Responsive Programming and Evaluation for Women in the Criminal Justice System: A Shift from 'What Works?' to 'What is the Work?'* Paper presented at the 51st Annual Meeting of the American Society of Criminology, Toronto, Canada.
- Covington, S. S. & Bloom, B. E. (2003). Gendered Justice: Women in the Criminal Justice System. In B. E. Bloom (Ed.) *Gendered Justice: Addressing Female Offenders* (pp. 3-23). Durham: Carolina Academic Press.
- Covington, S. S. & Bloom, B. E. (2007). Gender-Responsive Treatment and Services in Correctional Settings. *Women and Therapy*, 29, 9-33.
- Covington, S. S. & Bloom, B. (2008; 2017). *Gender-Responsive Program Assessment Tool (Criminal Justice Version)*. Available at: http://www.centerforgenderandjustice.org/assets/files/5796-65-genderresponsiveassessment_corrections.pdf
- Cowburn, M. (2010). Principles, Virtues and Care: Ethical Dilemmas in Research with Male Sex Offenders. *Psychology, Crime and Law*, 16, 65-74.
- Cowe, F., Deering, J. & Vanstone, M. (2007). From Eclecticism to Orthodoxy in Practice: Can Theory and Practice Merge in Probation? In S. R. Smith (Ed.) *Applying Theory to Policy and Practice: Issues for Critical Reflection* (pp. 37-52). Oxon: Routledge.
- Creswell, J. W. (2002). *Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research*. Upper Saddle River, NJ: Pearson Education.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative & Mixed Methods Approaches*. Thousand Oaks, CA: Sage Publications.
- Cullen, F. T. & Gendreau, P. (2000). Assessing Correctional Rehabilitation: Policy, Practice, and Prospects. *Criminal Justice* 2000, 3, 109-175.
- Cullen, F. T. & Gendreau, P. (2001). From Nothing Works to What Works: Changing Professional Ideology in the 21st Century. *The Prison Journal*, 81, 313-338.
- Dahlgren, L. & Willander, A. (1989). Are Special Treatment Facilities for Female Alcoholics Needed? A Controlled 2 Year Followup Study from a Specialized Female Unit (EWA) Versus a Mixed Male/Female Treatment Facility. *Alcoholism: Clinical and Experimental Research*, 13, 499-505.
- Davies, J. T. & Peters, E. (2014). Relationships Between Gatekeepers and Researchers: The Experiences of Conducting Evaluations into Parenting Programmes in Community and Penal Settings. In K. Lumsden & A. Winter (Eds.) *Reflexivity in Criminological Research: Experiences with the Powerful and the Powerless* (pp. 35-46). Basingstoke: Palgrave Macmillan.

- Day, A., Casey, S., Ward, T., Howells, K. & Vess, J. (2010). *Transitions to Better Lives: Offender Readiness and Rehabilitation*. Cullompton: Willan.
- Day, J. C., Zahn, M. A. & Tichavsky, L. P. (2014). What Works and for Whom? The Effects of Gender Responsive Programming on Girls and Boys in Secure Detention. *Journal of Research in Crime & Delinquency*, 52, 93-129.
- Deering, J. (2011). *Probation Practice and the New Penology: Practitioner Reflections*. Aldershot: Ashgate.
- Denzin, N. (1989). *The Research Act, 3rd edition*. Englewood Cliffs, NJ: Prentice Hall.
- Dey, I. (1993). *Qualitative Data Analysis: A User-Friendly Guide for Social Scientists*. London: Routledge.
- Disley, E., Giacomantonio, C., Kruithof, K. & Sim, M. (2015). *The Payment by Results Social Impact Bond Pilot at HMP Peterborough: Final Process Evaluation Report*. London: Ministry of Justice.
- Dolowitz, D. P. & Marsh, D. (2000). Learning from Abroad: The Role of Policy Transfer in Contemporary Policy-Making. *Governance: An International Journal of Policy and Administration*, 13, 5-24.
- Dosher, A. & Terry, J. (1999). An International Evolutionary Design for Evaluation of Community Development. In *New Designs for Youth Development*. Washington, DC: National Network for Youth.
- Dowden, C. (2005). *What Works for Women Offenders? A Meta-Analytic Exploration of Gender-Responsive Treatment and their Role in the Delivery of Effective Correctional Intervention*. Paper presented at What Works with Women Offenders: A cross-national dialogue about effective responses to female offenders, Prato, Italy.
- Dowden, C. & Andrews, D.A. (1999a). What Works for Female Offenders: A Meta-Analytic Review. *Crime and Delinquency*, 45, 438-452.
- Dowden, C. & Andrews, D.A. (1999b). What Works in Young Offender Treatment: A Meta- Analysis. *Forum on Corrections Research*, 11, 21-24.
- Dowden, C. & Andrews, D.A. (2000). Effective Correctional Treatment and Violent Reoffending: A Meta-Analysis. *Canadian Journal of Criminology*, 42, 449-467.
- Dowden, C. & Andrews, D. A. (2004). The Importance of Staff Practice in Delivering Effective Correctional Treatment: A Meta-Analytic Review of Core Correctional Practice. *International Journal of Offender Therapy and Comparative Criminology*, 48, 203-214.

Duwe, G. & Clark, V. (2015). Importance of Program Integrity: Outcome Evaluation of a Gender-Responsive, Cognitive-Behavioral Program for Female Offenders. *Criminology & Public Policy*, 14, 1-26.

Dweck, C. (2014). *The Power of Believing that You Can Improve*. Presentation at TEDx Norrköping. Available at: https://www.ted.com/talks/carol_dweck_the_power_of_believing_that_you_can_improve

Dweck, C. (2017). *Mindset*. London: Robinson.

Earle, J; Nadin, R; and Jacobson, J. (2014). *Brighter Futures: Working Together to Reduce Women's Offending*. London: Prison Reform Trust.

Easton, S. & Piper, C. (2016). *Sentencing and Punishment: The Quest for Justice*, 4th edition. Oxford: Oxford University Press.

Ebong, B. (2004). *The Use of Indigenous Techniques of Communication for Language Learning" the Case of Cameroon*. Available at: http://monarch.qucosa.de/fileadmin/data/qucosa/documents/4863/data/ebong_indigenoustechniques.pdf

Economic and Social Research Council. (2015). *ESRC Framework for Research Ethics, Updated January 2015*. Available at: http://www.esrc.ac.uk/_images/framework-for-research-ethics_tcm8-33470.pdf

Egan, G. (2002). *The Skilled Helper*, 7th edition. Pacific Grove, CA: Brooks Cole.

Ellickson, P. & Petersilia, J. (1983). *Implementing New Ideas in Criminal Justice*. Santa Monica, CA: RAND.

Elliott, D. S., Huizinga, D. & Ageton, S. S. (1985). *Explaining Delinquency and Drug Use*. Beverly Hills, CA: Sage.

Erikson, K. T. (1962). Notes on the Sociology of Deviance. *Social Problems*, 9, 307-314.

Falshaw, L., Friendship, C., Travers, R. & Nugent, F. (2003). *Searching for 'What Works': An Evaluation of Cognitive Skills Programmes*. Home Office Research Findings No. 206. London: Home Office.

Farrall, S. (2002). *Rethinking What Works With Offenders*. Cullompton: Willan.

Farrall, S. (2004). Social Capital, Probation Supervision and Desistance from Crime. In S. Maruna and R. Immarigeon (Eds.) *After Crime and Punishment: Ex-Offender Reintegration and Desistance from Crime* (pp. 57-82). Cullompton: Willan.

- Farrall, S. and Calverley, A. (2006). *Understanding Desistance from Crime: Theoretical Directions in Resettlement and Rehabilitation*. London: Open University Press.
- Flick, U. (2002). *An Introduction to Qualitative Research, 2nd edition*. London: Sage Publications.
- Folkard, M. S., Smith, D. E. & Smith, D. D. (1976). *IMPACT Volume II: The Results of the Experiment*. Research Study 36. London: Her Majesty's Stationery Office
- Friendship, C., Blud, L., Erikson, M. & Travers, R. (2002). *An Evaluation of Cognitive Behavioural Treatment for Prisoners*. Home Office Research Findings 161. London: Home Office.
- Funk, S. J. (1999). Risk Assessment for Juveniles on Probation: A Focus on Gender. *Criminal Justice and Behaviour*, 26, 44-68.
- Gallagher, C.A., Wilson, D.B., & MacKenzie, D.L. (2000). *A Meta-Analysis of the Effectiveness of Sex Offender Treatment Programs*. Unpublished paper. University of Maryland, College Park, MD.
- Gauke, D. (2018). *Press Release: Justice Secretary Outlines Future Vision for Probation*. London: Ministry of Justice, National Probation Service and Her Majesty's Prison and Probation Service. Available at: <https://www.gov.uk/government/news/justice-secretary-outlines-future-vision-for-probation>
- Gelsthorpe, L. & Hedderman, C. (2012). Providing for Women Offenders: the Risk of Adopting a Payment by Results Approach. *Probation Journal*, 59, 374-390.
- Gelsthorpe, L. & Padfield, N. (2018). Editorial: Female Offender Strategy. *Criminal Law Review*, 9, 693-694.
- Gelsthorpe, L., Sharpe, G. & Roberts, J. (2007). *Provision for Women Offenders in the Community*. London: Fawcett Society.
- Gelsthorpe, L. & Wright, S. (2015). The Context: Women as Lawbreakers. In J. Annison, J. Brayford & J. Deering. *Women and Criminal Justice: From the Corston Report to Transforming Rehabilitation*. Bristol: Policy Press.
- Gendreau, P. (1996a). The Principles of Effective Interventions with Offenders. In A. T. Harland (Ed.) *Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply* (pp. 117-130). Thousand Oaks, CA: Sage
- Gendreau, P. (1996b). Offender Rehabilitation: What We Know and What Needs to Be Done. *Criminal Justice and Behavior*, 23, 144-161.

- Gendreau, P. & Andrews, D. A. (1979). Psychological Consultation in Correctional Agencies: Case Studies and General Issues. In J. J. Platt & R. J. Wicks (Eds.) *The Psychological Consultant* (pp.177-212). New York: Grune & Stratton.
- Gendreau, P., Andrews, D. A. & Thériault, Y. (2010). *Correctional Program Assessment Inventory – 2010 (CPAI-2010): Scoring Manual and Interview Guide*. Ottawa, ON: Available from paulgend@bell.net or yvetheri@yahoo.ca
- Gendreau, P., French, S. A. & Gionet, A. (2004). What Works (What Doesn't Work): The Principles of Effective Correctional Treatment. *Journal of Community Corrections*, 13, 4-30.
- Gendreau, P., French, S. A. & Taylor, A. (2002a). *What Works (What Doesn't Work), revised 2002*. Invited submission to the International Community Corrections Association Monograph Series Project. Reproduced in E. J. Latessa, S. J. Listwan & E. J. Latessa. (2013). *What Works (and Doesn't) in Reducing Recidivism*. Cincinnati, OH: Anderson.
- Gendreau, P. & Goggin, C. (1991). *Evaluation of Correctional Service of Canada Substance Abuse Programs*. Research Report No. 16. Ottawa, Ontario: Research and Statistics Branch, Correctional Service of Canada.
- Gendreau, P., Goggin, C., Cullen, F. T. & Paparozzi, M. (2002b). The Common-Sense Revolution and Correctional Policy. In J. McGuire (Ed.) *Offender Rehabilitation and Treatment: Effective Programmes and Policies to Reduce Re-offending* (pp. 359-386). Chichester: John Wiley & Sons.
- Gendreau, P., Goggin, C. & Smith, P. (1999). The Forgotten Issue in Effective Correctional Treatment: Program Implementation. *International Journal of Offender Therapy and Comparative Criminology*, 43, 180-187.
- Gendreau, P., Goggin, C. & Smith, P. (2001). Implementation Guidelines for Correctional Programs in the "Real World". In G. A. Bernfeld, D. P. Farrington & A. W. Leschied (Eds.) *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs* (pp.247-268). Chichester: John Wiley & Sons, Ltd.
- Gendreau, P., & Ross, R. R. (1979). Effective Correctional Treatment: Bibliotherapy for Cynics. *Crime & Delinquency*, 1979, 25, 463-489.
- Gendreau, P. & Thériault, Y. (2010). *Correctional Program Assessment-Inventory – 2010 Administration and Scoring Protocol*. Ottawa, ON: Available from paulgend@bell.net or yvetheri@yahoo.ca
- Giordano, P. C., Cernokovich, S. A. & Rudolph, J. L. (2002). Gender, Crime and Desistance: Toward a Theory of Cognitive Transformation. *American Journal of Sociology*, 107, 990-1064.

- Glaser, B. G. & Strauss, A. L. (2010). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New Brunswick: Aldine Transaction.
- Glazer, S. (2004). Lost in Translation. *New York Times Book Review*, 22 August 2004, 13.
- Gobeil, R., Blanchette, K. & Stewart, L. (2016). A Meta-Analytic Review of Correctional Interventions for Women Offenders. *Criminal Justice and Behavior*, 43, 301-322.
- Gomm, R. (2013). What Will ‘Count’ and be Transformed for Women in the Criminal Justice System? *British Journal of Community Justice*, 11, 153-157.
- Grattet, R., Janetta, J. & Lin, J. (2006). *Evidence-Based Practices in Corrections: A Training Manual for the California Program Assessment Process (CPAP)*. Office of Research, California Department of Corrections and Rehabilitation and Center for Evidence Based Corrections, University of California, Irvine.
- Gray, G. A. (1997). *Does Coercion Play a Significant Role in Community Treatment Programs that Reduce Offender Recidivism?* University of New Brunswick. Unpublished Master Thesis.
- Grieger, L. & Hosser, D. (2014). Which Risk Factors Are Really Predictive? An Analysis of Andrews and Bonta’s “Central Eight” Risk Factors for Recidivism in German Youth Correctional Facility Inmates. *Criminal Justice and Behavior*, 41, 613-634.
- Grierson, J. (2018). Former Chief Inspector to Chair Labour Review of Probation. *The Guardian*. Available at: <https://www.theguardian.com/society/2018/jul/08/former-uk-chief-inspector-to-chair-labour-review-of-probation>
- Gutierrez, L., Wilson, H. A., Rugge, T. & Bonta, J. (2013). The Prediction of Recidivism with Aboriginal Offenders: A Theoretically Informed Meta-Analysis. *Canadian Journal of Criminology and Criminal Justice*, 55, 55-99.
- Hannah-Moffat, K. (2004). Gendering Risk at What Cost: Negotiations of Gender and Risk in Canadian Women’s Prisons. *Feminism and Psychology*, 14, 243-249.
- Hannah-Moffat, K. (2009). Gridlock or Mutability: Reconsidering “Gender” and Risk Assessment. *Criminology & Public Policy*, 8, 209-219
- Hannah-Moffat, K. & Shaw, M. (2000). Thinking About Cognitive Skills? Think Again! *Criminal Justice Matters*, 39, 8-9.
- Hannah-Moffat, K. & Shaw, M. (2001). *Taking Risks: Incorporating Gender and Culture into Classification and Assessment of Federally Sentenced Women in Canada*. Ontario: Status of Women, Canada.

- Hannah-Moffat, K. & Shaw, M. (2003). The Meaning of “Risk” in Women’s Prisons: A Critique. In B. E. Bloom (Ed.) *Gendered Justice: Addressing Female Offenders* (pp. 69-96). Durham: Carolina Academic Press.
- Harland, A. T. (Ed.) (1996) *Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply*. Thousand Oaks, CA: Sage.
- Harland, A., Warren, M. & Brown, E. (1979). *A Guide to Restitution Programming*. Working Paper 17. Albany, NY: Criminal Justice Research Centre.
- Harris, M. K. (2005). In Search of Common Ground: The Importance of Theoretical Orientations in Criminology and Criminal Justice. *Criminology and Public Policy*, 4, 311-328.
- Harris, P. & Smith, S. (1996). Developing Community Corrections: An Implementation Perspective. In A. T. Harland (Ed.) *Choosing Correctional Options that Work: Defining the Demand and Evaluating the Supply* (pp.183-222). Thousand Oaks, CA: Sage.
- Heath, B. & Raynor, P. (2014). *Probation in Europe: Jersey Probation and After Care Service (JPACS)*. Utrecht: CEP, Confederation of European Probation. Available at: <https://www.cep-probation.org/wp-content/uploads/2015/03/Chapter-Jersey-final.pdf>
- Heath, B., Raynor, P. & Miles, H. (2002). What Works in Jersey: The First Ten Years. *VISTA*, 7, 202-208.
- Heim, M. H. & Tymowski, A. W. (2006). *Guidelines for the Translation of Social Science Texts*. New York: American Council of Learned Societies. SSTP Guidelines is available in multiple languages at: www.acls.org/sstp.htm
- Henley, A. (2018). Mind the Gap: Sentencing, Rehabilitation and Civic Purgatory. *Probation Journal*, 65, 285-301.
- Herzog-Evans, M. (2018). The Risk-Need-Responsivity Model: Evidence Diversity and Integrative Theory. In P. Ugwu-dike, P. Raynor & J. Annison (Eds.) *Evidence-Based Skills in Criminal Justice: International Research on Supporting Rehabilitation and Desistance* (pp. 99-126). Bristol: Policy Press.
- Hewish, S. (1999). Cited in E. D. Girolamo (2000). *A Critical Examination of the Use of Drama with Offenders in Prison and on Probation*. Thesis submitted for examination for a PhD at Middlesex University, England. Available at: https://eprints.mdx.ac.uk/6452/7/di_Girolamo_Emiliana_2000.pdf
- Higgs, C. (2001). Technology and the Depersonalisation of Knowledge and Practice. In J. Higgs & A. Titchen (Eds.) *Professional Practice in Health, Education and the Creative Arts* (pp.114-124). Oxford: Blackwell Science Ltd.

Hirschi, T. (1969). *Causes of Delinquency*. Berkeley, California: University of California Press.

HM Inspectorate of Prisons (2017). *HM Chief Inspector of Prisons for England and Wales Annual Report 2016-17*. London: HM Inspectorate of Prisons.

HM Inspectorate of Probation (2016a). *An Inspection of Through the Gate Resettlement Services for Short-Term Prisoners: A Joint Inspection by HM Inspectorate of Probation and HM Inspectorate of Prisons*. London: HMIP.

HM Inspectorate of Probation (2016b). *A Thematic Inspection of the Provision and Quality of Services in the Community for Women who Offend: An Inspection by HM Inspectorate of Probation*. London: HMIP.

HM Inspectorate of Probation (2017). *2017 Annual Report*. Manchester: HMIP.

Hoge, R., Leschied, A. & Andrews, D. A. (1993). *An Investigation of Young Offender Services in the Province of Ontario: A Report of the Repeat Offender Project*. Toronto, Ontario: Ministry of Community and Social Services.

Hollin, C. & Palmer, E. (2006). Criminogenic Need and Women Offenders: A Critique of the Literature. *Legal and Criminal Psychology*, 11, 179-195.

Hollin, C., McGuire, J., Hounscome, J., Hatcher, R., Bilby, C. & Palmer, E. (2008). Cognitive Skills Offending Behaviour Programs in the Community: A Reconviction Analysis. *Criminal Justice and Behaviour*, 35, 269-283.

Hollin, C., Palmer, E., McGuire, J., Hounscome, J., Hatcher, R., Bilby, C. & Clark, C. (2004). *Pathfinder Programmes in the Probation Service: A Retrospective Analysis*. Home Office Online Report 66/04. London: Home Office.

Holloway, K., Williams, K. & Brayford, J. (2017). *Evaluation of the Pan-Wales Women's Triage (The Diversion Scheme)*. Welsh Centre for Crime and Social Justice and University of South Wales.

Holsinger, A. M. (1999). *Opening the 'Black Box': Assessing the Relationship Between Program Integrity and Recidivism*. University of Cincinnati. Unpublished Ph.D. dissertation.

Holtfreter, K. & Cupp, R. (2007). The Empirical Status of the LSI-R for Women. *Journal of Contemporary Criminal Justice*, 23, 363-382.

Holtfreter, K. & Morash, M. (2003). The Needs of Women Offenders: Implications for Correctional Programming. *Women and Criminal Justice*, 14, 137-160.

Home Office. (1958). *Penal Practice in a Changing Society: Aspects of Future Development (England and Wales)* (Draft). London: Her Majesty's Stationery Office. Available at: <http://www.butlertrust.org.uk/wp-content/uploads/2015/07/RAB-white-paper-1959.pdf>

Home Office. (2013). *Guidance: Integrated Offender Management (IOM)*. Available at: <https://www.gov.uk/guidance/integrated-offender-management-iom>

Hosking, N. & Rico, J. (2018). Can the Recruitment of Ex-Offenders Enhance Offender Engagement? An Assessment of the London Probation Trust's Engagement Worker Role. In P. Ugwudike, P. Raynor & J. Annison (Eds.) *Evidence-Based Skills in Criminal Justice: International Research on Supporting Rehabilitation and Desistance* (pp. 335-352). Bristol: Policy Press.

Hough, C. V. (2016). Transforming Rehabilitation and its Impact on a Locally Based Rehabilitation Programme for BAME and Muslim Offenders. *European Journal of Probation*, 8, 68-81.

House of Commons Committee of Public Accounts. (2016). *Transforming Rehabilitation: Seventeenth Report of Session 2016-17*. London: House of Commons. Available at: <https://publications.parliament.uk/pa/cm201617/cmselect/cmpubacc/484/484.pdf>

House of Commons Committee of Public Accounts. (2018). *Government Contracts for Community Rehabilitation Companies: Twenty-Seventh Report of Session 2017-19*. London: House of Commons. Available at: <https://publications.parliament.uk/pa/cm201719/cmselect/cmpubacc/897/897.pdf>

House of Commons Debate. (2012). *18 December 2012*. Available at: <https://publications.parliament.uk/pa/cm201213/cmhansrd/cm121218/debtext/121218-0001.htm>

Howard, P. & Moore, R. (2009). *Research Summary 10/09: Measuring Changes in Risk and Need Over Time Using OASys*. London: Ministry of Justice. Available at: <http://webarchive.nationalarchives.gov.uk/20110201134349/http://www.justice.gov.uk/publications/docs/research-summary-oasys-10-09.pdf>

Howe, D. (1991). Knowledge, Power and the Shape of Social Work Practice. In M. Davies (Ed.) *The Sociology of Social Work*. London: Routledge.

Hubbard, D. J. & Matthews, B. (2008). Reconciling the Differences Between the "Gender-Responsive" and the "What Works" Literatures to Improve Services for Girls. *Crime & Delinquency*, 54, 225-258.

Hunt, D. E., Butler, L. F., Noy, J. E. & Rosser, M. E. (1978). *Assessing Conceptual Level by the Paragraph Completion Method*. Toronto: O.I.S.E.

International Board of Standards for Training, Performance and Instruction. (2006). *Evaluation Competencies*. Available at: <http://ibstpi.org/evaluator-competencies/>

Jersey Probation and After Care Service. *Programmes*. Available at: <http://probation.je/wp-content/uploads/2014/11/Group-Programmes.pdf>

Jolliffe, D., Hedderman, C., Palmer, E. & Hollin, C. (2011). *Re-offending Analysis of Women Offenders Referred to Together Women (TW) and the Scope to Divert from Custody*. London: Ministry of Justice.

Jones, T. & Newburn, T. (2007). *Policy Transfer and Criminal Justice: Exploring US Influence Over British Crime Control Policy*. Maidenhead: Open University Press.

Kelling, G., Edwards, S. & Moore, M. (1986). Federally Funded Community Crime Control: Urban Initiatives Anti-Crime Program. *Criminal Justice Policy Review*, 1, 59-75.

Kendall, K. (2002). Time to Think Again About Cognitive Behavioural Programmes. In P. Carlen (Ed.) *Women and Punishment: The Struggle for Justice* (pp. 182-198). Cullompton: Willan.

Kendall, K. (2004). Dangerous Thinking: A Critical History of Correctional Cognitive Behaviouralism. In G. Mair (Ed.) *What Matters in Probation* (pp. 53-89). Cullompton: Willan.

Kennedy, S. M. (2004). A Practitioner's Guide to Responsivity: Maximizing Treatment Effectiveness. *Journal of Community Corrections*, 13, 22-30.

Kerr, B. (2014). *Correctional Program Assessment Inventory-2010 Report Conducted on a Women's Project in Wales*. Swansea: Swansea Service Evaluation Team, Swansea University.

Kerr, B. (2016). *Correctional Program Assessment Inventory-2010 Report Conducted on Jersey Probation and Aftercare Service's Community Supervision Service*. Swansea: Swansea Service Evaluation Team, Swansea University. Available at: <http://probation.je/reports-v3/?mdocs-cat=mdocs-cat-11>

King, J. A., Morris, L. L. & Fitz-Gibbon, C. T. (1987). *How to Assess Program Implementation*. Newbury Park: Sage Publications.

King, J. A. & Stevahn, L. (2015). Competencies for Program Evaluators in Light of Adaptive Action: What? So What? Now What? In J. W. Altschuld & M. Engle (Eds.) *Accreditation, Certification, and Credentialing: Relevant Concerns for U.S. Evaluators*. *New Directions for Evaluation*, 145, 21-37.

King, R. D. (2000). Doing Research in Prisons in R. D. King & E. Wincup (Eds.) *Doing Research on Crime and Justice*. Oxford: Oxford University Press.

- Koons, B. A., Burrow, J. D., Morash, M. & Bynum, T. (1997). Expert and Offender Perceptions of Program Elements Linked to Successful Outcomes for Incarcerated Women. *Crime and Delinquency*, 43, 515-532.
- Lab, S. P. & Whitehead, J. T. (1990). From “Nothing Works” to “the Appropriate Works”: The Latest Stop on the Search for the Secular Grail. *Criminology*, 28, 405-417.
- Landenberger, N. A., & Lipsey, M. W. (2005). The Positive Effects of Cognitive-Behavioral Programs for Offenders: A Meta-Analysis of Factors Associated with Effective Treatment. *Journal of Experimental Criminology*, 1, 451-476.
- Langan, N. P. & Pelissier, B. M. M. (2001). Gender Differences Among Prisoners in Drug Treatment. *Journal of Substance Abuse*, 13, 291-301.
- Latessa, E. J. (2004). The Challenge of Change: Correctional Programs and Evidence-Based Practices. *Criminology & Public Policy*, 3, 547-560.
- Latessa, E. J. (2012). *Evaluating Correctional Programs*. Resource Material Series No.88 – Work product of the 151st International Training Course of the United Nations Asia and Far East Institute for the Prevention of Crime and the Treatment of Offenders (UNAFEI).
- Latessa, E. J. (2018). Does Treatment Quality Matter? Of Course it Does, and There Is Growing Evidence to Support It. *Criminology & Public Policy*, 17, 181-188.
- Latessa, E. J., Cullen, F. T. & Gendreau, P. (2002). Beyond Correctional Quackery – Professionalism and the Possibility of Effective Treatment. *Federal Probation*, 66, 43-49.
- Latessa, E. J. & Holsinger, A. (1998). The Importance of Evaluating Correctional Programs: Assessing Outcome and Quality. *Corrections Management Quarterly*, 2, 22-29.
- Latessa, E. J. & Holsinger, A. (1999). *Evaluation of the Ohio Department of Youth Service’s Community Correctional Facilities (CCF’s)*. Subgrant number 96-JJ-SI1-0697. Cincinnati, OH: Division of Criminal Justice.
- Latessa, E. J. & Smith, P. (2015). *Corrections in the Community*, 6th edition. New York: Routledge
- Latessa, E. J., Smith, P., Schweitzer, M. & Labrecque, R. M. (2013). *Evaluation of the Effective Practices in Community Supervision Model (EPICS) in Ohio*. Draft report, available at:
<http://www.uc.edu/content/dam/uc/ccjr/docs/reports/Final%20OCJS%20Report%202.2.13.pdf>
- Lederman, M. (1986). *Organizational Supportiveness Index – System Risk Inventory*. Unpublished manuscript, Carleton University, Ottawa, Canada.

- Leschied, A. W., Bernfeld, G. A. & Farrington, D. P. (2001). Implementation Issues. In G. A. Bernfeld, D. P. Farrington & A. W. Leschied (Eds.) *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs* (pp. 3-19). Chichester: John Wiley & Sons, Ltd.
- Lewis, S., Maguire, M., Raynor, P., Vanstone, M., Raybould, S. & Rix, A. (2003). *The Resettlement of Short-Term Prisoners: An Evaluation of Seven Pathfinders*. RDS Occasional Paper No. 83. London: Home Office.
- Lex, B. W. (1995). Alcohol and Other Psychoactive Substances Dependence in Women and Men. In M. V. Seeman (Ed.) *Gender and Psychopathology* (pp. 311-357). Washington DC: American Psychiatric Press.
- Liebling, A., Price, D. & Elliott, C. (1999). Appreciative Inquiry and Relationships in Prison. *Punishment & Society*, 1, 71-98.
- Link, B. G., Cullen, F. T., Struening, E., Shrout, P. E. & Dohrenwend, B. P. (1989). A Modified Labeling Theory Approach to Mental Disorders: An Empirical Assessment. *American Sociological Review*, 54, 400-423.
- Link, B. G. & Phelan, J. C. (1999). The Labelling Theory of Mental Disorder (II): The Consequences of Labelling. In *A Handbook for the Study of Mental Health: Social Contexts, Theories and Systems*. Cambridge: Cambridge University Press.
- Lipsey, M. W. (1989). The Efficacy of Intervention for Juvenile Delinquency: Results from 400 Studies. Paper presented at the 41st annual meeting of the American Society of Criminology, Reno, Nevada. Cited in D. A. Andrews & J. Bonta (2010). *The Psychology of Criminal Conduct*, 5th edition. New Providence, NJ: Anderson.
- Lipsey, M. W. (1992). Juvenile Delinquency Treatment: A Meta-Analytic Inquiry into the Variable of Effects. In T. D. Cook, H. Cooper, D. S. Cordray, H. Hartmann, L. V. Hedges, R. J. Light, T. A. Lewis, & F. Mosteller (Eds.) *Meta- Analysis for Explanation: A Casebook* (pp. 83-127). New York: Russell Sage.
- Lipsey, M. (1999). Can Rehabilitative Programs Reduce the Recidivism of Juvenile Offenders? An Inquiry into the Effectiveness of Practical Programs. *Virginia Journal of Social Policy and the Law*, 6, 611-641.
- Lipsey, M. W. (2009). The Primary Factors that Characterize Effective Interventions with Juvenile Offenders: A Meta- Analytic Overview. *Victims and Offenders*, 4, 124-147.
- Lipsey, M. W., Landenberger, N. A. & Wilson, S. J. (2007). Effects of Cognitive-Behavioral Programs for Criminal Offenders. *Campbell Systematic Reviews* 2007, 6.

Lipsey, M. W. & Wilson, D. B. (1998). Effective Intervention for Serious Delinquency in Adolescence and Early Adulthood: A Synthesis of Longitudinal Research. In R. Loeber & D. P. Farrington (Eds.) *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions* (pp. 86-105). Thousand Oaks: Sage.

Lipsey, M.W., & Wilson, D.B. (2001). *Practical Meta-Analysis*. Thousand Oaks: Sage Publications.

Lipton, D., Martinson, R., & Wilks, J. (1975). *The Effectiveness of Correctional Treatment: A Survey of Treatment Evaluation Studies*. New York: Praeger.

Logan, C. H. & Gaes, G. G. (1993). Meta-Analysis and the Rehabilitation of Punishment. *Justice Quarterly*, 10, 245-263.

Lösel, F. (1995). The Efficacy of Correctional Treatment: a Review and Synthesis of Meta-Evaluations. In J. McGuire (Ed.) *What Works: Reducing Reoffending. Guidelines from Research and Practice* (79-111). Chichester: Wiley.

Lösel, F. (2001). Evaluating the Effectiveness of Correctional Programs: Bridging the Gap Between Research and Practice. In G. A. Bernfeld, D. P. Farrington & A. W. Leschied (Eds.) *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs* (pp. 247-268). Chichester: John Wiley & Sons, Ltd.

Lösel, F. (2012). Offender Treatment and Rehabilitation: What Works? In M. Maguire, R. Morgan & R. Reiner (Eds.) *The Oxford Handbook of Criminology, 5th edition* (pp. 986-1016). Oxford: Oxford University Press.

Lovins, L. B., Lowenkamp, C. T., Latessa, E. J. & Smith, P. (2007). Application of the Risk Principle to Female Offenders. *Journal of Contemporary Criminal Justice*, 23, 383-398.

Lowenkamp, C. T. (2004) *Correctional Program Integrity and Treatment Effectiveness: A Multi-site, Program-level Analysis*. University of Cincinnati. Unpublished Ph.D. dissertation.

Lowenkamp, C. T., Holsinger, A. M. & Latessa, E. J. (2005). Are Drug Courts Effective: A Meta-Analytic Review. *Journal of Community Corrections*, 28, 5-10.

Lowenkamp, C. T., Holsinger, A. M., Robinson, C. R. & Alexander, M. (2014). Diminishing or Durable Treatment of STARR? A Research Note on 24-Month Re-Arrest Rates. *Journal of Crime & Justice*, 37, 275-283.

Lowenkamp, C. T. & Latessa, E.J. (2003). *Evaluation of Ohio's Halfway Houses and Community Based Correctional Facilities*. Cincinnati, OH: Center for Criminal Justice Research, University of Cincinnati.

Lowenkamp, C. T. & Latessa, E.J. (2005a). *Evaluation of Ohio's CCA Programs*. Cincinnati, OH: Center for Criminal Justice Research, University of Cincinnati.

Lowenkamp, C. T. & Latessa, E.J. (2005b). *Evaluation of Ohio's Reclaim Funded Programs, Community Correctional Facilities, and DYS Facilities*. Cincinnati, OH: Center for Criminal Justice Research, University of Cincinnati.

Lowenkamp, C. T., Latessa, E. J. & Holsinger, A. M. (2006). The Risk Principle in Action: What Have We Learned From 13,676 Offenders and 97 Correctional Programs? *Crime & Delinquency*, 52, 77-93.

Lowenkamp, C. T., Latessa, E. J. & Smith, P. (2006). Does Correctional Program Quality Really Matter? The Impact of Adhering to the Principles of Effective Intervention. *Criminology and Public Policy*, 5, 575-594.

Lowenkamp, C. T., Makarios, M. D., Latessa, E. J., Lemke, R. & Smith, P. (2010). Community Corrections Facilities for Juvenile Offenders in Ohio: An Examination of Treatment Integrity and Recidivism. *Criminal Justice and Behaviour*, 37, 695-708.

MacKenzie, D., Shaw, J. & Gowdy, V. (1993). An Evaluation of Shock Incarceration Programs in Louisiana. In *Research in Brief*. Washington DC: National Institute of Justice.

Maguire, M. (2004). The Crime Reduction Programme in England and Wales: Reflections on the Vision and the Reality. *Criminal Justice*, 4, 213-237.

Maguire, M., Grubin, D., Lösel, F. & Raynor, P. (2010). 'What Works' and the Correctional Services Accreditation Panel: Taking Stock from an Inside Perspective. *Criminology & Criminal Justice*, 10, 37-58.

Maguire, M. & Raynor, P. (2006). How the Resettlement of Prisoners Promotes Desistance from Crime: Or Does It? *Criminology and Criminal Justice*, 6, 19-38.

Maguire, M., Raynor, P., Vanstone, M. & Kynch, J. (2000). Voluntary After-Care and the Probation Service: A Case of Diminishing Responsibility. *The Howard Journal*, 39, 234-248.

Mair, G. (2000). Credible Accreditation? *Probation Journal*, 47, 268-271.

Mair, G. (2004). The Origins of What Works in England and Wales: a House Built on Sand? In G. Mair (Ed.) *What Matters in Probation* (pp. 12-33). Cullompton: Willan Publishing.

Malloch, M., McIvor, G. & Burgess, C. (2014). 'Holistic' Community Punishment and Criminal Justice Interventions for Women. *The Howard Journal*, 53, 395-410.

Mance, H. (2016). Britain Has Had Enough of Experts, Says Gove. *Financial Times*. Available at: <https://www.ft.com/content/3be49734-29cb-11e6-83e4-abc22d5d108c>

Mann, R. E. (2009). Sex Offender Treatment: The Case for Manualization. *Journal of Sexual Aggression*, 15, 121-131.

- Martinson, R. (1974). What Works? – Questions and Answers About Prison Reform. *Public Interest*, 35, 22-54.
- Martinson, R. (1979). New Findings, New Views: A Note of Caution Regarding Sentencing Reform. *Hofstra Law Review*, 7, 243-258.
- Maruna, S. (2001). *Making Good: How Ex-Convicts Reform and Rebuild Their Lives*. Washington D.C.: American Psychological Association.
- Maruna, S. (2015). Qualitative Research, Theory Development and Evidence-Based Corrections: Can Success Stories Be “Evidence”? In J. Miller & W. R. Palacios (Eds.) *Qualitative Research in Criminology* (pp. 311-337). New York: Routledge.
- Maruna, S. & LeBel, T. P. (2003). Welcome Home? Examining the Reentry Court Concept from a Strengths-Based Perspective. *Western Criminology Review*, 4, 91-107.
- Maruna, S. & LeBel, T. P. (2010). The Desistance Paradigm in Correctional Practice: From Programmes to Lives. In F. McNeill, P. Raynor & C. Trotter (Eds.) *Offender Supervision: New Directions in Theory, Research and Practice* (pp. 65-87). Oxon: Willan Publishing.
- Matthews, B., Hubbard, D. J. & Latessa, E. (2001). Making the Next Step: Using Evaluability Assessment to Improve Correctional Programming. *The Prison Journal*, 81, 454-472.
- Mawby, R. C. & Worrall, A. (2013). *Doing Probation Work: Identity in a Criminal Justice Occupation*. Oxon: Routledge.
- McCulloch, T. (2005). Probation, Social Context and Desistance: Retracing the Relationship. *Probation Journal*, 52, 8-22.
- McGuire, J. (Ed.). (1995). *What Works: Reducing Reoffending*. Chichester: Wiley.
- McGuire, J. (2000). *Cognitive-Behavioural Approaches: An Introduction to Theory and Research*. London: HMIP.
- McGuire, J. (2004). *Understanding Psychology and Crime: Perspectives on Theory and Action*. Maidenhead: Open University Press.
- McGuire, J., Bilby, C., Hatcher, R., Hollin, C., Hounscome, J. & Palmer, E. (2008). Evaluation of Structured Cognitive-Behavioural Treatment Programs in Reducing Criminal Recidivism. *Journal of Experimental Criminology*, 4, 21-40.
- McIvor, G. (1990). *Sanctions for Serious or Persistent Offenders*. Sterling: Social Work Research Centre.

McIvor, G. (2007). The Nature of Female Offending. In R. Sheehan, G. McIvor & C. Trotter (Eds.) *What Works with Women Offenders* (pp. 1-22). Cullompton: Willan Publishing.

McMahon, G., Hall, A., Hayward, G., Hudson, C. & Roberts, C. (2004). *Basic Skills Programmes in the Probation Service: An Evaluation of the Basic Skills Pathfinder*. Home Office Research Findings 203. London: Home Office.

McMurrin, M. (2009). Motivational Interviewing with Offenders: A Systematic Review. *Legal and Criminological Psychology*, 14, 83-100.

McNeill, F. (2006). A Desistance Paradigm for Offender Management. *Criminology and Criminal Justice*, 6, 39-62.

McNeill, F. (2012). Four Forms of Offender Rehabilitation: Towards an Interdisciplinary Perspective. *Legal and Criminological Psychology*, 17, 18-36.

McNeill, F., Farrall, S., Lightowler, C. & Maruna, S. (2012) Reexamining “Evidence-Based Practice” in Community Corrections: Beyond “a Confined View” of What Works. *Justice Research and Policy*, 14, 35-60.

McNeill, F. and Maruna, S. (2007). Giving Up and Giving Back: Desistance, Generativity and Social Work with Offenders. In G. McIvor and P. Raynor (Eds.) *Developments in Work with Offenders* (pp. 224-339). London: Routledge.

McNeill, F. Raynor, P. & Trotter, C. (Eds.) (2010). *Offender Supervision: New Directions in Theory, Research and Practice*. Oxon: Willan Publishing.

McNeill, F. & Weaver, B. (2010). *Changing Lives? Desistance Research and Offender Management*. SCCJR Project Report; No.03/2010.

McNeill, F. and Whyte, B. (2007). *Reducing Reoffending: Social Work and Community Justice in Scotland*. Cullompton: Willan.

Merrick, R. (2017). Women Hit Hardest by ‘Shameful’ Short Prison Sentences, New Figures Reveal. *Independent*. Available at: <https://www.independent.co.uk/news/uk/politics/women-prison-sentences-short-jailed-periods-less-month-inmates-minor-offences-children-labour-a8124896.html>

Messina, N., Grella, C. E., Cartier, J. & Torres, S. (2010). A Randomized Experimental Study of Gender-Responsive Substance Abuse Treatment for Women in Prison. *Journal of Substance Abuse Treatment*, 38, 97-107.

Miles, H. & Raynor, P. (2014). *Reintegrative Justice in Practice: The Informal Management of Crime in an Island Community*. Farnham: Ashgate.

Miles, H., Raynor, P. & Coster, B. (2009). *Community Sentences and their Outcomes in Jersey: the Third Report*. Available at: <http://www.cep-probation.org/wp-content/uploads/2015/03/Community-Sentences-and-their-Outcomes-in-Jersey.pdf>

Miles, H., Raynor, P. & Waterman, B. (2015). *Community Sentences and their Outcomes in Jersey: the Fourth Report*. Available at: <http://probation.je/wp-content/uploads/2015/07/Community-Sentences-and-their-Outcomes-in-Jersey-the-fourth-report-.pdf>

Miles, M. & Huberman, A. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks, CA: Sage Publications.

Miller, J. B. (1986). *What Do We Mean By Relationships?* Work in Progress No. 33. Wellesley: Stone Centre, Working Paper Series.

Miller, W. R. & Rollnick, S. (2013). *Motivational Interviewing: Preparing people for change, 3rd edition*. New York: The Guilford Press.

Miller, W. R. & Rose, G. S. (2009). Toward a Theory of Motivational Interviewing. *American Psychologist*, 64, 527-537.

Millson, B., Robinson, D. & Van Dieten, M. (2010). *Women Offender Case Management Model: The Connecticut Project*. Washington, DC: National Institute of Corrections.

Ministry of Justice. (2007). *Background Briefing on the Crown Dependencies: Jersey, Guernsey and the Isle of Man*. London: Ministry of Justice. Available at: https://www.justice.gov.uk/downloads/about/moj/our-responsibilities/Background_Briefing_on_the_Crown_Dependencies2.pdf

Ministry of Justice. (2011). *Offender Management Caseload Statistics 2010 Tables (Population A1.20)*. London: Ministry of Justice.

Ministry of Justice. (2012). *Statistics on Women and the Criminal Justice System 2011*. London: Ministry of Justice. Available at: www.justice.gov.uk/statistics/criminal-justice/women

Ministry of Justice. (2013a). *Transforming Rehabilitation: A Strategy for Reform*. London: Her Majesty's Stationery Office.

Ministry of Justice. (2013b). *Strategic Objectives for Female Offenders*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/177038/strategic-objectives-female-offenders.pdf

Ministry of Justice. (2014). *Offender Management Caseload Statistics Annual Tables 2013*. London: Ministry of Justice.

Ministry of Justice. (2015). *Justice Data Lab Re-offending Analysis: Women's Centres Throughout England*. London: Ministry of Justice.

Ministry of Justice. (2017a). *Offender Management Statistics (Quarterly) July – September 2016, Prison Population: 31 December 2016*. London: Ministry of Justice.

Ministry of Justice. (2017b). *Offender Management Statistics (Quarterly) July – September 2016, Prison Receptions: July to September 2016*. London: Ministry of Justice.

Ministry of Justice. (2018a). *Final and Interim Proven Reoffending Statistics for the Community Rehabilitation Companies and the National Probation Service*. London: Ministry of Justice. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/706194/crc-nps-interim-results-bulletin-apr18.pdf

Ministry of Justice. (2018b). *Population and Capacity Briefing for Friday 1st June 2018*. Available at: <https://www.gov.uk/government/statistics/prison-population-figures-2018>

Ministry of Justice. (2018c). *Female Offender Strategy*. London: Ministry of Justice. Available at: <https://www.gov.uk/government/publications/female-offender-strategy>

Ministry of Justice. (2018d). *Strengthening Probation, Building Confidence*. London: Ministry of Justice. Available at: https://consult.justice.gov.uk/hm-prisons-and-probation/strengthening-probation-building-confidence/supporting_documents/strengtheningprobationbuildingconfidence.pdf

Ministry of Justice. (2018e). *A Whole System Approach for Female Offenders: Emerging Evidence*. London: Ministry of Justice. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/719771/guide-to-whole-system-approach.pdf

Ministry of Justice Social Exclusion Task Force. (2009). *Short Study on Women Offenders*. London: Ministry of Justice. Available at: http://webarchive.nationalarchives.gov.uk/+/http://www.cabinetoffice.gov.uk/media/209663/setf_shortstudy_womenoffenders.pdf

Mohr, W. K., Martin, A., Olson, J. N., Pumariega, A. J. & Branca, N. (2009). Beyond Point and Level Systems: Moving Toward Child-Centred Programming. *American Journal of Orthopsychiatry*, 79, 8-18.

National Assembly for Wales. (2018). *NDM6665 – Member Debate Under Standing Order 11.21 (iv)*. Available at: <http://www.senedd.assembly.wales/ieIssueDetails.aspx?IId=21292&Opt=3>

National Association of Probation Officers. (2001). AGM Resolutions 2001. *Napo News* 134, 10-15.

National Association of Probation Officers (2015) *Probation Under Threat*. Available at: <http://www.napo.org.uk/about/probationunderthreat.cfm>

National Offender Management Service. (2010). *Guidance for the Voluntary and Community Sector: Working with NOMS to manage the Risk of Serious Harm*. London: Ministry of Justice, NOMS.

National Offender Management Service. (2014). *Research Applications*. London: Ministry of Justice, NOMS.

National Offender Management Service. (2015). *Better Outcomes for Women Offenders*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/457922/Better_Outcomes_for_Women_Offenders_September_2015.pdf

NBC. (2017). Conway: Press Secretary Gave ‘Alternative Facts’. *Meet the Press*. Available at: <https://www.nbcnews.com/meet-the-press/video/conway-press-secretary-gave-alternative-facts-860142147643>

Nelson-Jones, R. (2011). *Basic Counselling Skills: A Helper’s Manual*, 3rd edition. London: Sage.

Nesovic, A. (2003). *Psychometric Evaluation of the Correctional Program Assessment Inventory (CPAI)*. Carleton University, Ottawa, Ontario. Unpublished Ph.D. dissertation.

Noaks, L. & Wincup, E. (2004). *Criminological Research: Understanding Qualitative Methods*. London: Sage Publications.

Offender Rehabilitation Act 2014. (2014). London: The Stationery Office Ltd.

Office of the Secretary of State for Wales. (2018). *New Plans for Probation in Wales Announced by UK Government*. Available at: <https://www.gov.uk/government/news/new-plans-for-probation-in-wales-announced-by-uk-government>

Ogloff, J. R. P., Davis, M. R. (2004). Advances in Offender Assessment and Rehabilitation: Contributions of the Risk-Needs-Responsivity Approach. *Psychology, Crime & Law*, 10, 229-242.

O’Loughlin, A. (2014). The Offender Personality Disorder Pathway: Expansion in the Face of Failure? *The Howard Journal of Criminal Justice*, 53, 173-192.

Olver, M. E., Lewis, K. & Wong, S. C. P. (2013). Risk Reduction Treatment of High-Risk Psychopathic Offenders: The Relationship of Psychopathy and Treatment Change to Violent Recidivism. *Personality Disorders: Theory, Research, and Treatment*, 4, 160-167.

Olver, M. E., Stockdale, K. C. & Wormith, J. S. (2014). Thirty Years of Research on the Level of Service Scales: A Meta-Analytic Examination of Predictive Accuracy and Sources of Variability. *Psychological Assessment*, 26, 156-176.

Palmer, E. J., Hatcher, R. M., McGuire, J. & Hollin, C. R. (2015). Cognitive Skills Programs for Female Offenders in the Community: Effect on reconviction. *Criminal Justice and Behaviour*, 42, 345-360.

Palmer, E. J. & Hollin, C. R. (2007). The Level of Service Inventory-Revised with English Women Prisoners: A Needs and Reconviction Analysis. *Criminal Justice and Behaviour*, 34, 971-984.

Palmer, E. J., McGuire, J., Hounscome, J., Hatcher, R., Bilby, C. & Hollin, C. (2007). Offending Behaviour Programmes in the Community: The Effects on Reconviction of Three Programmes with Adult Male Offenders. *Legal and Criminological Psychology*, 12, 251-264.

Palmer, T. (1963). *Types of Probation Officers and Types of Youth on Probation: Their Views and Interactions*. University of California: Youth Studies Center.

Palmer, T. (1965). Types of Treaters and Types of Juvenile Offenders. *Youth Authority Quarterly*, 18, 14-23.

Palmer, T. (1967). *Personality Characteristics and Professional Orientations of Five Groups of Community Treatment Project Workers: A Preliminary Report on Differences Among Treaters*. Community Treatment Project Report Series: No. 1.

Palmer, T. (1971). California's Community Treatment Program for Delinquent Adolescents. *Journal of Research in Crime and Delinquency*, 8, 74-92.

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

Palmer, T. (1992). *The Re-Emergence of Correctional Intervention*. Newbury Park, CA: Sage Publications.

Palmer, T. (1996). Programmatic and Nonprogrammatic Aspects of Successful Intervention. In Alan T. Harland (Ed.) *Choosing Correctional Options That Work: Defining the Demand and Evaluating the Supply* (pp. 131-182). Thousand Oaks, CA: Sage

Paparozi, M. A. & Gendreau, P. (2005). An Intensive Supervision Program that Worked: Service Delivery, Professional Orientation, and Organizational Supportiveness. *The Prison Journal*, 85, 445-466.

Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods*, 3rd edition. Thousand Oaks, CA: Sage Publications.

- Paul, G. L. (1967). Strategy of Outcome Research in Psychotherapy. *Journal of Consulting Psychology*, 31, 109-118.
- Pearce, S., Murray, D. & Lane, M. (2015). *HMP Doncaster Payment by Results Pilot: Final Process Evaluation Report*. London: Ministry of Justice.
- Pearson, D. A., McDougall, C., Kanaan, M., Bowles, A. & Torgerson, D. J. (2011). Reducing Criminal Recidivism: Evaluation of Citizenship, an Evidence-Based Probation Supervision Process. *Journal of Experimental Criminology*, 7, 73-102.
- Pearson, F. S., Lipton, D. S., Cleland, C. M., & Yee, D. S. (2002). The Effects of Behavioral/Cognitive-Behavioral Programs on Recidivism. *Crime and Delinquency*, 48, 476-496.
- Pease, K., Billingham, S. & Earnshaw, I. (1977) *Community Service Assessed in 1976*. Home Office Research Study No. 39. London: HMSO
- Petersilia, J. (1990). Conditions that Permit Intensive Supervision Programs to Survive. *Crime and Delinquency*, 36, 126-145.
- Phillips, J. (2014). Probation in the News: Transforming Rehabilitation. *British Journal of Community Justice*, 12, 7-48.
- Phillips, J., Westaby, C. & Fowler, A. (2016). 'It's Relentless': The Impact of Working Primarily with High-Risk Offenders. *Probation Journal*, 63, 182-192.
- Polaschek, D. L. L. (2012). An Appraisal of the Risk-Need-Responsivity (RNR) Model of Offender Rehabilitation and its Application in Correctional Treatment. *Legal and Criminological Psychology*, 17, 1-17.
- Pollack, J. (1986). *Sex and Supervision: Guarding Male and Female Inmates*. New York: Greenwood Press.
- Pollack, S. (2005). Taming the Shrew: Regulating Prisoners Through Women-Centred Mental Health Programming. *Critical Criminology*, 13, 71-87.
- Pollock, I. & Glyn, A. (2018). Women Who Commit Minor Crimes 'Need Help Not Prison'. *BBC News*. Available at: <http://www.bbc.com/news/uk-wales-43906626>
- Porporino, F. J. (2010). Bringing Sense and Sensitivity to Corrections: From Programmes to 'Fix' Offenders to Services to Support Desistance. In J. Brayford, F. Cowe & J. Deering (Eds.) *What Else Works? Creative Work with Offenders* (pp. 61-85). Cullompton: Willan Publishing.
- Pratt, T. (2002). Meta-Analysis and its Discontents: Treatment Destruction Techniques Revisited. *Journal of Offender Rehabilitation*, 35, 23-40.

- Prison Reform Trust Women's Justice Taskforce. (2014). *Transforming Lives: Reducing Women's Imprisonment*. Available at: prisonreformtrust.org.uk
- Prison Reform Trust. (2017a). *Why Focus on Reducing Women's Imprisonment?: Prison Reform Trust Briefing, February 2017*. London: Prison Reform Trust.
- Prison Reform Trust. (2017b). *Bromley Briefings Prison Factfile, Autumn 2017*. London: Prison Reform Trust. Available at: <http://www.prisonreformtrust.org.uk/Portals/0/Documents/Bromley%20Briefings/Autumn%202017%20factfile.pdf>
- Prochaska, J. O. & DiClemente, C. C. (1982). Transtheoretical Therapy: Toward a More Integrative Model of Change. *Psychotherapy: Theory, Research, and Practice*, 19, 276-288.
- Quay, H. C. (1987). Institutional Treatment. In H. C. Quay (Ed.) *Handbook of Juvenile Delinquency* (pp. 244-265). New York: John Wiley & Sons.
- Rautkis, E. (2016). "When You First Get There, You Wear Red": Youth Perceptions of Point and Level Systems in Group Home Care. *Child and Adolescent Social Work Journal*, 33, 91-102.
- Raynor, P. (1998). Attitudes, Social Problems and Reconvictions in the STOP Probation Experiment. *Howard Journal*, 37, 1-15.
- Raynor, P. (2003). Evidence-Based Probation and its Critics. *Probation Journal*, 50, 334-345.
- Raynor, P. (2004). The Probation Service 'Pathfinders': Finding the Path and Losing the Way? *Criminal Justice*, 4, 309-325.
- Raynor, P. (2012). Community Penalties, Probation, and Offender Management. In M. Maguire, R. Morgan & R. Reiner (Eds.) *The Oxford Handbook of Criminology*, 5th edition (pp. 928-944). Oxford: Oxford University Press.
- Raynor, P. (2014). Consent to Probation in England and Wales: How it Was Abolished and Why it Matters. *European Journal of Probation*, 1, 1-12.
- Raynor, P. (2017). *The Design, Implementation and Consequences of 'Transforming Rehabilitation': Evidence Submitted to the Justice Committee Inquiry*. Available at: www.parliament.uk.
- Raynor, P. (2018a). The Search for Impact in British Probation: From Programmes to Skills and Implementation. In P. Ugwu-dike, P. Raynor & J. Annison (Eds.) *Evidence-Based Skills in Criminal Justice: International Research on Supporting Rehabilitation and Desistance* (pp. 37-56). Bristol: Policy Press.
- Raynor, P. (2018b). From 'Nothing Works' to 'Post-Truth': The Rise and Fall of Evidence in British Probation. *European Journal of Probation*, 10, 59-75.

Raynor, P. (2018c). *A Short History of Evidence in Twelve Quotations*. Paper given at The Welsh Centre for Crime and Social Justice Ninth Annual Conference, 30th April – 1st May 2018, Gregynog Hall, Newtown, Wales.

Raynor, P. (2018d). Back to the Future? The Long View of Probation and Sentencing. *Probation Journal*, 65, 335-347.

Raynor, P. & Robinson, G. (2009). *Rehabilitation, Crime and Justice*. Basingstoke: Palgrave Macmillan.

Raynor, P. & Ugwudike, P. (2013). Skills and Training in British Probation: A Tale of Neglect and Possible Revival. *Federal Probation*, 77, 49-53.

Raynor, P., Ugwudike, P. & Vanstone, M. (2009). *The Jersey Supervision Interview Checklist Version 7C*. © The Jersey Crime and Society Project.

Raynor, P., Ugwudike, P. & Vanstone, M. (2014). The Impact of Skills in Probation Work: A Reconviction Study. *Criminology and Criminal Justice*, 14, 235-249.

Raynor, P. & Vanstone, M. (1996). Reasoning and Rehabilitation in Britain: The Results of the Straight Thinking On Probation (STOP) Programme. *International Journal Offender Therapy and Comparative Criminology*, 40, 272-284.

Raynor, P. & Vanstone, M. (1997). *Straight Thinking On Probation (STOP): The Mid Glamorgan Experiment*. Probation Studies Unit Report No. 4. Oxford: University of Oxford Centre for Criminological Research.

Raynor, P. & Vanstone, M. (2001). "Straight Thinking On Probation": Evidence-Based Practice and the Culture of Curiosity. In G. A. Bernfeld, D. P. Farrington & A. W. Leschied (Eds.) *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs* (pp. 247-268). Chichester: John Wiley & Sons, Ltd.

Raynor, P. & Vanstone, M. (2016). Moving Away from Social Work and Half Way Back Again: New Research on Skills in Probation. *British Journal of Social Work*, 46, 1131-1147.

Redondo, S., Sánchez-Meca, J. & Garrido, V. (1999). The Influence of Treatment Programs on the Recidivism of Juvenile and Adult Offenders: An European Meta-Analytic Review. *Psychology, Crime & Law*, 5, 251-278.

Reid, W. J. & Epstein, L. (1972). *Task-Centred Casework*. New York: Columbia University Press.

Reisig, M., Holtfreter, K. & Morash, M. (2006). Assessing Recidivism Risk Across Female Pathways to Crime. *Justice Quarterly*, 23, 384-405.

- Reitsma-Street, M. & Leschied, A. W. (1988). The Conceptual-Level Matching Model in Corrections. *Criminal Justice and Behaviour*, 15, 92-108.
- Rettinger, J. & Andrews, D. A. (2010). General Risk and Need, Gender Specificity, and the Recidivism of Female Offenders. *Criminal Justice and Behaviour*, 37, 29-46.
- Rex, S. (1999). Desistance from Offending: Experiences of Probation. *The Howard Journal*, 38, 366-383.
- Rex, S. & Hosking, N. (2013). A Collaborative Approach to Developing Probation Practice: Skills for Effective Engagement, Development and Supervision (SEEDS). *Probation Journal*, 60, 332-338.
- Rex, S., Gelsthorpe, L., Roberts, C. & Jordan, P. (2004). *What's Promising in Community Service: Implementation of Seven Pathfinder Projects*. Home Office Research Findings 231. London: Home Office.
- Richie, B. (2001). Challenges Incarcerated Women Face as they Return to their Communities: Findings from Life History Interviews. *Crime and Delinquency*, 47, 368-389.
- Robinson, C. R., Lowenkamp, C. T., Holsinger, A. M., VanBenschoten, S., Alexander, M. & Oleson, J. C. (2012). A Random Study of Staff Training Aimed at Reducing Re-Arrest (STARR): Using Core Correctional Practices in Probation Interactions. *Journal of Crime and Justice*, 35, 167-188.
- Robinson, G. & Crow, I. (2009). *Offender Rehabilitation: Theory, Research and Practice*. London: Sage.
- Robinson, G. & McNeill, F. (2010). Probation in the United Kingdom. In M. Herzog-Evans (Ed.) *Transnational Criminology Manual, Volume 3*, 741-762.
- Robinson, G., Priede, C., Farrall, S., Shapland, J. & McNeill, F. (2012). Doing 'Strengths-Based' Research: Appreciative Inquiry in a Probation Setting. *Criminology & Criminal Justice*, 13, 3-20.
- Robinson, G., Priede, C., Farrall, S., Shapland, J. & McNeill, F. (2014). Understanding 'Quality' in Probation Practice: Frontline Perspectives in England & Wales. *Criminology & Criminal Justice*, 14, 123-142.
- Rogers, P. J. & Fraser, D. (2003). Appreciating Appreciative Inquiry. *New Directions for Evaluation*, 100, 75-83.
- Rorty, R. (1983). Pragmatism Without Method. In R. Rorty (1991). *Objectivity, Relativism, and Truth*. Cambridge: Cambridge University Press.
- Rorty, R. (1990). Pragmatism as Anti-Representationalism. In J. P. Murphy. *Pragmatism: From Peirce to Davidson* (pp. 1-6). Boulder, CO: Westview.

- Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.
- Rosenthal, R. & Rubin, D. B. (1982). A Simple, General Purpose Display of Magnitude of Experimental Effect. *Journal of Educational Psychology*, 74, 166-169.
- Ross, R. & Fabiano, E. (1986). *Female Offenders: Correctional Afterthoughts*. Jefferson: McFarland.
- Rothschild, T. & Selvaggi, K. *The TIER System*. Contact: metam1@aol.com
- Royka, J. G. (2002). Overcoming the Fear of Using Drama in English Language Teaching. *The Internet TESL Journal*, 8. Available at: <http://iteslj.org/>
- Rumgay, J. (2004). Scripts for Safer Survival: Pathways Out of Female Crime. *Howard Journal of Criminal Justice*, 43, 405-419.
- Russ-Eft, D., Bober, M. J., de la Teja, I. Foxon, M. J. & Koszalka, T. A. (2008). *Evaluator Competencies: Standards for the Practice of Evaluation in Organizations*. San Francisco, CA: Jossey-Bass.
- Salisbury, E. J. (2015). Program Integrity and the Principles of Gender-Responsive Interventions: Assessing the Context for Sustainable Change. *Criminology & Public Policy*, 14, 1-10.
- Savage, M. (2018). Private Probation Firms Fail to Cut Rates of Reoffending. *The Observer*. Available at: <https://www.theguardian.com/society/2018/feb/03/private-firms-fail-cut-rates-reoffending-low-medium-risk-offenders>
- Saxena, P., Messina, N. J. & Grella, C. E. (2014). Who Benefits from Gender-Responsive Treatment? Accounting for Abuse History on Longitudinal Outcomes for Women in Prison. *Criminal Justice and Behavior*, 41, 417-432.
- Selvaggi, K. S., Wolf, A. & Long, C. (2013). *Webinar: A Trauma-Informed Effective Reinforcement System for Girls*. Available at: <https://nicic.gov/webinar-trauma-informed-effective-reinforcement-system-girls>
- Severiens, S. & ten Dam, G. T. M. (1994). Gender Differences in Learning Styles: A Narrative Review and Quantitative Meta-Analysis. *Higher Education*, 27, 487-501.
- Shaw, M. & Hannah-Moffat, K. (2004). How Cognitive Skills Forgot About Gender and Diversity. In G. Mair (Ed.) *What Matters in Probation* (pp. 90-121). Cullompton: Willan.
- Shear, L. & Smith, N. (1999). *Training and Technical Assistance Component. Final Evaluation Report*. Office of Child Abuse Prevention: Crime Prevention Program.

- Sheehan, R., McIvor, G. & Trotter, C. (Eds.) (2007). *What Works with Women Offenders*. Cullompton: Willan Publishing.
- Skeem, J. L., Louden, J. E., Polaschek, D. & Camp, J. (2007). Assessing Relationship Quality in Mandated Community Treatment: Blending Care with Control. *Psychological Assessment*, 19, 397-410.
- Smith, L. (1982). 'Day Training Centres'. *Home Office Research Bulletin*, 14, 34-7.
- Smith, P., Cullen, F. T. & Latessa, E. J. (2009). Can 14,737 Women Be Wrong? A Meta-Analysis of the LSI-R and Recidivism for Female Offenders. *Criminology & Public Policy*, 8, 183-208.
- Smith, P., Schweitzer, M., Labrecque, R. M., & Latessa, E. J. (2012). Improving Probation Officers' Supervision Skills: An Evaluation of the EPICS Model. *Journal of Crime and Justice*, 35, 189-199.
- Snyder, H. N. and Mulako-Wangota, J. Bureau of Justice Statistics. Generated using the *Arrest Data Analysis Tool* at www.bjs.gov. (February, 2015).
- Sorbello, L., Eccleston, L., Ward, T. & Jones, R. (2002). Treatment Needs of Female Offenders: A Review. *Australian Psychologist*, 37, 196-295.
- Sorsby, A., Shapland, J. & Robinson, G. (2017). Using Compliance with Probation Supervision as an Interim Outcome Measure in Evaluating a Probation Initiative. *Criminology and Criminal Justice*, 17, 40-61.
- Spiegler, M. D. & Guevremont, D. C. (2009). *Contemporary Behaviour Therapy*, 5th edition. Belmont, CA: Wadsworth.
- Spiropoulos, G. V., Spruance, L., Van Voorhis, P. & Schmitt, M. M. (2005). Pathfinders and Problem Solving: Comparative Effects of Two Cognitive-Behavioural Programs Among Men and Women Offenders in Community and Prison. *Journal of Offender Rehabilitation*, 42, 69-94.
- Staffordshire and West Midlands Community Rehabilitation Company. *Programmes*. Available at: <http://www.swmcrc.co.uk/our-work/programmes/>
- Stein, D. M., & Lambert, M. L. (1984). On the Relationship Between Therapist Experience and Psychotherapy Outcome. *Clinical Psychology Review*, 4, 127-142.
- Stevahn, L., King, J. A., Ghore, G. & Minnema, J. (2005). Establishing Essential Competencies for Program Evaluators. *American Journal of Evaluation*, 26, 43-59.
- Sutherland, A. (2017). A Propensity Score Analysis of a Community Resettlement Programme for Women Prisoners. *Criminology & Criminal Justice*, 0, 1-18.

- Tafreshi, D., Slaney, K. L. & Neufield, S. D. (2016). Quantification in Psychology: Critical Analysis of an Unreflective Practice. *Journal of Theoretical and Philosophical Psychology*, 36, 233-249.
- Tannenbaum, F. (1938). *Crime and Community*. New York: Columbia University Press.
- Taxman, F. S. (2008). No Illusions: Offender and Organizational Change in Maryland's Proactive Community Supervision Efforts. *Criminology & Public Policy*, 7, 275-302.
- Taxman, F. S. (2014). Second Generation of RNR: The Importance of Systemic Responsivity in Expanding Core Principles of Responsivity. *Federal Probation*, 78, 32-40.
- Taxman, F. S., Shepardson, E. S. & Bryne, J. M. (2004). *Tools of the Trade: A Guide to Incorporating Science into Practice*. National Institute of Corrections, US Department of Justice and Maryland Department of Public Safety and Correctional Services. Available at: <https://s3.amazonaws.com/static.nicic.gov/Library/020095.pdf>
- Taxman, F. S., Yancey, C. & Bilanin, J. E. (2006). *Proactive Community Supervision in Maryland: Changing Offender Outcomes*. Maryland: Maryland Division of Parole and Probation. Available at: https://www.dpscs.state.md.us/publicinfo/publications/pdfs/PCS_Evaluation_Feb06.pdf
- Tew, J. & Atkinson, R. (2013). The Chromis Programme: From Conception to Evaluation. *Psychology, Crime & Law*, 19, 415-431.
- Thomas, D. R. (2003). *A General Inductive Approach for Qualitative Data Analysis*. Available at: <http://www.frankumstein.com/PDF/Psychology/Inductive%20Content%20Analysis.pdf>
- Thornton, D. M. (1987). Correctional Evaluation of Custodial Regimes. In B. J. McGurk, D. M. Thornton & M. Williams (Eds.) *Applying Psychology to Imprisonment* (pp. 467-481). London: Her Majesty's Stationery Office.
- Tonry, M. (2010). "Public Criminology" and Evidence-Based Policy. *Criminology & Public Policy*, 9, 783-797.
- Toronjo, H. & Taxman, F. S. (2018). Supervision Face-to-Face Contacts: The Emergence of an Intervention. In P. Ugwu-dike, P. Raynor & J. Annison (Eds.) *Evidence-Based Skills in Criminal Justice: International Research on Supporting Rehabilitation and Desistance* (pp. 217-242). Bristol: Policy Press.
- Trevisan, M. S. & Walser, T. M. (2015). *Evaluability Assessment: Improving Evaluation Quality and Use*. Thousand Oaks, CA: Sage
- Traux, C. B. & Carkhuff, R. (1976). *Toward Effective Counseling and Psychotherapy: Training and Practice*. New Brunswick: Aldine Transaction.

- Trice, H. M. & Roman, P. M. (1970). Delabeling, Relabeling and Alcoholics Anonymous. *Social Problems*, 17, 538-546.
- Trotter, C. (1996). The Impact of Different Supervision Practices in Community Corrections: Cause for Optimism. *Australian & New Zealand Journal of Criminology*, 29, 29-46.
- Trotter, C. (2013). Effective Supervision of Young Offenders. In P. Ugwudike & P. Raynor (Eds.) *What Works in Offender Compliance: International Perspectives and Evidence-Based Practices* (pp. 227-241). Basingstoke: Macmillan.
- Tsai, Y. (2011). Relationship Between Organizational Culture, Leadership Behaviour and Job Satisfaction. *BMC Health Services Research*. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3123547/>
- Ugwudike, P. (2013). Compliance with Community Orders: Front-Line Perspectives and Evidence-Based Practices. In P. Ugwudike & P. Raynor (Eds.) *What Works in Offender Compliance: International Perspectives and Evidence-Based Practices* (pp. 165-182). Basingstoke: Macmillan.
- Ugwudike, P. (2016). *Symposium on Innovative Approaches to Evaluating Practices in the Justice System*. Swansea: Swansea Service Evaluation Team, Swansea University. Available at: <https://www.swansea.ac.uk/media/Programme%20NEW.pdf>
- Ugwudike, P. & Morgan, G. (2015). *From 'What Works' to 'How Things Work': Strategies for Harmonising Criminal Justice Practices with the Evidence Base*. Paper presented in the 'Changing Probation' panel session at the British Society of Criminology Conference, Plymouth University, July 2015.
- Ugwudike, P. & Morgan, G. (2018). Bridging the Gap Between Research and Frontline Youth Justice Practice. *Criminology & Criminal Justice*, 0, 1-22.
- Ugwudike, P., Raynor, P. & Annison, J. (Eds.) (2018). *Evidence-Based Skills in Criminal Justice: International Research on Supporting Rehabilitation and Desistance*. Bristol: Policy Press.
- VanderVen, K. (1995). Point and Level Systems: Another Way to Fail Children and Youth. *Child & Youth Care Forum*, 24, 345-367.
- VanderVen, K. (2016). *Updated 2016 Digital Point Pack: A Documented Analysis of the Literature on the Destructiveness of "Point and Level Systems" Commonly Employed in Group and Residential Settings, and Schools*. Available at: <http://www.cyc-net.org/Documents/VanderVen%20Point%20Pack.pdf>
- Van Kalmthout, A. M. & Durnescu, I. (Eds.) (2008). *Probation in Europe*. Nijmegen: Wolf Legal Publishers.

- Vanstone, M. (2000). Cognitive-Behavioural Work with Offenders in the UK: A History of Influential Endeavour. *The Howard Journal*, 39, 171-183.
- Vanstone, M. & Raynor, P. (2012). *Observing Interview Skills: A Manual for Users of the Jersey Supervision Interview Checklist*. © The Jersey Crime and Society Project, c/o Jersey Probation and Aftercare Service.
- Van Voorhis, P. (2012). On Behalf of Women Offenders: Women's Place in the Science of Evidence-Based Practice. *Criminology & Public Policy*, 11, 111-145.
- Van Voorhis, P. & Brown, K. (1996). *Evaluability Assessment: A Tool for Program Development in Corrections*. Washington DC: National Institute of Corrections.
- Van Voorhis, P., Cullen, F. T. & Applegate, B. (1995). Evaluating Interventions with Violent Offenders: A Guide for Practitioners and Policymakers. *Federal Probation*, 50, 17-27.
- Van Voorhis, P. & Presser, L. (2001). *Classification of Women Offenders: A National Assessment of Current Practices*. Washington: US Department of Justice, National Institute of Corrections. Available at: <http://static.nicic.gov/Library/017082.pdf>
- Van Voorhis, P., Wright, E. M., Salisbury, E. & Bauman, A. (2010). Women's Risk Factors and Their Contributions to Existing Risk/Needs Assessment: The Current Status of a Gender-Responsive Supplement. *Criminal Justice and Behaviour*, 37, 261-288.
- Von Mackensen, S., Campos, I. G., Acquadro, C. & Strandberg-Larsen, M. (2013). Cross-Cultural Adaptation and Linguistic Validation of Age-Group-Specific Haemophilia Patient-Reported Outcome (PRO) Instruments for Patients and Parents *Haemophilia*, 19, 73-83.
- Vose, B., Lowenkamp, C. T., Smith, P. & Cullen, F. T. (2009). Gender and the Predictive Validity of the LSI-R: A Study of Parolees and Probationers. *Journal of Contemporary Criminal Justice*, 25, 459-471.
- Wallerstein, I. (1981) Concepts in the Social Sciences: Problems of Translation. In M. G. Rose (Ed.) *Translation Spectrum: Essays in Theory and Practice* (pp. 88-98). Albany: State University of New York Press.
- Ward, T. (2002). The Management of Risk and the Design of Good Lives. *Australian Psychologist*, 37, 172-179.
- Ward, T. (2010). The Good Lives Model of Offender Rehabilitation: Basic Assumptions, Aetiological Commitments, and Practice Implications. In F. McNeill, P. Raynor & C. Trotter (Eds.) *Offender Supervision: New Directions in Theory, Research* (pp. 41-64). Cullompton: Willan.

Ward, T. & Hudson, S. M. (1998). The Construction and Development of Theory in the Sexual Offending Area: A Metatheoretical Framework. *Sexual Abuse: A Journal of Research and Treatment*, 10, 47-63.

Ward, T. & Maruna, S. (2007). *Rehabilitation: Beyond the Risk Paradigm*. London: Routledge.

Ward, T. & Stewart, C. (2003). Criminogenic Needs and Human Needs: A Theoretical Model. *Psychology, Crime & Law*, 9, 125-143.

Ward, T., Yates, P. M. & Willis, G. M. (2012). The Good Lives Model and the Risk Need Responsivity Model: A Critical Response to Andrews, Bonta, and Wormith (2011). *Criminal Justice and Behaviour*, 39, 94-110.

Warren, M. (1971). Classification of Offenders as an Aid to Efficient Management and Effective Treatment. *Journal of Criminal Law, Criminology and Police Science*, 62, 239-258.

Weaver, B. and McNeill, F. (2010). Travelling Hopefully: Desistance Research and Probation Practice. In J. Brayford, F. Cowe & J. Deering (Eds.) *What Else Works? Creative Work with Offenders* (pp. 36-60). Cullompton: Willan.

Wellisch, J., Anglin, M. D. & Prendergast, M. L. (1993). Treatment Strategies for Drug-Abusing Women Offenders. In J. A. Inciardi (Ed.) *Drug Treatment and Criminal Justice* (pp. 5-25). Newbury Park, CA: Sage Publications.

Wholey, J. S. (1979). *Evaluation: Promise and Performance*. Washington, DC: The Urban Institute.

Wilkins, L. T. (1958). A Small Comparative Study of the Results of Probation. *British Journal of Delinquency*, 8, 201-209.

Wilkins, L. T. (1969). *Evaluation of Penal Measures*. New York: Random House.

Williams, K. (2018). *Submission to the Commission on Justice in Wales*. Available at: <https://beta.gov.wales/sites/default/files/publications/2018-07/Submission-to-the-justice-commission-from-professor-kate-williams.pdf>

Wilson, D. B., Bouffard, L. A., & MacKenzie, D. L. (2005). A Quantitative Review of Structured, Group-Oriented, Cognitive-Behavioral Programs for Offenders. *Journal of Criminal Justice and Behavior*, 32, 172-204.

Wong, S. & Hare, R. D. (2005). *Guidelines for a Psychopathy Treatment Program*. Toronto: Multi-Health Systems.

Wooditch, A., Tang, L. & Taxman, F. (2014). Which Criminogenic Need Changes are Most Important in Promoting Desistance from Crime and Substance Abuse? *Criminal Justice and Behavior*, 41, 276-299.

Worell, J. & Remer, P. (2003). *Feminist Perspectives in Therapy: Empowering Diverse Women*, 5th edition. Hoboken: John Wiley & Sons.

Wormith, J. S., Gendreau, P. & Bonta, J. (2012). Deferring to Clarity, Parsimony, and Evidence in Reply to Ward, Yates and Willis. *Criminal Justice and Behaviour*, 39, 111-120.

Wright, K. A. (2014) The Private Prison. In F. T. Cullen, C. L. Jonson, M. K. Stohr (Eds.) *The American Prison: Imagining a Different Future*. Thousand Oaks, CA: Sage.