The Efficacy of a psychoeducational intervention for the stabilisation of complex interpersonal trauma symptomatology in female offenders

A thesis submitted in partial fulfilment of the requirements of

Edinburgh Napier University, for the award of Doctor of Philosophy

Adam Mahoney

May 2019

Abstract

Interpersonal trauma is endemic in female prisons. Implementing trauma informed interventions to assist the recovery of women in custody has long been advocated for. It has been argued that psychoeducation should constitute a critical first phase of trauma informed interventions. The main objective of psychoeducation being to stabilise symptoms and behaviours thereby enabling survivors to cope with subsequent trauma memory processing (TMP). Although group based psychoeducation interventions have been frequently delivered they have received scant empirical testing particularly within forensic environments.

The first part of this thesis involved conducting a systematic review and meta-analysis to investigate the efficacy of group based interventions. The results were considered with respect to five outcome domains (PTSD symptoms, Depression, Psychological Distress, Substance Misuse and Dissociation). This was the first time that a detailed analysis was conducted specifically for group interventions. Whilst TMP treatments computed large statistically significant effect sizes, for PTSD symptoms, compared to usual care (*k*=6, g= -0.98 [95%CI, -1.53 to -0.43], psychoeducation interventions (after outliers were removed) had only small non-significant effect sizes *k*=7, g=−0.25 [95%CI −0.66 to 0.16]. However, when TMP and psychoeducation were directly compared only small non-significant differences were apparent in favour of the former (*k*=4, g= -0.34 [95%CI, -1.05 to 0.36]) for the amelioration of PTSD symptoms. Similarly, trauma informed interventions were also as efficacious as non-trauma informed interventions (*k*=5, g= 0.36 [95%CI, -0.24 to 0.96]).

The second part of this thesis concerned a randomised control trial (RCT) which investigated the efficacy of Survive & Thrive, a pure psychoeducational intervention, which was delivered to female prisoners. This brief 10 session intervention (was compressed to a 2 session per week format to accommodate short sentences. Participants who received this intervention (n=44) were compared to those who received usual care (n=42). Results from an intent-to-treat (ITT) analysis indicated that there were no statistically significant differences between the two arms across the three assessment time points (including one month after the intervention) for the main outcomes (Behavioural Assessment Checklist-Revised, β= 4.60 [95%CI, -1.60 to

10.88], p= 0.148; PTSD Checklist, β= -1.47 [95%CI, -4.30 to 1.36], p= 0.303). Subscales from other measures however indicated that participants in the intervention arm reported significantly more Depression (β= 0.95 [95%CI, 0.11 to 1.79], p=.027) and less emotional Non-Acceptance (β= -1.65 [95%CI, -3.22 to -0.07], p=.041). All ITT results were only statistically significant at follow up. However, an adequate dose (≥7 sessions) analysis indicated that interactions between time and study arm were significant at post assessment. This included for the Distress subscale in the main behaviour outcome measure (β= -3.51 [95%CI, -6.55 to -0.47], p=.024). *Post hoc* Reliable Change analyses suggested twice an many AD participants made progress in addressing PTSD symptoms compared to usual care (30.3% vs 17.6%, OR 2.03 [95%CI, 0.64 to 6.43]).

The trial undertaken for this thesis is the first comprehensive RCT for a group based psychoeducational intervention with female prisoners. The clinical and research utility of the results from this trial and the meta-analysis are discussed with respect to the stabilisation and amelioration of symptomatology associated with interpersonal trauma. It suggested that there is still further work to be done if psychoeducational interventions are to demonstrate greater efficacy than usual care.

Acknowledgements

This thesis is dedicated to the memory of my father Terrance Mahoney who sadly passed away during the early stages of this thesis. It is also dedicated to my wonderful son Oliver who was also born during the early stages of this thesis. A crucial member of my support team is of course my wife the gorgeous and incredibly intelligent Dr Jenny Mahoney whose patience and fortitude has enabled me to complete this.

Similarly, none of this would have been possible without the encouragement, support and supervision from Prof Thanos Karatzias, Dr Paul Hutton and Dr Nadine Dougal. I am extremely grateful for all of their time and feedback.

It is also important to acknowledge the support that I received from colleagues in the Scottish Prison Service whose kindness enabled this thesis to happen. Not least is the very wise and kind Kirsty Halliday, my line manager for many years. There are also many, many other colleagues to thank, including Dr Sandra Ferguson, and of course all of the women who participated in the trial.

Contents

		page
Abstract		ii
Acknowledgements		iv
Contents		v
List of appendices		x
List of tables		xii
List of figures		xiv
Glossary of terms		xv
1. Introduction to the thesis		1
1.1.	The importance of addressing interpersonal trauma	1
1.2.	Interpersonal trauma: definition, conceptualisation and classification	2
1.3.	Theories explaining the impact of interpersonal trauma	4
1.4.	Psychological interventions for interpersonal trauma	6
1.4.1.	Phase 1 psychoeducational interventions for interpersonal trauma	8
1.5.	The importance of considering interpersonal trauma in forensic populations	10
1.5.1.	Trauma informed interventions in prisons	12
1.5.2.	The evidence base for interpersonal trauma interventions in forensic populations	13
1.6.	Developing the evidence base for psychoeducational interventions in prison populations	20
1.6.1.	Survive & Thrive: a psychoeducational intervention for complex trauma	22
1.6.2.	Aims and objectives of the thesis	23
2. A systematic review and meta-analysis of group treatments for adults with symptoms associated with complex post-traumatic stress disorder		25
2.1.	Introduction	25
2.1.1.	Complex post-traumatic stress disorder and phase based treatments	25

2.1.2.	Group versus individual treatment modalities	26
2.1.3.	Previous meta-analyses	27
2.1.4.	Research questions	29
2.2.	Methods	29
2.2.1.	Search and inclusion criteria	29
2.2.2	Data-extraction and categorisation	31
2.2.3.	Risk of bias and coding of methodological quality	33
2.2.4.	Outcomes	33
2.2.5.	Meta-analysis	34
2.2.6.	Heterogeneity	35
2.2.7.	Moderator analysis	35
2.3.	Results	36
2.3.1	Study characteristics	36
2.3.2.	Methodological quality of included studies	48
2.3.3.	Treatment effects	51
2.3.4.	Heterogeneity	56
2.3.5.	Publication bias	56
2.3.6.	Moderator analysis	57
2.4.	Discussion	58
2.4.1.	Summary of findings	58
2.4.2.	Are TMP treatments more effective than psychoeducational treatments?	59
2.4.3.	Comparison of findings with other meta-analysis	59
2.4.4.	Implications for clinical practice	60
2.4.5.	Implications for future research	62
2.4.6.	Conclusions	64
3.	**Methods: Randomised control trial**	**66**
3.1.	Introduction to Survive & Thrive (Prison Version) Randomised Control Trial	66
3.1.1.	Developing an evidence based for the treating female prisoners with interpersonal trauma	66

3.1.2.	Intervention resource: selection and justification	66
3.2.	Survive & Thrive Prison Version	68
3.2.1.	Mutualisation and responsivity of the intervention	68
3.2.2.	Treatment theory and psychoeducational techniques	68
3.2.3	Treatment targets	70
3.2.4.	Dose of treatment and method of delivery	72
3.3.	Research questions	72
3.3.1.	Hypothesis	73
3.4.	Design and rationale of the randomised control trial	73
3.5.	Trial registration	75
3.6.	Ethical approval	75
3.7.	Study setting: delivery teams and locations	76
3.8.	Standardisation and training	77
3.8.1.	Treatment supervision	77
3.9.	Sample and selection	78
3.9.1.	Sample size calculations	80
3.10.	Control comparator	83
3.11.	Outcome measurements	83
3.11.1.	Demographics and sentencing characteristics	83
3.11.2.	Baseline trauma profile measures	84
3.11.3.	Main outcome measures (MOM): behavioural change and PTSD symptom stabilisation	86
3.11.4.	Other outcome measures (OOM): alleviation of psychological distress	89
3.11.5.	Theoretical outcome measures (TOM): rehabilitation of criminal thinking styles	87
3.11.6.	Administration schedule	92
3.12.	Trial procedures and management	92
3.12.1.	Recruitment	94
3.12.2.	Informed consent	95
3.12.3.	RCT management and fidelity	95
3.12.4.	Quality assurance / treatment integrity and fidelity	98

3.12.5.	Data storage and protection	98
3.12.6.	Safety and ethical considerations	99
3.13.	Analysis plan	100
3.13.1.	Data checking and cleaning	100
3.13.2.	Missing data	101
3.13.3.	Primary data analysis – Intention-to-treat (ITT) analysis	101
3.13.4.	Secondary data analysis –– adequate dose (as per completer protocol)	104
3.13.5.	Reliable and clinically significant change analysis	104
3.14.	Summary	105
3.	**Results: Randomised control trial**	**107**
4.1.	Introduction	107
4.2.	Study participants	107
4.2.1.	Demographic and sentencing characteristics	108
4.2.2.	Trauma characteristics at baseline	109
4.3.	Missing data analysis	115
4.3.1.	Comparisons between outcome completers and missing data Participants	116
4.3.2.	Missing completely at random (MCAR) / Missing at random (MAR) analysis	117
4.4.	Psychometric analysis	119
4.5.	Fidelity analysis: quality assurance of treatment facilitation	121
4.6.	Efficacy of Survive & Thrive: Intent-to-treat analysis	123
4.7.	Efficacy of Survive & Thrive: Adequate dose analysis	132
4.7.1.	Comparison of adequate dose and non-adequate dose participants	132
4.7.2.	Comparison of adequate dose and WL/TAU arm participants	137
4.8.	Clinically significant and reliable change analysis	144
4.9.	Summary	150
5.	**Discussion**	153
5.1.	Summary of the thesis	153
5.1.1.	Meta-analysis findings	154
5.1.2.	RCT findings	156

5.2. Comparisons with previous research 158
5.3. Summary of contributions to the literature 160
5.4. Clinical practice implications 161
5.5. Strength and limitations 167
5.6. Directions for future research 172
5.7. Overall conclusions 175
6. References 178
7. Appendix
1. Appendix for systematic review and meta-analysis 208
2. Appendix for methods: Randomised control trial 221
3. Appendix for results: Randomised control trial 243

List of appendices

1.	**Appendices for systematic review and meta-analysis**	**209**
1.1.	Literature search strategy	209
1.2.	Authors categorisation decisions for treatment arms	210
1.3.	Funnel for group-based trauma interventions compared to usual care	226
1.4.	GRADE and risk of bias outcomes	227
1.5.	Summary of moderator findings across outcome domains	242
2.	**Appendices for methods: Randomised control trial methods**	**245**
2.1.	Psychometric properties of outcome measures	245
2.2.	Missing data	252
2.3.	Statistical analysis and interpretation	255
2.3.1.	Effect size	255
2.3.2.	Linear mixed models	257
2.3.3.	Interpretation of LMM output	265
3.	**Appendix for results: Randomised control trial**	**267**
3.1.	Complete results for the Trauma Antecedent Questionnaire (TAQ)	269
3.2.	Complete results for the SIDES –Self Report	270
3.3	Missing data analysis	273
3.4.	Investigation into quality assurance of treatment facilitation	278
3.5.	Differences between intervention arms for outcome measures	279
3.5.1	PTSD Checklist –Civilian Version	279
3.5.2.	Behavioural Assessment Checklist - Revised	282
3.5.3.	Difficulties in Emotional Regulation Scale	286
3.5.4.	Dissociative Experiences Scale	289
3.5.5.	Hospital Anxiety Depression Scale	292
3.5.6.	Criminal Cognitions Scale	293
3.6.	Differences between adequate dose and non-adequate dose participants	296
3.6.1	PTSD Checklist –Civilian Version	296
3.6.2	Behavioural Assessment Checklist - Revised	299
3.7.	Differences between Adequate Dose and Control Participants	303

3.7.1	PTSD Checklist –Civilian Version	303
3.7.2	Behavioural Assessment Checklist -Revised	305
3.7.3	Difficulties in Emotional Regulation Scale	307
3.7.4	Dissociative Experiences Scale	310
3.7.5.	Hospital Anxiety Depression Scale	312
3.7.6.	Criminal Cognitions Scale	313
3.8.	Plots for reliable change scores	316

List of Tables

Introduction		
1	Summary of previous RCTs with incarcerated populations	18
Systematic Review and Meta-analysis		
2	Summary of study characteristics	37
3	Summary of study and treatment characteristics	44
4	Summary of participant characteristics	46
5	Cochrane risk of bias ratings and methodological quality for included studies	49
6	Effect size estimates for different treatment categorisations and comparisons	53
Methods: Randomised control trial		
7	Components of the Survive & Thrive Prison Version intervention	70
8	Administration schedule for measures included in the RCT	92
Results: Randomised control trial		
9	Sociodemographic, forensic and clinical characteristics of participants	112
10	Frequency and type of traumatic events at baseline	113
11	Summary of lifetime occurrence, current presence and current severity of trauma symptom clusters at baseline	114
12	Missing data count	116
13	Results from Little's missing completely at random (MCAR) test	118
14	Cronbach's alpha for the Behavioural Assessment Checklist-Revised and subscales	119
15	Cronbach's Alpha for the Criminal Cognitions Scale and subscales	121
16	ITT outcomes for all measures across time points and study arms	127
17	ITT analysis of the interaction between study arm and time	130
18	PCL-C and BAC-R outcome data and analysis of the interaction between study arm and time based on adequate dose and non-adequate dose participants	135
19	ITT outcomes for all measures across time points and study arms as adjusted for AD S&T participants	139

20	ITT comparison of the control and AD S&T treatment arms at T1 - T3	142
21	PCL- C clinically significant and reliable change for S&T AD and control arm at T2 and T3	146
22	Clinically significant and reliable change for AD S&T and control arms at post treatment	149

List of figures

Introduction		
1	PTSD and complex PTSD symptoms: in classification hierarchy	4
Systematic review and meta-analysis		
2	PRISMA flow diagram	31
3	Forest plot for PTSD symptoms post treatment effect sizes estimates as grouped by comparators	56
Methods: Randomised control trial		
4	Study enrolment and randomisation procedure	93
5	Planned study and design and assessment points	93
Results: Randomised control trial		
6	CONSORT diagram	111
7	Frequency of participants meeting the diagnostic criteria for SIDES-SR symptom domains	115
8	Frequency and range of the composite treatment quality assurance ratings achieved for delivery of the Survive & Thrive intervention	122
9	Comparison between WL/TAU control and S&T intervention outcome data for PCL-C total scores.	125
10	Comparison between WL/TAU control and S&T intervention outcome data for BAC-R total scores	126
11	Frequency of sessions completed by S&T arm participants	132
12	Comparison between AD and Non-AD S&T intervention participants outcome data for PCL-C Total Scores	133
13	Comparison between AD and non-AD S&T intervention participants outcome data for BAC-R Total Scores	133
14	PCL-C total reliable change score for AD participants in the S&T arm at T2	146
15	PCL-C total reliable change score for participants in the control arm at T2	147
16	BAC-R at T2 change for AD participants in the S&T arm	147
17	BAC-R at T2 changefor participants in the control arm	148

Glossary of terms

Abbreviation	***Term***
PTSD	Post-traumatic stress disorder
CPTSD	Complex post-traumatic stress disorder
RCT	Randomised control trial
S&T	Survive & Thrive (i.e. experimental arm)
WL / TAU	Waiting list / treatment as usual (i.e. control arm)
ITT	Intent-to-treat
AD	Adequate dose
LMM	Linear mixed model
ICC	Intra-class correlation coefficient
MPD	Missing participant data
FIML	Full information maximum likelihood
MI	Multiple imputation
T1	1st time point i.e. 'pre' intervention psychometric administration
T2	2nd time point i.e. 'post' intervention psychometric administration
T3	3rd time point i.e. 'follow up' intervention psychometric administration
SIDES-SR	Self-report instrument for disorders of extreme stress (baseline measure)
TAQ	Trauma antecedents questionnaire (baseline measure)
PCL-C	PTSD checklist- civilian version (main outcome measure)
BAC-R	Behavioural assessment measure – revised (main outcome measure)
DERS	Difficulties with emotional regulation scale (outcome measure)
DES	Dissociative experiences scale (outcome measure)
HADS	Hospital anxiety and depression scale (outcome measure)
CCS	Criminal cognitions scale (outcome measure)
CSC	Clinically significant change
RC	Reliable change (RCI= Reliable change index)
RA	Research assistant
PI	Principle Investigator

1. Introduction to the thesis

1.1. The importance of addressing the impact of interpersonal trauma

Interpersonal trauma is recognised internationally as being one of the leading causes of mental health and physical health disabilities (Felitti, Anda, Nordenberg, Williamson, Spitz, Edwards, Koss & Marks, 1998) and is associated with poorer social, economic and criminal justice outcomes as well as shorter life spans (Bowen, Jarrett, Stahl, Forrester & Valmaggia, 2018; Bywaters, Bunting, Davidson, Hanratty, Mason, McCartan & Steils, 2016). Interpersonal trauma is also highly prevalent. Bowen et al (2018), for example, summarised the prevalence of childhood abuse rates from international surveys and concluded that 5.3–10.8% of individuals have experienced physical abuse, 0.6–2.4% experienced sexual abuse and 3.6–5.2% from neglect. These rates are substantially higher for prison based populations with studies indicating that childhood sexual abuse may be as high as 68% for female prisoners with even higher rates (91%) of childhood and adulthood trauma (Browne, Miller & Maguin, 1999; Karatzias, Power, Woolston, Apurva, Begley et al, 2018).

From a public health perspective, there is also a growing recognition by governments, policy makers and institutional bodies of the need to address the burden that interpersonal trauma places on individuals and societies (Magruder, McLaughlin & Borbon, 2017; Scottish Government, 2017). This mental health burden is particularly evident in cases of repeated and extensive interpersonal violence and the resulting symptoms, often associated with complex posttraumatic stress disorder, can persist over the life course (Kessler, Aguilar-Gaxiola, Alonso et al, 2017). Similarly, it is important to note that experiences of interpersonal trauma vary with socio-economic circumstances and gender. For example, Kessler et al (2017) found that women are significantly more likely to experience intimate partner sexual violence and men more likely to experience physical violence and accidents. Understanding gendered ‘pathways’ into, and of course leading out of, interpersonal trauma is important if interventions are to be effective. This is particularly so for institutions where there are high rates of interpersonal trauma such as in women’s prisons (Ney, Van Voohris & Lerner, 2011; Mahoney, 2011). Indeed, understanding and developing the efficacy of

psychological interventions is an ethical and social imperative. It is the intention of this thesis to contribute towards this endeavour.

1.2. Interpersonal trauma: definition, conceptualisation and classification

It is important from the outset to provide an operational definition of interpersonal trauma and complex posttraumatic stress disorder (CPTSD). It is intended that this definition will both help to provide a focus for the thesis as a whole, but also to ensure, where possible, that there is consistency within the literature being referenced. Interpersonal trauma refers to the psychological distress that is often a result of repeated abuse, particularly at key developmental stages, and which can have a profound impact throughout an individual's life (Courtois & Ford, 2016, Enlow et al, 2013). The interpersonal violence constituting this trauma can be emotional, sexual and physical as well as other forms of neglect and intimate partner violence. Experiences of war, torture and exploitation are also included in this respect (Herman, 1992). As the World Health Organisation (WHO) argues intentionality aside the resulting impact on an individual's health and wellbeing is of most concern when defining interpersonal violence (WHO, 2002). A number of authors including Courtois & Ford (2009, pg.1) have described traumatic stressors as: '1) repetitive and prolonged; 2) involving direct harm and/or neglect and abandonment by caregivers or ostensibly responsible adults; 3) occur at developmentally vulnerable times in a survivor's life; and 4) have the potential to compromise severely a child's development. For example, Beck et al (2009) highlights the traumatic stress associated with survivors of motor vehicle accidents (MVA) as often being less pronounced than survivors of interpersonal traumas, such as sexual assault. Indeed, childhood sexual abuse is regarded as a particularly devastating form of interpersonal violence (Briere & Elliot, 2003). As such, it is the developmental sequelae from multiple and often sustained abusive experiences that authors have sought to distinguish from succinct 'single-event' PTSD (Cloitre, Stolbach, Herman, van der Kolk, Pynoos, Wang & Petkova, 2009). Herman (1992) also used the term 'complex' trauma in her seminal text in which she explored the impact of various forms of interpersonal abuse. In the foreword to Courtois & Ford (2009) Judith Herman writes:

…complex PTSD…always begin[s] with the social ecology of prolonged and repeated interpersonal trauma. There are two main points to grasp here. The first is that such

trauma is always embedded in a social structure that permits the abuse and exploitation of a subordinate group… The second point is that such trauma is always relational. It takes place when the victim is in a state of captivity, under the control and domination of the perpetrator (page xiv).

Prior to the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V, APA, 2013) there had been little formal recognition of the complex manifestations of PTSD (APA, 2000). However, in practice this has long been challenged and a DSM task force sought to provide a comprehensive set of diagnostic criteria to define complex trauma which at one stage became termed as Disorders of Extreme Stress Not Otherwise Specified (DESNOS) (Roth, Newman, Pelcovitz, Van der Kolk, & Mandel, 1997). Similarly, the term Developmental Trauma Disorder (DTD) has also been used in recognition of the impact that chronic traumatic events in childhood and adolescence can have particularly at key developmental stages (van der Kolk, 2005). The current DSM V (APA, 2013) introduced a new criterion accounting for symptoms associated with negative alterations in mood and cognitions as well as a dissociative and pre-school subtype. The former subtype, which includes depersonalisation and derealisation, accounts for developmental differences in the modulation and reaction to trauma and expressions of PTSD. As such it is perhaps most similar to the definition of complex trauma that has been argued for (UKPTS, 2017).

Whilst a full understanding of interpersonal trauma may still be evolving the diagnostic criteria of PTSD has moved from beyond being simply defined as an anxiety disorder to including a range of other phenotypes such as dysphoric/anhedonic and those involving externalising behaviours (Friedman, 2013). Similarly, Cloitre, Garvert, Brewin, Bryant & Maercke (2013) described how work on the recently released *International Classification of Diseases*, 11th version (ICD-11, WHO, 2018) has approached the definition of complex PTSD. This definition includes the presence of core PTSD symptoms (avoidance, re-experiencing, alterations in arousal and reactivity such as hyperarousal/hypervigilance) and also persistent and pervasive symptoms in three specific domains:

- *Affect functioning*: Dysregulation as seen in emotional reactivity, violent outbursts and a tendency towards dissociative states when under stress
- *Self-functioning*: Pervasive feelings of shame and guilt as well as beliefs concerning the self as diminished, defeated or worthless.
- *Relational functioning*: Difficulties in sustaining interpersonal relationships or feeling close to others.

(Cloitre et al, 2013)

This is summarised in Figure 1 below.

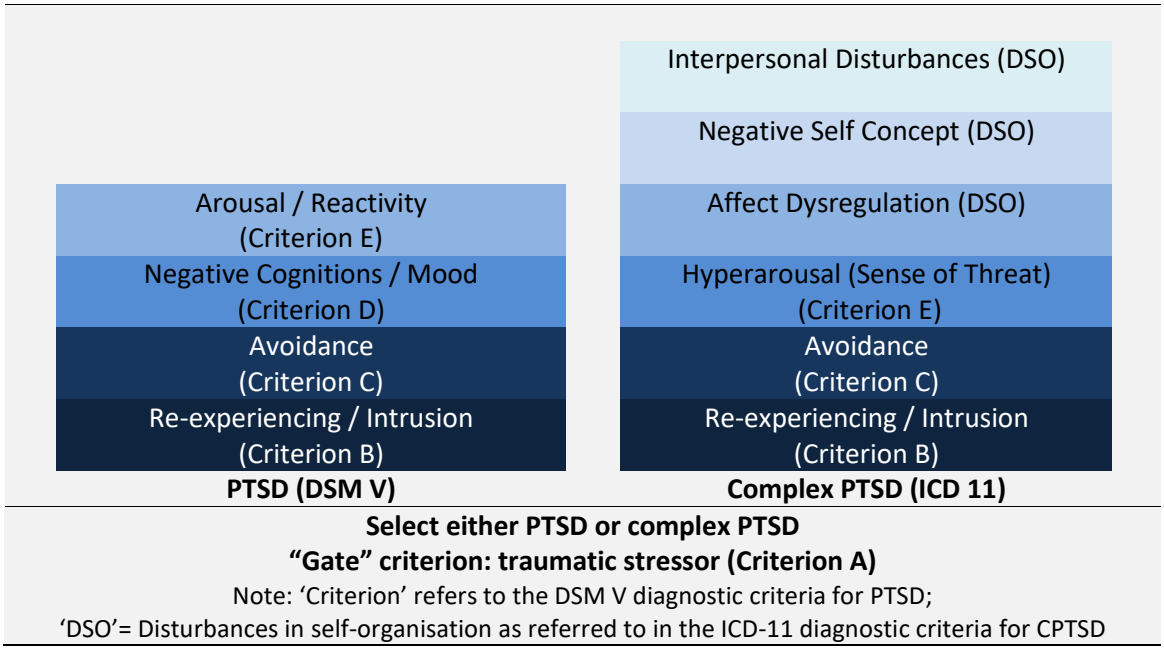


Figure 1: PTSD and complex PTSD symptoms in a classification hierarchy

Considerable work has been undertaken to distinguish 'complex' PTSD from single incident PTSD (Hyland, Shevlin, Fyvie & Karatzias, 2018). However, conceptualising the relational aspect of interpersonal violence and trauma is key to understanding the associated psychopathology and thereby avoiding a 'one size fits' all approach to symptom manifestation (Cloitre, 2015).

1.3. Theories explaining the impact of interpersonal trauma

Physiological explanations have considered the impact of traumatic experiences on the sympathetic nervous system and other associated neurological and endocrinological systems (Lazaratou, 2017). The intergenerational transmission of

environmental stress from the mother to the foetus as well as pre-verbal experiences, such as witnessing violence and the mother's emotional state, have also been accounted for by biological based explanations (Opendak & Sullivan, 2016). Neurobiological studies, for example, have highlighted how molecular and cellular mechanisms mediate stress reactions including the imprinting of traumatic experiences that directly activate neurobiological circuits (e.g. the amygdala) associated with emotion. These neurological repercussions can result in a constant state of alarm and hypervigilance compromising psychological, cognitive and social functioning. Longitudinal studies have, as already mentioned, demonstrated a range of social outcomes (e.g., homelessness, prostitution, delinquency, criminal behaviour, and inability to maintain employment) as well as health outcomes (e.g., heart disease, cancer, chronic lung disease, liver disease, skeletal fractures, HIV-AIDS) from such stress reactions (Felitti et al, 1998). Cloitre et al (2009) also demonstrated in a series of studies that a greater frequency of adult and particularly childhood trauma can result in more severe and qualitatively different outcomes across multiple domains.

The relevance of clinical presentation associated with interpersonal trauma has been of interest to those seeking to address offending behaviour (Jones, 2015; Lewis, 2010). The impaired ability to access cortical areas associated with rational thought and problem-solving in highly stressful situations may have important implications for violent offending. For example, a number of studies have outlined the developmental impact that prolonged abuse can have on important neurophysiological structures that place individuals at a higher risk of violence responses (Heide & Solomon (2006). These studies have increasingly evidenced the need to formulate the emotional and impulsive responses associated with violent offenders who have a history of complex PTSD and a tendency towards prolonged dissociative states when distressed (Heide & Solomon, 2009; Howard, Karatzias, Power & Mahoney, 2017; Jones, 2015; Mahoney & Karatzias, 2012; Weierstall, Huth, Knecht, Nandi, Elbert, 2012).

Understanding trauma based adaptations to the limbic system and the Hypothalamic-Pituitary-Adrenal (HPA) axis is particularly critical when comprehending the development of problematic coping strategies and externalising behaviours such as violence (Howard, Karatzias, Power, & Mahoney, 2017a; Puetz, Zweerings, Dahmen, Ruf, Scharke, Herpertz-Dahlmann & Konrad, 2016; Sadeh & McNiel, 2015). Haller

(2018), for example, hypothesises that over time exposure to abusive circumstances can downregulate critical HPA endocrine systems involved in glucocorticoids (a class of corticosteroids) release. The subsequent impact on neural functioning results in a 'proneness' to violence through an increased stress sensitivity ('allostatic crash') as well as deficits in prefrontal cortex functioning. Similarly, a hypo-functioning HPA axis is linked to brain circuitry involved in predatory aggression as well as other response structures (such as the central amygdala) also resulting disproportionate forms of aggression in socially challenging situations.

As Moretti, Odgers Reppucci, Catherine (2011) have articulated, factors such as personality and temperament along with interpersonal expectations and attributions (e.g. rejection sensitivity, attachment style and rumination) all have an important mediating role within the behavioural deficits that lead to offending. Elevated levels of maladaptive schemas, particularly those involving 'Disconnection', 'Impaired Autonomy', 'Vulnerable to Harm' have also been observed as relevant to psychopathological outcomes (Karatzias, Jowett, Begley & Deas, 2016). In this respect, many offenders undertaking interventions in prisons report disturbances within the self-concept domain, often as a result of adverse life events (Mahoney, Chouliara & Karatzias, 2015). Such disturbances in self-organisation (DSO), as associated with complex PTSD, are often linked to the relational difficulties that are frequently cited as important correlates to female offending (de Vogal & Nicholls, 2016). Whilst anti-social peers and a lack of pro-social personal support are frequently cited as risk factors within the male offending literature they also have relevance to female offenders and may be further manifestations of DSO symptoms (Stewart, 2015).

1.4. Psychological interventions for interpersonal trauma

The efficacy of standard evidence-based treatments for PTSD needs to be considered when treating more complex forms of trauma. There is evidence to suggest that complex forms of interpersonal trauma may be less amenable to standard psychological interventions and may require different types of treatment as well as a specific sequence of treatments (Corrigan & Hull, 2015; Finn, Warner, Price & Spinazzola, 2018; Gene-Cos, Fisher, Ogden & Cantrel, 2016). As such the treatment of complex trauma focuses on both the reductions of PTSD symptoms as well as

important self-regulation capacities and developing psychosocial and environmental strengths (Cloitre, Courtois, Ford, Green, Alexander, Briere, Herman, Lanius, Stolbach, Spinazzola, van der Kolk, & van der Hart, 2012).

Herman (1992) described recovery from interpersonal trauma as first requiring the restoration of power and control. As such creating a sense of relational safety for survivors may be an essential prerequisite before trauma memory processing (TMP) and for the development of supportive social connections to occur. Various authors have also advocated that complex trauma responsive interventions should adopt a phased based approach to promoting recovery (Ford, Cortois, Steele, van der Hart & Mijenhuis, 2005; Harris & Fallot, 2001; Herman, 1992; van der Hart, Brown & van der Kolk, 1989). Herman (1992) described a 3 phased model to trauma recovery, and whilst there is some variation in how authors describe these stages, phase 1 has been conceptualised as a psychoeducation stage where coping skills are provided and safety and stabilisation established. Phase 2 interventions are conceptualised as assisting survivors in TMP thereby developing an understanding of their experiences often with an emphasis on remembrance and mourning. Whilst PTSD interventions have often focused on the recall and reconsolidation of a single event CPTSD interventions may be more focused on assisting the individual develop a more coherent organisation of a lifetime of abusive experiences (Cloitre et al, 2012; Miller, 2016; Parnell, 1999). In phase 3, the final stage, integration and reconnection with ordinary life are promoted through consolidating progress particularly in the areas of safety and trust.

The current clinical guidelines question the effectiveness of standard PTSD treatment protocols such as trauma-focused cognitive behavioural therapy (TF-CBT) and eye movement desensitisation and reprocessing (EMDR) for complex interpersonal trauma (NICE, 2018). Although there is good evidence for these treatments with PTSD it is recognised that individuals with more complex presentations and comorbidities often do not fully recover (Cloitre et al, 2012). As such the three-stage process, as outlined above, has been advocated although it is also recognised that there is limited evidence for the effectiveness of the stabilisation and reintegration/reconnection protocols within current CPTSD treatment (NICE, 2018).

As noted although the implementation of phase 1 interventions appears to make clinical sense and have been used in practice based settings for a long time, there is a limited evidence for their effectiveness (de Jongh et al, 2016). Similarly, whilst there is a dearth of research detailing the transition between these phases it may be speculated that not all individuals will require assistance and therapeutic input at each stage. Sufficient recovery may be achieved in the initial stages to enable some survivors to continue addressing trauma symptoms either on their own or through informal support networks. Whilst, it can make intuitive sense to provide efficacious psychoeducational and supportive interventions, particularly in organisations that serve large populations of survivors, there is little empirical evidence that in themselves they promote recovery (Covington, Burke, Keaton & Norcott, 2008; Hoge & Chard, 2018; Zlotnick, Johnson & Najavtis, 2009).

1.4.1. Phase 1 psychoeducational interventions for interpersonal trauma

The essential components of trauma focused interventions have been considered as including improvements in the emotional processing of trauma memories, through repeated exposure; and improvements in the 'meaning' attributed to traumatic events (Resick, O'Brien Uhlmansiek, Clum, Galovski, Scher & Young-Xu, 2008). In this respect, an important function of psychoeducation can be regarded as helping survivors develop their understanding of their experiences and ameliorating associated distress. Many such interventions use CBT approaches to enable an understanding of how maladaptive beliefs can maintain levels of psychopathology (Kubany, 2003). Other psychoeducational approaches seek to stabilise maladaptive behaviours through practical strategies that decrease affective distress and perceptual biases as well as increasing self-awareness and adaptive coping strategies (Cohen, Mannarino, Kliethermes, Murray, 2012; Ford et al, 2005; Harris & Fallot, 2001).

Considering which components of psychoeducational interventions are efficacious is important given the considerable variability in their content, delivery and length (Dorrepaal, Thomaes, Smit, van Balkom, Veltman, Hoogendoorn, & Draijer, 2012; Wessely, Bryant, Greenberg, Earnshaw, Sharpley & Hacker Hughes, 2008). Whilst most psychoeducational interventions are 'brief' some may be extremely brief and this may as a consequence have either very limited or no benefit. Rose, Brewin, Andrews & Kirk

(1999), for example, demonstrated that clients who received 30 minutes of psychoeducation did not experience any change in their presenting trauma symptomology and were also unlikely to seek further assistance. However, it can also be hypothesised that psychoeducation of a longer duration could have an important preparatory role in successfully orientating survivors towards further treatment. As such, psychoeducation might help to overcome difficulties that result in higher dropout rates for other types of interventions (McDonagh, Friedman, McHugo, Ford, Sengupta, Mueser, Demment, Fournier, Schnurr, Descamps, 2005). In this respect, psychoeducation has often been incorporated as an important aspect of trauma based treatment protocols (UKPTSD, 2017).

The potential cost effectiveness of psychoeducational interventions also needs consideration. Factors such as being time-efficient, reducing waiting lists, being more accessible and responsive to the competing demands in participant's lives and reducing the need for highly trained therapists can all promote the decision to deliver such interventions (Nollett, Lewis, Kitcher et al, 2018). In addition, the large scale delivery of widely accessible interventions may have direct cost implication at an individual, service and societal level (Brookes, Barrett, Netten & Knapp, 2013). In this respect considering the unit costs of delivering psychoeducational interventions versus potentially more complex TMP interventions, often delivered on an individual basis, needs to be calculated. Indeed, the cost of unresolved trauma, including from survivors who become involved in the criminal justice system, may influence the perceived overall effectiveness of psychoeducational interventions and other low-intensity forms of treatment (Ali, Rhodes, Moreea et al, 2017; Kezelman, Hossack, Stravropoulos & Burley, 2015; Peleikis & Dahl, 2005).

Whilst group treatment modalities have been used for both psychoeducational and TMP interventions, it has been particularly linked to the former in areas such as skill acquisition (Sayın, Candansayar, & Welkin, 2013). A number of authors have also described the advantages of a group format for helping lessen the feelings of stigma, isolation and shame that frequently follow interpersonal trauma (Herman, 2007; Talbot, Houghtalen, Cyrulik, Betz, Barkun, Duberstein, Wynne, 1998). However, previous meta-analyses have suggested that individual treatments or treatments containing individual sessions might be more efficacious than group based

interventions alone, although they have not distinguished between psychoeducational and TMP interventions (Ehring, Welboren, Morina, Wicherts, Freitag & Emmelkamp, 2014; Sloan, Feinstein, Gallagher, Beck, & Keane, 2013; Taylor & Harvey, 2010). Whilst in general there is limited evidence on this issue it can be concluded that most clients will receive a combination of individual and group psychotherapy (Bilić, Nemčić-Moro, Karšić, Grgić, Stojanović-Špehar & Marčinko, 2010).

De Jongh et al (2016) noted that there is currently limited research using randomised control trials (RCTs) to establish the efficacy of a phased based approach. Similarly, to date there are only a limited number of RCT studies evaluating what can be identified as psychoeducation or 'present-centred' interventions (Classen, Palesh, Cavanaugh, et al, 2011; Dorrepaal et al, 2012). In this respect, a number of interventions whilst including elements of stabilisation and psychoeducation also include protocols that assist with TMP or which focus on the dual task of managing substance misuse and traumatic stress (for example *Seeking Safety*, Najavits, 2002). Similarly, most research into the efficacy of interpersonal trauma treatment is from the USA (Taylor & Harvey, 2010). Given the different health care, penal, social and even diagnostic systems in different countries careful, consideration needs to be given when generalising any research findings beyond specific groups of participants (Moloney & Kelly, 2004). As such conducting research into establishing how social and cultural factors, including regionally and ethnically experienced deprivation and trauma, is important in understanding the applicability of internationally recognised guidelines for PTSD/CPTSD (Bass, Annan, McIvor-Murray, Kaysen, Griffiths, Cetinoglu, Wachter, Murray & Bolton, 2013; Johnstone & Boyle, 2018; Kaslow, Leiner, Reviere, Jackson, Bethea, Bhaju, Rhodes, Gantt, Senter & Thompson, 2010; Yeomans, Forman, Herbert & Yuen, 2010).

1.5. The importance of considering interpersonal trauma in forensic populations

As noted previously, the prevalence of abusive and adverse life experiences is not randomly distributed and can vary greatly depending upon gender (Kessler et al, 2017). Differences in the reported prevalence rates have also been noted in forensic populations although it is frequently concluded that female offenders have particularly high rates of adulthood and childhood abuse. For example, Browne, Miller

& Maguin (1999) noted in their particular sample of female prisoners that the prevalence of severe childhood physical abuse was 70%, child sexual abuse 59%, and that more than half of childhood sexual experiences occurred before the age of 10 years. Browne et al (1999) also noted that three-quarters of female prisoners reported physical violence by an intimate partner in adulthood and 77% of respondents indicated that they had experienced interpersonal (physical or sexual) violence by individuals exclusive of childhood or intimate partners. In the UK, Loucks (1997) reported that 46% of Scottish female prisoners had experienced sexual abuse and 82% abuse of any type across the lifespan. Recently, Karatzias et al (2018) also noted the high prevalence of multiple trauma histories (91%) in their sample of Scottish female prisoners and that 58% met the diagnostic criteria for DSM-V PTSD. This recent study established that PTSD symptoms, as measured by the Post-Traumatic Stress Disorder Checklist (PCL-5), did not mediate the relationship between childhood trauma, sentence length, age at first offence or emotional regulation. However, adulthood psychological trauma was a statistically significant mediator between childhood trauma and sentence length. Other findings from this study included a statistically significant association between PTSD and age of first offence. These findings provide further evidence for interpersonal trauma, i.e. multiple traumatic life experiences, are an important correlate to various forensic outcomes.

The studies reported above have based their prevalence data on self-report rather than clinical interviews and as such outcomes maybe subject to over diagnosis (Bowen et al, 2018). However, as Greene, Ford, Wakefield & Barry (2014) note clinical interview procedures operate on the basis of a defined categorical diagnosis and therefore may exclude important sub-clinical levels of PTSD. As Moloney, van den Bergh & Moller (2009) argue in their review of the international literature on the prevalence of trauma histories of incarcerated women the empirical evidence is ‘scarce, contradictory and subject to under-reporting’. Considering the dearth of international prevalence studies it is easy to conclude that this continues to be the case. The importance of addressing these concerns can be seen in the greater occurrence of mental health problems and reconviction rates in prisoners with interpersonal trauma histories (Bowen et al, 2018; Mahoney & Karatzias, 2012;

Maniglio, 2009; Ministry of Justice, 2010; Sarchiapone, Carli, Cuomo, Marchetti, & Roy 2009).

1.5.1. Trauma informed interventions in prisons

There is very little literature on the treatment of interpersonal trauma based difficulties within the offender literature with some notable exceptions (Koltz, 2012; Jones, 2015). The psychological therapies available to offenders in prison therefore consist of offending behaviour programmes, largely designed according to the Risk-Need-Responsivity model which specifies which factors will reduce recidivism (Andrews & Bonta, 2010; McDougall, Clarbour, Perry, & Bowles, 2009). These have been largely considered as separate to mental health interventions designed to alleviate psychological distress (Polaschek, 2011). Obviously, a distinction needs to be made between offending behaviour, which is not attributable to the mental health or behavioural sequelae associated with interpersonal trauma and those behaviours which have a clear trauma aetiology. Whilst such a distinction may present as simplistic, and may indeed be unhelpful, the social and emotional reactions both to offending as well as to interpersonal trauma needs to be considered. As Herman (1992, pg. 388) noted:

Observers who have never experienced prolonged terror, and who have no understanding of coercive methods of control, often presume that they would show greater psychological resistance than the victim in similar circumstances. The survivor's difficulties are all too easily attributed to underlying character problems, even when the trauma is known. When the trauma is kept secret, as is frequently the case in sexual and domestic violence, the survivor's symptoms and behavior may appear quite baffling, not only to lay people but also to mental health professionals. The clinical picture of a person who has been reduced to elemental concerns of survival is still frequently mistaken for a portrait of the survivor's underlying character.

Another important concern relevant to offenders as well as survivors of interpersonal trauma is an avoidance of treatment (van der Kolk et al, 1996). Indeed, seeking assistance with depression, anger, self-destructive behaviours and other mental health difficulties may be similar treatment goals. An equally important consideration might be how different agencies conceptualise an individual's difficulties as trauma

symptoms as opposed to simply criminal behaviours (Welfare & Hollin, 2012). However, for survivors, whose experiences of trauma may be central to their offending developing interventions that are ‘‘trauma informed’ and ‘trauma responsive’ could help to ensure a full account of the relevant psychopathology and offending behaviour (Harris & Fallot, 2001).

1.5.2. The evidence base for interpersonal trauma interventions in forensic populations

The limited research base for psychoeducational group based interventions seems to be particularly apparent for incarcerated offenders. As with non-forensic populations there is a dearth of quality controlled studies and those that do exist have often focused on female prisoners in less restrictive regimes (Cole, Sarlund-Heinrich & Browne, 2007; Messina, Grella, Cartier & Torres, 2010). An initial review of the literature indicated five randomised control prison based studies which investigated the efficacy of trauma interventions with female prisoners. These are described below and summarised in Table 1.

The first controlled trial to examine the effectiveness of group therapy for prisoners who had experienced interpersonal violence was Bradley & Follingstad (2003). The 18 session treatment protocol focused the first 9 sessions on psychoeducation and emotions management skills based on the well-validated dialectic behavioural therapy (DBT) model by Linehan (1993). It also included a substantial narrative TMP component in which participants were asked to write about and ‘make sense’ of their lives. This study reported significant reductions and large within subject effect sizes for PTSD, depression and interpersonal symptoms in the treatment arm (n=13) compared to the ‘no contact control’ arm (n=18). Although n=24 participants were assigned to the treatment arm considerable attrition, due to release and other scheduling difficulties, resulted in just over half completing the treatment. The results reported by Bradley & Follingstad (2003) are therefore based only on this very limited subset of completers. Cole et al (2007) also implemented an RCT with very small treatment completion numbers (n=9). The authors reported that this pilot study, which included recently incarcerated female offenders, investigated the efficacy of TMP treatment protocols which the authors describe as consisting of CBT and an overall ‘feminist’

approach. The four phases to their intervention are described as: 1) self-soothing and safety (7 sessions), 2) psychoeducation about the impact of trauma (4 sessions), 3) processing through writing 'personal stories of trauma' (1 session), 4) further processing and termination (4 sessions). Cole et al (2007) also described difficulties in implementing the intervention in a secure environment and the 'mixed results' obtained. As such whilst some non-trauma specific symptoms showed improvement for the treatment arm there were no differences between the arms across any scores that could be reliably attributed to the intervention.

Messina et al (2010) investigated the efficacy of a psychoeducational trauma informed intervention (*Beyond Trauma*; 11 sessions) delivered together with a 17 session gender responsive treatment programme (Covington, 2008) as compared to usual care, a 6 month Therapeutic Community (TC), for substance misusing female offenders (n=115). Statistically significant reductions for the treatment arm were apparent for substance misuse as well as longer retention in aftercare and less re-incarceration. However, both arms in the study reported significant improvements in psychological wellbeing. There are important limitations to this study including using the Addiction Severity Index (ASI) as a main outcome measure for trauma and mental health symptomatology (Melberg, 2004). In addition, substantial demographic differences between the arms were noted and the treatment milieu, in which the trauma informed intervention is situated, makes it difficult to clearly establish whether this was an important active therapeutic constituent.

Zlotnick, Johnson & Najavtis (2009) recruited participants (n=49) from an intensive substance abuse treatment program located in a minimum security women's prison and compared the enhanced psychoeducation / dual diagnosis intervention, *Seeking Safety,* to usual care. The *Seeking Safety* participants attended an average of 15 group sessions and 3 individual booster sessions in addition to usual care. Given that usual care involved intensive programming for approximately 30 hrs per week for 3–6 months, methodological issues concerning the apparent sophistication of usual care, as seen in Messina et al's (2010), need to be considered. Participants in both arms improved significantly on assessed outcomes for PTSD, SUD, psychopathology, and legal problems. However, the study was underpowered and there was potential for contamination between the two conditions as the same clinicians provided both

treatments. The authors also reflected on the potential of participants who lived in a communal setting to share information and materials.

Both the interventions used by Messina et al (2010) and Zlotnick et al (2009) sought to assist women with co-occurring substance use disorder (SUD) and PTSD. Substance misuse is regarded as an important treatment need in offender populations (Fazel, Yoon, Hayes, 2017; Wolff, Huening, Shi, Frueh, Hoover, McHugo, 2015). Both Messina et al (2010) and Zlotnick et al (2009) used the ASI which is an internationally validated semi-structured interview schedule capable of detecting changes during treatment (Denis, Cacciola, & Alterman, 2013; McLellan, Kushner, Metzger, et al. 1992). However, it is noted, and indeed acknowledged by the authors of these previous studies, that the lack of immediate post treatment information using the ASI makes it difficult to ascertain the effectiveness of these interventions.

Ford, Chang, Levine and Zhang (2013) completed the most recent RCT into the effectiveness of a psychoeducational intervention for incarcerated women (n=72). This involved comparing the 12 session *Trauma Affect Regulation: Guide for Education and Therapy* (TARGET) with a supportive group therapy that provided a matched treatment arm. Whilst the supportive group therapy provided psychoeducation on symptoms of traumatic stress, personal boundaries and attachment styles, it did not provide the detailed material on the neurological-behavioural impact of stress and emotional regulation skills that were hypothesised to be essential to TARGET's previous successful outcomes (Ford et al, 2013). Both arms/interventions returned significant reductions in PTSD symptoms and increased self-efficacy with small effect sizes in favour of TARGET (*d*= .13 and .39). Ford et al (2013) focused their evaluation primarily on the alleviation of PTSD symptom reduction and negative mood regulation rather than on behavioural stabilisation which may be more appropriate given the assumed limitations of any brief group therapy. In fact, it is noted that there was a slight increase in negative mood (*d*= .32) in the experimental group highlighting the challenges in working with this group. The only superior outcome that TARGET demonstrated was an increase in forgiveness which can be theorised to have important links to emotional resolution (Reed & Enright, 2006).

Ford et al (2013) has conducted the most robust of the RCT studies on psychoeducational approaches to date with incarcerated females, particularly as their dropout rate was relatively low (<5% compared to a reported average of 33%). However, Ford et al (2013) also has important limitations: a lack of follow up assessments, lack of an intent-to-treat (ITT) analysis, being statistically under powered and a lack of information on participants' sentencing and offending histories. It was also noted that when testing for forgiveness, the only statistically significant outcome, due to an 'administration error' not all participants were included further questioning the power involved in this study and its overall integrity. The authors speculated that a potential drawback of TARGET is that unlike the control, it did not encourage participants to apply emotion management skills to real life issues.

To date the five studies described above (see also Table 1) which have employed an RCT design with incarcerated offenders have not shown any robust statistically significant positive results when compared to control conditions. Only two of these studies have investigated psychoeducational interventions for complex trauma (Messina et al, 2010; Ford et al, 2013). Both of these studies have clearly identifiable methodological issues and only Ford et al (2013) reported effect sizes from which to gauge the magnitude of effect for such an intervention. These methodological deficits also include the lack of an ITT analysis. Mohr, Hopewell, Schulz, Montri, Gotzsche, Devereaux, Elbourne, Egger & Altman (2010) describe such analyses as crucial for allowing the full strength of an RCT to be realised. As such 'preserving' the benefits of randomisation requires retaining all participants into the arms that they were allocated thereby including all available data. The Cochrane review guidelines also have long regarded this as the least bias way for an RCT to estimate an intervention's effectiveness (Higgins & Green, 2011).

As Zlotnick et al (2009) note control trials with disadvantaged women experiencing chronic disorders and co-occurring life problems out with prisons have been shown to be effective. In this respect, a question that is invariably posed is whether there are unique considerations that need to be accounted for when delivering such treatments, and indeed RCTs, in prisons? Such considerations may include unique ethical challenges, scant research expertise, poorly understood and unique mental health needs as well as the security and administration difficulties relevant to prisons.

All of which can add to the considerable length of time it takes to undertaking a RCT (Cislo & Trestman, 2013). There is clearly a need for further work in this area to establish the efficacy for interventions that have been delivered to socially disadvantaged and multiply traumatised populations. Many of these interventions have been based on brief group treatment modalities yet to date there are no systematic reviews synthesizing the outcomes of studies investigating the efficacy of group based interventions for complex trauma (Pelekis & Dahl, 2005; Sloan et al, 2013).

Table 1: Summary of previous RCTs with incarcerated populations

Study Country	**N final sample /** Treatment vs control arm	**Security level** Treatment focus	**Trauma measure**	**Other measures**	**Main findings (study arm x time)** Summary of study limitations
Bradley & Follingstad (2003) USA	N= 13, DBT and narrative TMP (18 sessions) N=18, WL ('no contact')	Medium security IPV	TSI	BDI, IIP	Significant reductions in PTSD, mood and interpersonal symptoms in treatment group. Pre-post only, underpowered, no ITT analysis, high dropout rate, limited analysis and trial reporting.
Cole et al (2007) USA	N= 4, 'feminist' TMP (16 sessions) N= 5, TAU	Medium security (assumed) CSA	TC/R, TSI	SCL-90-R	No statistical differences between arms. Pre-post only, no ITT analysis, limited statistical analysis, substantively underpowered, potential reporting bias (treatment facilitated by author), and treatment protocols ambiguous.
Zlotnick, et al (2009) USA	N= 27, SS (15 sessions – on average) N= 22, TAU	Minimum security Substance misuse and PTSD	CAPS, TSC	SCID, ASI, BSI	No significant difference between conditions both arms significantly improved. SS significant low return rate to prison at 6 month FU. Potential contamination between arms, SS case management option not implemented, not assessed at end of treatment but 4-6 wks after end, limited power, no ITT analysis, inadequate power
Messina et al (2010) USA	N= 60, GRT- incudes *Beyond Trauma* (28 sessions) N= 55, TAU	Minimum security Substance misuse and IPV: prison based therapeutic community	-	ASI (subscales)	No differences in psychopathology reduction – both arms improved. GRT arm greater reductions in drug use and remained longer in further treatment. ASI subscales used to measure psychopathology, lack of immediate post treatment data (6 and 12 months FU), lack of information regarding analysis, missing data and no ITT, lack of trauma outcome measure, substantive demographic differences between arms.

Table 1 continued

Study Country	**N final sample /** Treatment vs control arm	**Security level** Treatment focus	**Trauma measure**	**Other measures**	**Main findings (study arm x time)** Summary of study limitations
Ford et al (2013) USA	N=38,TARGET (12 sessions) N= 34, SGT	Medium security (assumed) PTSD	CAPS, TSI	CORE-OM, HFS, HS, NMR	No differences - both arms had significant reductions. TARGET achieved greater forgiveness (HFS). Inconsistent administration of measures (e.g. HFG), underpowered, active control arm, no ITT analysis, no FU assessment, and partial as well as full PTSD included which may reduce statistical power.

Note: ITT= Intent-to-treat analysis, PTSD= post-traumatic stress disorder, IPV= interpersonal violence, CSA= childhood sexual abuse, FU= follow up assessment, SS= Seeking Safety (intervention), TARGET= Trauma Affect Regulation Guide for Education and Therapy (intervention).
Measures referred to: BDI= Becks Depression Inventory, TSI, Trauma Symptom Inventory, IPP = Inventory of Interpersonal Problems. TC/R = Trauma Content of the Rorschach Inkblot Method. SCL-90-R= Symptom Checklist-90-Revised. CAPS Clinician-Administered Posttraumatic Stress Disorder Scale-I. ASI=Addiction Severity Index. GRT= Gender Responsive Treatment, HFS= Heartland Forgiveness Scale, HS Hope Scale, NMRS = Negative Mood Regulation Scale, SS= Seeking Safety. Other study details are available in Table 2 on page 34.

1.6. Developing the evidence base for psychoeducational interventions in prison populations

As noted above there is a very limited evidence base for the efficacy of group based psychoeducational interventions for interpersonal trauma in prisons. Existing studies often have quasi-experimental designs and a limited number of trauma specific outcomes (Ball, Karatzias, Mahoney, Ferguson, Pate, 2013; Lynch et al, 2012; Zolntick, Najavits, Rohsehow, & Johnson, 2003). Similarly, to date no prison based studies have been able to achieve an adequately powered sample size making it difficult to ascertain whether the positive findings in community populations are applicable to this specialist setting (Benish, Imel & Wampold, 2008; Ford et al, 2013 (Santa Ana, Martino, Ball, Nich, Frankforter & Carroll, 2008). In addition, previous research has often been carried out in conjunction with other treatment protocols making it difficult to isolate the impact of trauma based psychoeducational components.

Whilst RCTs represent a gold standard, essential for developing an empirical understanding of treatment efficacy, it is clear that implementing this methodology in prison settings is challenging (Liebman et al 2013; Lynch et al, 2012). There is, however, precedence for successfully implementing large scale RCT investigations out with the trauma literature. Drawing on such experiences may help manage environmental challenges whilst also more accurately accounting for differences between study participants and non-completers. For example, McDougall et al (2009) in their RCT investigating the efficacy of *Enhanced Thinking Skills*, a problem solving skills group intervention, utilised a third non-randomised arm to provide an ethically and risk appropriate response to including prisoners who were being released sooner from custody and who were assessed as being at a high risk of harm.

Given the paucity of evidence based treatments currently delivered in prisons management processes have advocated the use of interventions designed on the best available and theoretically sound evidence. However, such situations can unintentionally result in interventions being delivered that may actually cause harm (Linden, 2013; Mews, Di Bella & Purver, 2017). These management based processes have also required that non-randomised cohorts be used in RCT to address concerns about withholding treatment. Importantly, as with McDougall et al (2009), those randomly allocated in the trial need to be of a large enough sample size to ensure

sufficient statistical power from which to detect differences in the outcome measures used and to avoid selection bias.

The appropriateness of RCTs in investigating the invariably long term recovery process undertaken by survivors of interpersonal trauma has also been argued (Corrigan & Hull, 2015). Nevertheless, such methodologies are essential if the positive claims regarding the efficacy of large scale psychoeducational interventions are to be investigated appropriately (Ali, Rhodes, Moreea, McMillan, Gilbody, Leach, Lucock, Lutz, & Delgadillo, 2017; Brookes, Barrett, Netten & Knapp, 2013; Tucker & Oei, 2007). Random allocation is therefore central to the ability of a study to make causal inferences about the efficacy of an intervention. In this respect endeavouring to account for all relevant known and unknown variables that may influence outcomes is important if the efficacy of a treatment is to be reliably and unequivocally demonstrated (Freedland, Mohr, Davidson, & Joseph, Schwartz, 2011; Lilienfeld, McKay, Hollon, 2018).

There are very few pure psychoeducational interventions for complex trauma. One of these interventions is *Survive & Thrive*, a brief 10 session psychoeducational programme developed in Scotland. Survive & Thrive has been adopted by the Scottish Prison Service (SPS) for delivery in the female prison estate to help respond to a clear treatment need within this population. This was a reasonable decision based on the wide spread delivery of Survive & Thrive in community settings throughout Scotland and the drive to provide equivalence of care to survivors in prison (NES, 2018; Scottish Government 2017; Tully, Forrester & Exworth, 2014). There is however only a very limited evidence base for Survive & Thrive's efficacy and to date there are only two uncontrolled studies on community based populations. For example, Ball, Karatzias, Mahoney, Ferguson & Pate (2013) in their preliminary evaluation of an earlier version of Survive & Thrive (Ferguson, 2008) reported positive results with a small sample of community based female offenders. This included statistically significant differences between pre and post-treatment scores across all dimensions of the PCL-C and CORE outcome measures with the exception of the CORE 'Risk' subscale. These results indicate that Survive & Thrive may have led to clinical improvements with regard to general psychological distress and other trauma symptomatology.

It is, however, notable that Karatzias, Ferguson, Chouliara, Gullone, Cosgrove, Douglas (2014) reported less favourable outcomes in their pilot evaluation of Survive & Thrive. In this non-forensic mixed gender study Survive & Thrive did not have any impact on clinical improvement with regard to general psychological distress, trauma symptomatology, self-esteem, and life satisfaction. There were however some mixed results with respect to self-harming, smoking and alcohol use. This included statistically significantly reduced numbers of participants reporting self-harm $(X^2)_{(1)}$ -21.0, $p \leq .001$), smoking $X^2_{(1)}$ -13.0, $p \leq .001$ and alcohol consumption $X^2_{(1)}$ -13.0, $p \leq .001$. However, the frequency of these behaviours was not reduced (cigarettes per day $t(4)= -1.0$, n.s; alcohol units per week: $t(8)= -1.3$, n.s). It was also noted that self-harm frequency (i.e. at least once per week: n=2, at pre and post assessment) was already low (out of the n=37 who started treatment and the n=21 who completed post assessment) suggesting that this was not a population with chronic levels of behavioural dysregulation. As such it could be hypothesised that the differences in the outcomes between these two repeat measures studies are because Survive & Thrive is more effective in populations, such as with a forensic sample, where there are higher levels of behavioural and emotional disturbance.

These preliminary studies suggest that the efficacy of Survive & Thrive requires further investigation particularly given the widespread delivery of this intervention across Scottish health boards without adequate empirical testing. Similarly, in respect to managing interpersonal trauma related psychological distress in a prison setting employing robust RCT protocols was required in order to ensure its suitability to a custodial setting. It is hypothesised that the stability provided by a prison or inpatient setting may help to manage previously reported high dropout rates although this has yet to be established (Ball et al, 2013; Lewis, 2006; Liebman, Burnette, Raimondi, Nichols-Hadeed, Merle & Cerulli, 2013).

1.6.1. Survive & Thrive: a psychoeducational intervention for complex trauma

As noted above, Survive & Thrive (Ferguson, 2013), is a psychoeducational group based intervention for the stabilisation of complex interpersonal trauma and has been widely delivered across mental health settings in Scotland. Survive & Thrive's initial usefulness was in response to large numbers of survivors on waiting lists for mental

health services. However, Survive & Thrive has been increasingly considered as a useful phase 1 intervention for promoting emotional and behavioural stabilisation and for helping survivors to understand the link between their past trauma and current symptoms (UKPTS, 2017). As a brief 10 session intervention Survive & Thrive is amongst a limited number of psychoeducational interventions for interpersonal trauma, none of which have an effective evidence base in prisons (see Table 1). However, Survive & Thrive's additional appeal is that it can be delivered by staff trained to deliver low intensity interventions (NES, 2018) thereby potentially making it a pragmatic and cost effective intervention.

1.6.2. Aims and objectives of the thesis

The primary objective for this thesis is to answer the following question: *Are group based psychoeducational interventions efficacious for the stabilisation of trauma symptomatology in a prison setting for female offenders?* However, as noted earlier, the evidence base for the efficacy of group psychoeducational interventions in prison settings is extremely limited. Therefore, it is imperative to contribute to the evidence base though establishing the efficacy of Survive & Thrive (S&T) as a prototypical psychoeducational intervention that has been widely delivered and piloted successfully in community treatment settings (Ball et al, 2013; Karatzias, Ferguson, Chouliara, Gullone, Cosgrove, Douglas, 2014).

As such the thesis will focus on whether S&T is efficacious when delivered in a female prison population. This includes establishing whether S&T helps to 'stabilise' (i.e. reduce) symptoms associated with trauma based disorders, including maladaptive behaviours, as important main outcomes as well as other outcomes commonly associated with mental health difficulties.

In light of the limited evidence base as outlined above this thesis will seek to establish the efficacy of group based interventions more generally through a meta-analytic review across a range of clinical populations.

The relevant research questions therefore are:

1. Will S&T be an efficacious intervention for promoting behavioural and emotional stability as associated with survivors of interpersonal trauma compared to a wait list control group (i.e. usual care) in a prison setting?
2. Will S&T be an efficacious intervention for stabilising symptoms associated with PTSD and compared to a wait list control group in a prison setting?
3. Will S&T be an efficacious treatment for stabilising general symptoms of psychopathology compared to a wait list control group in a prison setting?
4. Will S&T be a more efficacious treatment for those participants who receive an 'adequate dose' compared to a waitlist control group in a prison setting?
5. What is the current efficacy of psychoeducational group based interventions in the stabilisation of trauma associated symptomatology in comparison to various control conditions (i.e. usual care as well as trauma and non-trauma interventions)?

The final question listed above necessitates a meta-analytic review; the aims of which are further discussed in the next chapter.

2. A systematic review and meta-analysis of group treatments for adults with symptoms associated with complex post-traumatic stress disorder

2.1. Introduction

2.1.1. Complex post-traumatic stress disorder and phase based treatments

Across various clinical populations, histories of interpersonal violence and its negative psychological sequelae have long been recognised as having a profound impact on survivor's lives (Loewenstein & Brand, 2014, Herman, 1992, van der Kolk & van der Hart, 1989). The recently published ICD-11 (WHO, 2018) has formally recognised 'complex post-traumatic stress disorder' (CPTSD) as a disorder that can arise from chronic and often inescapable interpersonal violence (Cloitre, Garvert, Brewin, Bryant, & Maercker, 2013; Karatzias, Shevlin, Fyvie, Hyland, Efthymiadou, Wilson, Roberts, Bisson, Brewin, Cloitre, 2016). As previously discussed in the Introduction chapter the evidence suggests CPTSD may involve a distinct symptom profile, including high levels of depression, psychological distress, dissociation and substance misuse (Brewin, Cloitre, Hyland, Shevlin, et al, 2017; Loewenstein & Brand, 2014; Mauritz, Goossens, Draijer & van Achterberg, 2013). It is therefore important to clearly evaluate the efficacy of interventions that have been offered to clinical populations where there may have been a high prevalence of CPTSD symptoms (Dorrepaal, Thomaes, Hoogendoorn, Veltman, Draijer, & van Balkom, 2014). One such example of this is in prisons where a recent study indicated that as many as 58% of women prisoners met the diagnostic criteria for PTSD and 91% reported experiences of interpersonal trauma over their lifetime (Karatzias, Power, Woolston, Apurva, Begley, Quinn, Jowett, Howard & Purdie, 2017).

As previously noted a number of authors have advocated that efforts should be made to avoid symptom exacerbation through trauma memory exposure. The concern being that without the appropriate intra and interpersonal resources exposure to the traumatic memory could impair engagement and negatively impact on the efficacy of trauma-focused treatment (Cloitre et al, 2012; de Jongh, Resick, Zoellner et al, 2016). Psychoeducational interventions have therefore been offered instead at the beginning of therapy (i.e. phase 1) and often focus on safety planning, coping, anxiety

management or interpersonal difficulties (Dorrepaal et al., 2010; Zlotnick et al., 1997; Krupnick et al., 2008). In general, group based stabilisation interventions have tended to be brief and psychoeducational in their approach (Pelekis & Dahl, 2005). Indeed, such interventions have tended to be much briefer than the 6 month generally regarded as reasonable for this phase (Cloitre, Courtois, Ford, Green, Alexander, Brier et al, 2012).

However, the evidence for a special stabilization phase is weak (de Jongh et al, 2016). Therefore, there has been some scepticism as to whether phase 1 interventions achieve greater levels of symptom and behavioural stabilisation as opposed to phase 2 interventions that are more orientated towards trauma memory processing (TMP). Despite this, recent head-to-head trials have also questioned whether TMP treatments are necessarily more efficacious than phase 1 or 'non-trauma focused' interventions (Foa, McLean, Zang et al, 2018). As such, questions still exist as to whether a phased based approach or a general compassionate and therapeutic response might help survivors make more substantive progress in addressing symptoms and disorders resulting from interpersonal violence (Hoge & Chard, 2018).

2.1.2. Group versus individual treatment modalities

There is also considerable ambivalence and indeed disagreement about the benefits and treatment efficacy that might be derived from group based interventions for complex trauma. Several meta-analyses have reported that the largest reductions in PTSD symptoms can be achieved through individual trauma-focused treatments (Ehring et al, 2014; Taylor & Harvey, 2010; Watts et al, 2013). Historically, those advocating for the benefits of group based treatments have relied on clinical experience and theory (Fritch & Lynch, 2008; Herman, 1992, p. 214). It is thought that group approaches help to normalise symptoms, counteract isolation, provide peer support and observational learning, and ameliorate important shame based cognitions (Burlingame, Fuhriman & Mosier, 2003; Dorrepaal et al., 2010; Herman, 1992; Mendelsohn, Herman, Schatzow, Coco, Kallivayalil, & Levitan, 2011; Mendelsohn, Zachary, and Harney, 2007; McCrone, Weeramanthri, Knapp, Rushton, Trowell, Miles & Kolvin, 2005; Shea, McDevitt-Murphy, Ready, Schnurr, 2009; Zlotnick et al., 1997).

Short-term group psychotherapy has been a major treatment modality offered to people suffering from the psychopathology associated with complex interpersonal trauma such as child sexual abuse (Pelekis & Dahl, 2005). The potential of group based trauma-focused treatments to be an effective response to potentially large populations of survivors is an important consideration (Wolff, Huening, Shi, Frueh, Hoover & McHugo, 2015). However, along with these potential benefits come the challenges of implementing processes that maintain treatment replicability and fidelity (Najavitis & Hien, 2013). The aim of this review is therefore to produce a synthesis of the current evidence relating to the efficacy of group interventions, as a distinct treatment modality, for survivors of interpersonal trauma. Synthesising treatment outcomes according to a phase based approach may also help to develop a more nuanced understand of this modality's effectiveness across a range of symptoms.

2.1.3. Previous meta-analyses

To date a number of meta-analyses and systematic reviews have investigated the efficacy of PTSD treatments in general (Barrera et al, 2013; Bisson & Andrews, 2005, 2007; Bisson, Ehlers, Matthews, Pilling, Richards & Turner, 2007; Bisson, Roberts, Andrew, Cooper, & Lewis, 2013; Callahan et al 2004; de Jong & Gorey, 1996; Ehring et al, 2014; Pelekis & Dahl, 2005; Lenz, Haktanir & Callender, 2016; Roberts et al, 2015; Sloan et al, 2013; Taylor & Harvey, 2009; Taylor & Harvey 2010; Watts et al, 2013). In Bisson et al (2013) extensive review of psychological therapies for 'chronic' PTSD, 70 RCT studies were identified; this included 10 group based studies of which only one study was categorised as having a group non-Trauma Focused Cognitive Behavioural Therapy (non-TFCBT) arm. Bisson et al (2013) concluded that group TFCBT was superior to waitlist/usual care control conditions but that this was not the case for group non-TFCBT. Other meta-analyses have also highlighted that survivors with CPTSD symptoms, may present specific challenges to PTSD treatments (Dorrepaal et al, 2014; Greger, Munder & Bath, 2014), however, Torchalla, Nosen, Rostam & Allen (2012) also demonstrated that individuals with concurrent substance misuse disorder and PTSD responded equally well to both integrated and non-integrated treatments.

Sloan et al (2013) and Barrera et al (2013) are currently the only meta-analytic reviews that have focused exclusively on the efficacy of group treatments for PTSD. However, Barrera et al (2013) was specifically limited to CBT group treatments (n= 12). Given the preponderance of CBT studies within the PTSD treatment literature, there are of course similarities between this review and Sloan et al (2013) who identified 16 studies. Both reviews concluded that group treatments lead to large and significant pre-post treatment reduction in PTSD symptoms. However, Sloan et al (2013) concluded that there was no relative superiority for group treatments when compared to active treatment controls (*d*= .09, 95% CI [-.03, .22]). Nevertheless, group treatments were better than waiting list (WL) control comparisons (d= .56, 95% CI [.31, .82]). Barrera et al (2013) did not undertake an analysis according to the type of control used and reported that there were no significant differences in effect sizes between group treatments that included both in-group exposure and those that did not. Recent, meta-analyses have computed large effect sizes when individual trauma-focused (i.e. TMP) treatments are compared against minimal or no treatment arms. However, small effect sizes have been obtained when compared to other, non-trauma-focused active interventions, which has led to the efficacy of TMP treatments being questioned (Erford, Gunther, Duncan, Bardhoshi, Dummett, Kraft, Deferio, Falco & Ross,2016; Lenz, Haktanir & Callender, 2017). Such comparisons have never been made in group therapies.

Although there is considerable evidence for the treatment of PTSD there has been no meta-analysis of the efficacy of the group based interventions for complex interpersonal trauma symptoms in the outcome domains of PTSD, Depression, Psychological Distress, Substance Misuse and Dissociation. Symptoms associated with these conditions are commonly reported in people with interpersonal trauma. Furthermore, no previous meta-analyses of interventions for complex interpersonal trauma have considered whether phase 1 interventions (i.e. psychoeducational approaches), as characterised by high levels of psychoeducation and stabilisation, are more effective than phase 2 approaches, which include TMP protocols.

2.1.4. Research questions

The aim of this current systematic review and meta-analysis is to investigate previously unaddressed questions in the developing literature for group based treatments for populations with complex traumatisation. The following questions were considered across a range of common outcomes, including PTSD, depression, psychological distress, substance misuse and dissociation.

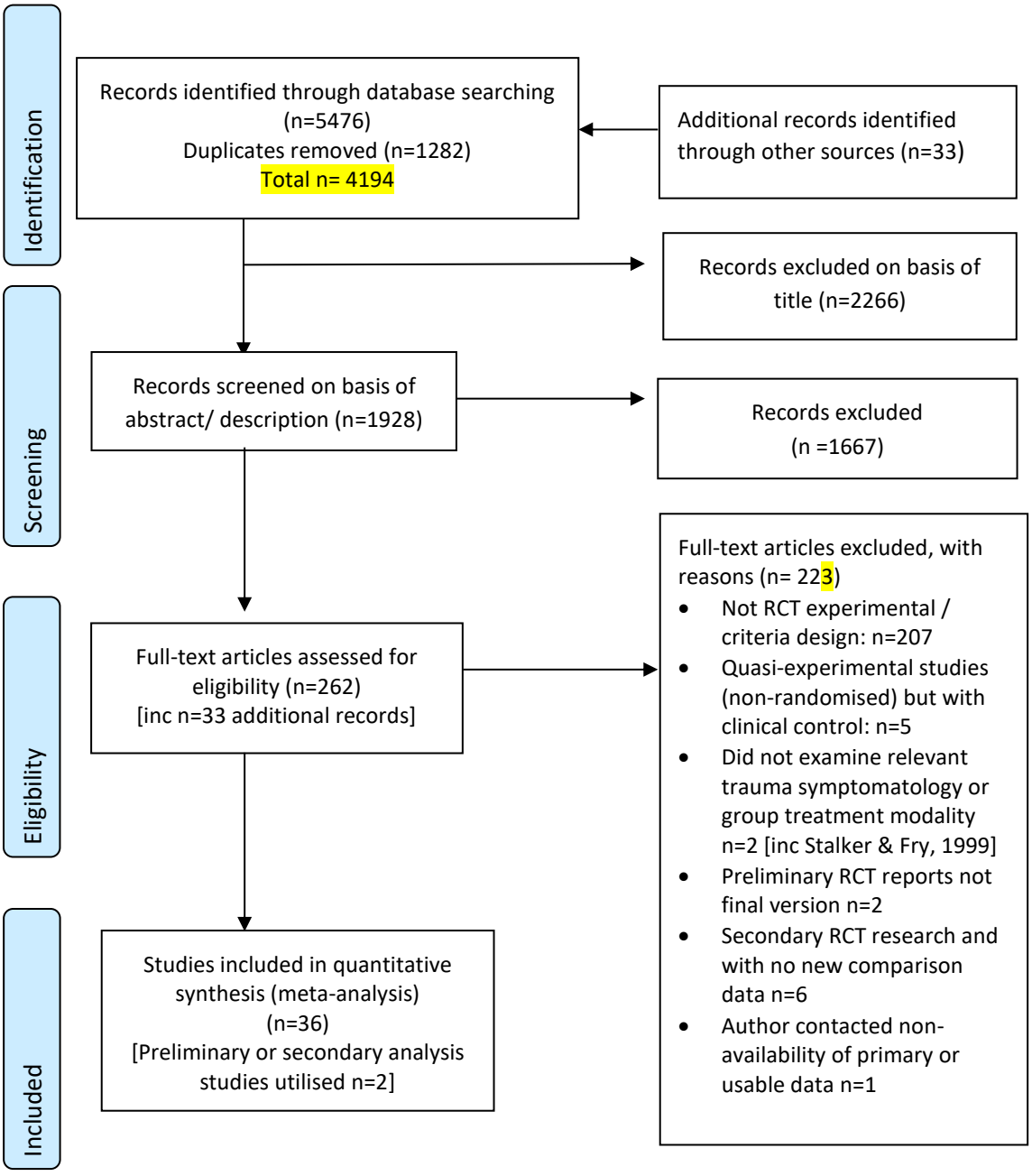
1. Are group based trauma interventions more effective than usual care?
2. Are group based trauma interventions more effective than other non-trauma group based treatments?
3. Are psychoeducational treatments more effective than usual care?
4. Are psychoeducational groups of greater intensity more effective than usual care?
5. Are TMP group treatments more effective than usual care?
6. Are TMP group treatments more effective than psychoeducational group treatments?

2.2. Method

2.2.1. Search and inclusion criteria

Search protocols were constructed with support from a senior healthcare research librarian and are detailed in Appendix 1.1. Inclusion criteria focused on identifying randomised clinical trials (RCT) of psychological interventions for people with histories of interpersonal trauma. As CPTSD has not until recently been diagnostically recognised, complexity was inferred from the interpersonal violence participants had experienced (e.g. childhood sexual abuse, intimate partner violence, genocide or war based experiences). All studies were based on interventions seeking to ameliorate at least one of the symptom domains included in this review and included those diagnosed with PTSD as well as those reporting traumatic stress from their experiences. The screening of articles not considered relevant to the literature review and the selection process is shown in Figure 2. Studies that focused exclusively on veteran populations were excluded to ensure that the primary focus was on participants with histories of interpersonal trauma and abuse.

A comprehensive search of relevant bibliographic databases included: Cochrane Database of Systematic Reviews, EMBASE, MedLine, PsychINFO, Social Services Abstracts, Sociological Abstracts, Web of Knowledge (including Science Citation Index and Social Science Citation Index), World Health Organisation ICTRP, CINAHL and Pubmed. This search process was undertaken in December 2016 and included all relevant available studies up until that date. The reference lists of earlier meta-analyses and systematic reviews were also screened for additional studies (Barrera et al, 2013; Bisson et al., 2007; Bradley et al., 2005; Taylor & Harvey, 2010; Sloan et al, 2013; Ehring et al 2014). As illustrated in Figure 2, n= 4194 studies were identified and screened, and 36 studies were included.



Notes: Adapted from: Moher, Liberati, Tetzlaff, Altman, The PRISMA Group (2009).

Figure 2: PRISMA flow diagram

2.2.2. Data extraction and categorisation

Interventions were classified as either *Psychoeducation, Psychoeducation Plus* or *TMP* treatments. In this review the former referred to phase 1 treatments with a defining focus on symptom stabilisation, safety and treatment relevant information and

included interventions that were either interactive or skills based (Lubin, Loris, Burt & Johnson, 1998; UKPT, 2017; Wessely, Bryant, Greenberg, Earnshaw, Sharpley & Hughes, 2008). The category *Psychoeducational Plus* differentiated more specific and substantive phase 1 treatments. Interventions in this category were defined as seeking to achieve greater treatment responsivity by focusing on specific symptoms and comorbidities. For example, *Seeking Safety*, designed to be a comprehensive intervention to treat comorbid substance use disorder (SUD) and PTSD would be included within this category (Najavtis, 2002). However, briefer versions of this intervention that just focused on PTSD psychoeducation would not (Ghee, Bolling & Johnson, 2009). TMP interventions were defined as 'trauma focused' interventions that assisted survivors through the exposure (imaginal or in vivo) to traumatic memories as well as cognitive restructuring through discussing traumatic memories and their associated faulty appraisal (Lenz et al, 2017).

Each study was also categorised according to the type of comparator used. Control arms involving waiting list (WL), minimal attention control or treatment as usual (TAU) were all categorised as 'usual care'. The description by Devilly & McFarlane (2009, pg. 1162) was utilised for these arms as there was an assumption that participants had received 'routine care, whether this was specifically mentioned in the original article or not, as long as this did not include active, trauma-focused treatment'. Control interventions involving a degree of psychotherapeutic sophistication were classified as 'active'. Therefore, complementary therapies such as acupuncture were not considered active treatments (Hollifield, Sinclair-Lian, Warner & Hammerschlag, 2007). However, therapies that assisted participants in developing somatic regulation skills, such as trauma informed yoga and introsceptive awareness or mindfulness, were considered active psychological therapies and categorised as *Psychoeducational* interventions (Kelly & Garland, 2016; Garland et al, 2016; Mitchell et al, 2014; van der Kolk, 2014).

Trials which had two or more group based treatment arms were combined following the Cochrane Handbook procedures (Higgins & Green, 2011, 7.7.3.8; 16.5.4). This approach was used when conducting meta-analyses comparing all group based treatments to usual care comparators. Where *TMP* and *Psychoeducational* arms had been combined, studies were categorised as *Psychoeducation Plus*. This included

Alexander et al, 1989; Classen et al (2001); Classen et al (2011) and Yeomans et al (2010) and the combination was used in the overall analyses of group-based interventions compared to usual care as well as the subgroup analysis of different treatment types compared to usual care. In these circumstances, *Psychoeducation Plus* refers to both single arm studies within this category and studies with the combined arms. Where analyses were conducted between different trauma-focused treatment arms, *TMP* interventions were considered the treatment group and compared to *Psychoeducation* and *Psychoeducation Plus* interventions. See Appendix 1.2. for decisions on categorisation.

2.2.3. Risk of bias and coding of methodological quality

The Cochrane Risk of Bias tool (Higgins & Green, 2011, version 5.1.) was used to assess overall methodological quality for each study. The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system (http://www.gradeworkinggroup.org) was used to rate the quality of outcomes and grade the strength of recommendations made across the various domains of clinical functioning (Thornton, Alderson, Tan, Turner, Latchem, Shaw, et al., 2013). The effect of randomisation and assessor blinding was also examined using moderator analyses.

2.2.4. Outcomes

The five outcome domains that were investigated included: *PTSD* (i.e. overall levels of trauma symptomatology); *Depression*; *Psychological Distress*; *Substance Misuse* and *Dissociation*. Where global measures of psychological distress were not available scores from anxiety and depression measures were combined following procedures detailed in the Cochrane Handbook (Higgins & Green, 2011). Appendix 1.4. (GRADE Assessment of outcome) includes a detailed record of which studies were included within the comparisons undertaken for each domain.

A sequential hierarchy was devised to account for the different measures used by studies to assess the same treatment outcomes or symptoms. For example, the Clinician Administered PTSD Scale (CAPS, Blake, Weathers, Nagy, Kaloupek, Klauminzer, Charney, Keane, & Buckley, 2000) and similarly the Structured Interview for Disorders of Extreme Stress (SIDES, Pelcovitz, van der Kolk, Roth, Mandel, Kaplan,

& Resick, 1997) were prioritised over other measures such as the Davidson Trauma Scale (Davidson, Tharwani & Connor, 2002). Substance misuse measures that were prioritised included biological testing and interview procedures over self-reported reductions of use. Intent-to-treat (ITT) data was also prioritised over completer samples where available.

Primary effect sizes were calculated using data from the first available time point after treatment ended. Whilst this was usually described as 'post' treatment, research protocols between studies varied in terms of delayed data collection and therefore the first available post-treatment data following treatment completion as included.

2.2.5. Meta-analysis

Outcome data for individual trials was entered into the Comprehensive Meta-Analysis (CMA) version 3.3.070, (Borenstein, Hedges, Higgins & Rothstein, 2014) for Windows software. The authors of all included studies were contacted and additional data requested where required.

Hedges's g was used to calculate effect sizes from the reported standardized mean difference (SMD) for continuous data using CMA, together with 95% confidence intervals (CIs). Mean effect sizes were calculated using a random-effects model, since this accounts for the dispersion of effect sizes where studies are unlikely to be functionally equivalent (Borenstein, Hedges, Higgins, & Rothstein 2009; Taylor & Harvey, 2010). Effect sizes calculated using Hedges's g were conservatively interpreted using Cohen's (1988) conventions where 0.2, 0.5 and 0.8 indicated small, medium and large effects respectively. Two-tailed hypotheses were used throughout and statistical significance was assumed if the probability of the observed difference arising under a true null hypothesis was less than 5% ($p<0.05$). Publication bias was investigated for each outcome domain using funnel plots (see Appendix 1.3.) and an outlier was defined as a study with an effect sizes at least one standard deviation beyond other effect-size values in either a positive or negative direction (see Weisz, Weiss, Han, Granger, & Morton, 1995).

2.2.6. Heterogeneity

Higgins's I^2 was used to express the amount of heterogeneity among studies. Moderate heterogeneity was assumed if the I^2 statistic was 40% or above (Higgins & Green, 2011, 9.5.2).

In order to manage heterogeneity within the included studies subgroup procedures available in CMA were used (Cuijpers, 2016). Two subgroup analyses were conducted for each outcome domain, corresponding to the research questions above. The first subgroup analysed studies according to comparator (see Figure 3) and the second according to treatment type.

Additional analyses, both statistical and visual plot, investigated potential bias in the study synthesis. This included calculating the fail-safe N, the number of studies required to support the null hypothesis and to reduce an effect size to a specified level (Orwin, 1983). The trim and fill method was also used to investigate whether 'trimming' potentially 'biased' studies would change the effect size.

2.2.7. Moderator analysis

Meta-regression procedures available in CMA V3 were utilised using a random-effects procedure to examine the potential moderating influence of study and treatment characteristics on treatment effect size estimates. Choice of moderator variables was informed predominantly a priori from previous meta-analytic reviews (Taylor & Harvey, 2010) as well as post hoc from other characteristics apparent in the included studies. See Appendix 1.5. for a full list of variables. A priori variables included publication details, participant characteristics, therapist context and the amount and type of treatment content. Variables were expanded post hoc in regards to treatment content variables and a prisons/forensic variable included in treatment settings as well as a summary risk of bias rating. Regression coefficients were calculated to identify which moderators explained a significant proportion of between study variance (Borenstein et al, 2009; Bowman, 2012). Regression coefficients were the estimated change in g per unit in each predictor variable. The *Q*-statistic was also calculated as an indicator of heterogeneity.

2.3. Results

2.3.1. Study characteristics

The PRISMA diagram details the search outcome (Figure 2). A total of 36 studies were identified and their characteristics are summarised in Table 2. This included 30 control WL/TAU (i.e. 'usual care') and 49 active treatment arms. Six were group based non-trauma active comparators, 15 arms were classified as TMP interventions and 28 were classified as interventions involving psychoeducation and the stabilisation of trauma related symptoms. In addition, one study presented results from psychoeducational and TMP arms combined (Classen et al, 2001). Within the psychoeducational arms, 8 were of sufficient intensity and focus to be classified as *Psychoeducation Plus* interventions. For example, Sikkema et al (2007, 2013) and Classen et al (2001, 2011) focused on addressing HIV risk behaviours whilst providing sufficient focus on treating trauma based symptoms. Non-trauma active comparators were defined as structured or manualised interventions that provided, often psychoeducational, treatment or support on other health or wellbeing issues not related to trauma. Stalker & Fry (1999) was the only RCT identified that compared TMP group treatment against individual based TMP treatment; as such this study was not included.

A distinct group of psychoeducational studies were based on mindfulness and yoga trauma informed therapeutic approaches (Garland et al, 2016; Kelly et al, 2016; Mitchell, Dick, DiMartino, Smith, Niles, Koenen & Street, 2014; van der Kolk, Stone, West, Rhodes, Emerson, Suvak & Spinazzoia, 2014). These therapies focused on affect tolerance and impulse regulation and differ from mainstream cognitive models through promoting somatic regulation and interoceptive awareness. Tables 3 and 4 provide summary data for study and participant characteristics.

Table 2: Summary of study characteristics

Study **Type of group treatment and control condition(s)** *Study relevant details*	**Population, Treatment Goal** *Symptom Domain*	**Treatment Categorisation** (No. of psychoed sessions) [1]	**No. planned sessions**[2] (No. of individual sessions)	***N*** in study *ITT if Inc.*	***N post***[3]	**Sample and Trauma Details** Primary abuse details	% female (% full PTSD)	Age M (SD)
Alexander et al (1989)	Incest survivors: treat CSA			65*		100% CSA	100%	36 (8.4)
Interpersonal transaction	*Dep, PDist*	Psychoed	10		16	(incest)		
Process (peer) group therapy		TMP	10		20			
WL		-	-		21			
** Total n randomised given only*								
Bass et al (2013)	Low income country: conflict					100%	100%	
Cognitive Processing Therapy	trauma	Psychoed Plus (1)	11 (1)	141	114	'sexual		36.9(13.4)
Individual Support (TAU)	*PTSD, PDist*	-	(3)	182	156	violence'		33.8(12.4)
Bohus et al (2013)	Treatment resistant PTSD:			*ITT*		100% CSA	100%	35.1(10.6)
DBT plus Exposure	CSA and co-occurring	TMP (11)	65 (25)	36	29			
TAU ('any treatment of choice')	psychopathology	-		38	29			
	PTSD, Dep, PDist, Diss		-					
Bradley & Follingstad (2003)	Prison: treat PTSD from					100% CSA	100%	36.7(8.3)
Narrative group and DBT skills	interpersonal violence	TMP (9)	18	24	13	and other		
WL ('no contact')	*PTSD, Dep, PDist, Diss*	-		25	18	abuse		
Chard (2005)	Treat CSA			*ITT*		100% CSA	100%	32.8(8.9)
Cognitive Processing Therapy	*PTSD, Dep, PDist*	TMP (1)	17 (10)	36	28			
WL ('minimal attention')		-	-	35	27			
Classen et al (2001)	Treat CSA					100% CSA	100%	38.4(11.7)
Present Focused / Trauma Focused	*PTSD, Dep, PDist, Diss*	Psychoed	24 (1)	19	19			
(results combined)		TMP	24 (1)					
WL				33	33			

Table 2 continued

Study **Type of group treatment and control condition(s)** *Study relevant details*	**Population, Treatment Goal** *Symptom Domain*	**Treatment Categorisation** (No. of psychoed sessions) [1]	**No. planned sessions**[2] (No. of individual sessions)	***N*** in study *ITT if Inc.*	***N post***[3]	**Sample and Trauma Details** Primary abuse details	% female (% full PTSD)	Age M (SD)
Classen et al (2011)	Treat CSA and HIV risk			*ITT*		100% CSA	100%	36.4 (9.7)
Present Focus	behaviours	Psychoed	24	56	43			
Trauma Focus	*PTSD, Dep, PDist, SubM, Diss*	TMP (2+)	24	55	42			
WL		-		55	44			
Cole et al (2007)	Prison: treat CSA						100%	31 (9.8)
Trauma Focused	*PTSD, Dep, PDist, Diss*	TMP (4)	16	4	4	60% CSA*		
WL		-	-	6	5	50% CSA*		
**Surmised from data may be higher*								
Constantino et al (2005)	DV Shelter Residents:					100% IPV	100%	35.5(7.3)
Social Support Intervention	stabilise / alleviate distress	Psychoed	8	13	13			
WL (No treatment)	*PDist*	-	-	11	11			
Crespo & Arinero (2010)	Treat IPV					100% IPV	100%	41(9.3)
Communication Skills	*PTSD, Dep, PDist*	Psychoed	8	28	DK		(0%*)	
Exposure		TMP (7)	8	25				
**Partial PTSD only included*								
Dorrepaal et al (2012)	Stabilise complex PTSD			*ITT*			100%	
CBT	symptoms	Psychoed	20	31	38	97% CSA		40.3 (10.7)
WL-TAU ('individual psychotherapy')	*PTSD*	-		28	33	91% CSA		37.1 (10.3)

Table 2 continued

Study **Type of group treatment and control condition(s)** *Study relevant details*	**Population, Treatment Goal** *Symptom Domain*	**Treatment Categorisation** (No. of psychoed sessions) [1]	**No. planned sessions**[2] (No. of individual sessions)	***N*** in study *ITT if Inc.*	***N post***[3]	**Sample and Trauma Details** Primary abuse details	% female (% full PTSD)	Age M (SD)
Falsetti et al (2008)	'Crime victims':			*ITT*		69% CSA	100%	35 (9.8)
Multiple Channel Exposure	treat panic disorder and	TMP (1)	12	25*	22		(100%)	
WL	PTSD	-		23	31			
**Completers only*	*PTSD, Dep, PDist*							
Ford et al (2013)	Prison: treat PTSD and					60% CSA	100%	
TARGET (Affect Regulation)	stabilise	Psychoed	12	41	38*		(82%)	34.6 (8.6)
Supportive Group Therapy	*PTSD, PDist, Diss*	NTG	12	39	34		(74%)	38.0 (7.8)
** Variations across different measures*								
Frisman et al (2008)	Complex PTSD with			*ITT*				
TARGET (Affect Regulation + TAU)	substance misuse: treat PTSD	Psychoed (8-9)	34.1*	141	91**	DK	63.1%	38.0
TAU (substance abuse care)	and stabilise	-	39*	72	50	(61.9%)	56.9%	
**Mean of total sessions with TAU. Actual TARGET sessions = 3.4*	*PTSD, Dep, PDist, SubM*						(100%)	
***Variations across different measures*								
Garland et al (2016)	Homeless/previously			*ITT*		100%	0%	
Brief trauma informed CBT	incarcerated*: treat	Psychoed	10	64	45	traumatic	(20%)	37.7(10.4)
MORE (Mindfulness)	traumatic distress and	NTG	10	64	48	event;	(27%)	36.5(11.2)
TAU (Therapeutic Community)	substance dependence	-	DK	52	52	81.1%	(29%)	38.7 (9.8)
** Lifetime incarcerated months*	*PTSD, Dep, PDist, SubM*					violence		
M(SD): 40.1 (55.9)								

Table 2 continued

Study **Type of group treatment and control condition(s)** *Study relevant details*	**Population, Treatment Goal** *Symptom Domain*	**Treatment Categorisation** (No. of psychoed sessions) [1]	**No. planned sessions**[2] (No. of individual sessions)	***N*** in study *ITT if Inc.*	***N post***[3]	**Sample and Trauma Details** Primary abuse details	% female (% full PTSD)	Age M (SD)
Ghee et al (2009)	Residential substance abuse					CSA: '*the*		34.7(8.7)
Seeking Safety*	clinic: reduce trauma-related	Psychoed	6	51	36	*majority*'	100%	
TAU: 'standard' addiction	symptoms	-	-	52	52			
treatment	*PTSD, SubM*							
**Condensed version plus TAU*								
Graham-Bermann et al (2013)	Treat IPV traumatic stress					IPV 100%	100%	33.1 (5.3)
Moms Empowerment Program	symptoms	TMP (DK)	10	61	60		(72%)	
Children only	*PTSD*	-	10	62	56		(82%)	
WL (combined)		-	-	58	57		(86%)	
Hien et al (2009)	Treat substance abuse &			*ITT*		CSA: 70.1%	100%	39.2 (SD
Seeking Safety	trauma	Psychoed	12 (1)	103	176		(76.7%)	not
Women's Health Education	*PTSD, SubM*	NTG	12 (1)	96	177		(84.2%)	reported)
Hinton et al (2011)	Latino 'treatment resistant':			*ITT*		DK	100%	
Applied Muscle Relaxation	treat PTSD	Psychoed	14	12	12		(100%)	51.4 (5.9)
Culturally Adapted CBT	*PTSD, PDist*	TMP (DK)	14	12	12		(100%)	47.6 (8.2)
Holllifield et al (2007)	Treat PTSD					33%	47.8%	
CBT	*PTSD, Dep, PDist*	TMP (3)	12	21	24	CSA/CPA	(100%)	40.9(13.4)
Acupuncture		-	12	19	25			42.3(12.1)
WL (combined)		-		21	24			43.4(13.5)

Table 2 continued

Study **Type of group treatment and control condition(s)** *Study relevant details*	**Population, Treatment Goal** *Symptom Domain*	**Treatment Categorisation** (No. of psychoed sessions) [1]	**No. planned sessions**[2] (No. of individual sessions)	***N*** in study *ITT if Inc.*	***N post***[3]	**Sample and Trauma Details** Primary abuse details	% female (% full PTSD)	Age M (SD)
Kaslow et al (2010)	Suicidal African Americans:					CSA 54%	100%	34.7 (9.4)
Nia: empowerment focused	Reduce PTSD symptoms	Psychoed	10	45	34, 31		(100%)	
TAU: 'standard psychiatric care' (including IPV support groups)	*PTSD, Dep, PDist*	-		44	35, 31			
Kelly et al (2016)	Reduced trauma symptoms,			*ITT*		100%	100%	41.5 (14.6)
TI-MBSR (Mindfulness)	psychoeducation	Psychoed	8	24	20	violence	(38%)	
WL	*PTSD, Dep, PDist*	-	-	21	19			
Krakow et al (2001)	Sexual Assault Survivors:			*ITT*		54% CSA	100%	
Imagery rehearsal for nightmares	treat chronic nightmares	Psychoed Plus (1)	3	80	39*	(72% CPA)	(95%)	36.0 (9.8)
WL	(PTSD)	-		88	41			40.2 (11.3)
Krupnick et al (2008)	Low income: treat 'highly			*ITT*		95.8% CSA	100%	32 (10.2)
Interpersonal psychotherapy	chronic' PTSD	Psychoed Plus (4)	16 (1)	32	20*		(100%)	
WL	*PTSD, Dep*	-		16	7		(100%)	
**Completers with over 50% attendance*								
Lau et al (2007)	Treat CSA			*ITT*		100% CSA	100%	
Systemic Group Therapy	*PDist*	Psychoed Plus (8)*	34*	54	46	(incest)	(100%)	32.4 (8.8)
Analytical Group Therapy		TMP	46*	52	40			34.2(10.5)
**Mean number of sessions*								
McWhirter (2011)	Treat IPV mothers and			*ITT*		100% IPV,	100	30 (18-47)
Emotion-focused	children*	Psychoed	5	22	21	CPA 89%*	(DK)	
Goal-orientated	*Dep, SubM*	NTG	5	24	21			
**Results from mothers only*								

Table 2 continued

Study **Type of group treatment and control condition(s)** *Study relevant details*	**Population, Treatment Goal** *Symptom Domain*	**Treatment Categorisation** (No. of psychoed sessions) [1]	**No. planned sessions**[2] (No. of individual sessions)	***N*** in study *ITT if Inc.*	***N post***[3]	**Sample and Trauma Details** Primary abuse details	% female (% full PTSD)	Age M (SD)
Messina et al (2010) Gender Responsive Treatment* TAU: Therapeutic Community **Includes Beyond Trauma (11 session trauma programme* *** 94 participants (83% of the total sample) completed 6mth measures.*	Prison: Treat substance misuse and reduce trauma symptoms *PDist, SubM*	 Psychoed Plus (11) -	 28 DK	*ITT* 60 55	 94**	55% SA 71% PA	100% (25%) (26%)	35.9 (9.6)
Messina et al (2012) Gender Responsive Treatment* TAU: Drug treatment programme **As described above* *** 57% and 58% respectively*	Community drug misusing offenders: treat substance misuse and reduce trauma symptoms *PTSD*	 Psychoed Plus (11) -	 28+ DK	*ITT* 85 65	 48** 38	55% SA	100% (31%) (26%)	36 (8.9)
Mitchell et al (2014) Yoga for PTSD / DBT skills WL	Treat traumatic stress and improve emotional regulation. *PTSD, Dep, PDist*	 Psychoed -	 12 -	*ITT* 14 12	 20 18	88% CSA	100% (70.7%) (100%)	44.3 (12.3)
Rieckert & Moller (2000) REBT for CSA WL	Treat post abuse symptoms in the absence of PTSD diagnosis *Dep, PDist*	 TMP (10) -	 10	 28 14	 26 14	100% CSA	100% (0%)	28
Sikkema et al (2007, 2013) LIFT (trauma coping group) Support Group Waiting List	Treat traumatic stress, HIV and drug use. *PTSD, SubM*	 Psychoed Plus (15) NTG -	 15 15	*ITT* 124 123	 73 77 48	100% CSA	54% (40%)	42.5(6.9)

Table 2 continued

Study **Type of group treatment and control condition(s)** *Study relevant details*	**Population, Treatment Goal** *Symptom Domain*	**Treatment Categorisation** (No. of psychoed sessions) [1]	**No. planned sessions**[2] (No. of individual sessions)	***N*** in study *ITT if Inc.*	***N post***[3]	**Sample and Trauma Details** Primary abuse details	% female (% full PTSD)	Age M (SD)
Tirado-Munoz et al (2015)	IPV safety and stabilisation;			*ITT*		IPV (100%)		
IPaVit-CBT	treat substance misuse and	Psychoed	10	7	7		100%	42 (5.6)
TAU: 'outpatient drug centre' *	depression	-	-	7	6			39.8 (11.6)
**'various professionals'*	*Dep, SubM*							
van der Kolk et al (2014)	Treatment resistant PTSD:			*ITT*		DK	100%	
Trauma- informed yoga	affect regulation, PTSD	Psychoed	10	32	31		(100%)	41.2 (12.2)
Woman's Health Education (WHE)	symptoms	NTG	10	32	29			44.3 (11.9)
	PTSD, Dep, PDist, Diss							
Yeoman et al (2010)	Low income country:					98.8%	44.4%	38.6 (12.8)
'Psychoed' workshop	provision of culturally	Psychoed	4*	41	38	Combat		
'Non-Psychoed' workshop	specific PTSD treatment	TMP	4*	41	37	trauma		
WL-ITT	*PTSD, PDist*	-		42	38			
**Full days*								
Zlotnick et al (1997)	Treat CSA					100% CSA	100%	39 (9.59)
Affect Management*	*PTSD, Diss*	Psychoed	15(DK)	24	17		(100%)	
WL		-	-	22	16			
Zlotnick et al (2009)	Prison*: treat PTSD and					Sexual	100%	34.6 (7.4)
Seeking Safety	substance misuse.	Psychoed Plus (15)	15 (3)**	27	23	abuse	(83.5%)	
TAU (residential 12 Step AA model)	*PTSD, PDist, SubM*	-		22	21	93.9%		
**Minimum security located in a residential treatment programme*								
***Mean No. sessions planned 25(12)*								

Abbreviations used. TF-IT: Trauma Focused Interpersonal Transaction psychotherapy; CPT-SA: Cognitive Processing Therapy - Sexual Abuse; IPT: Interpersonal Therapy; Psychoed:

Table 2 continued

psychoeducation; CBT: cognitive behavioural therapy; TF-CBT: trauma focused CBT; NTG: Non-Trauma-Group (active control); WL:. Waiting List; TAU: Treatment As Usual. *Symptom domain abbreviations used.* PTSD: Post Traumatic Stress Disorder; Dep: Depression; PDist: Psychological Distress; SubM: Substance Misuse; Diss: Dissociation. *Treatment categorisation abbreviations used.* Psychoed: Psychoeducation; Psychoed Plus: Psychoeducation Plus; TMP: Trauma Memory Processing.

[1] Psychoeducational content within TMP and Psychoed Plus interventions that specifically relates to PTSD / CPTSD symptoms as specified or estimated from available information.

[2] Total number of group sessions planned unless otherwise indicated as mean sessions M (SD).

[3] N: N based on completers (i.e. completers analysis) at post intervention evaluation

Table 3: Summary of study and treatment characteristics

	Frequency	Percentage		Frequency	Percentage
Year of Publication			*Number of group sessions*		
≤ 1999	2	5.6%	<10	7	19.4%
2000-2010	19	52.7%	10-20	22	61.1%
2011-2016	15	41.6%	>20	7	19.4%
Country of Origin			*Session Length*		
United States	29	80.6%	50-60 mins	2	5.5%
United States / Africa	2	5.6%	61-90 mins	17	47.2%
Europe	4	11.1%	91-180 mins	17	47.2%
Other	1	2.8%			
			Frequency		
RCT Comparator			Once per week	27	75.0%
Waiting List / Minimal Contact only	12	33.3%	Twice per week	4	11.1%
Treatment As Usual only	9	25.0%	Three or more times per week	5	13.9%
Active Comparison only	8	22.2%			
Active Comparison and WL/TAU	7	19.4%	*Treatment duration*		
			<10 weeks	12	33.3%
Pre-therapy group differences			10-20 weeks	22	61.1%
None	23	63.9%	>20 weeks	2	5.6%
Some, unclear importance	6	16.7%			
Some, important	7	19.4%	*Structure*		

Table 3 continued

	Frequency	Percentage		Frequency	Percentage
			Manualised	34	94.4%
Drop/out completer differences			Semi-structured	2	5.6%
None	30	83.3%			
Some / unclear importance	3	8.3%	*Therapy Process*		
Some, important	3	8.3%	Instructional/Psychoed	18	50.0%
			Dialogue based	5	13.9%
Setting			Exposure only	5	13.9%
Community	24	66.7%	Mixed	8	22.2%
Inpatient / Shelter	6	16.7%			
Forensic (Prisons/Probation)	6	16.7%	*Number of Psychoeducational Sessions*		
			0 -dk	2	5.6%
*Treatment Type**			1-5 sessions	11	30.6%
Cognitive Behavioural	30	69.8%	6-10 sessions	13	36.1%
Insight orientated	7	16.3%	11-20 sessions	10	27.7%
Eclectic	2	4.7%			
Mindfulness / Yoga	4	9.3%	*Therapist experience*		
			Students, Assistant Practitioners	3	8.3%
Modality			Students and Practitioners	2	5.6%
Group only	34	94.4%	Standard Practitioners	7	19.4%
Combined group and individual	2	5.6%	'Experienced' Practitioners only	24	66.6%

Notes: Frequencies that do not add to 36 indicate missing data (i.e. not applicable or not reported) or multiple arms

Table 4: Summary of participant characteristics

	Frequency	Percentage		Frequency	Percentage
Age			*Marital Status*		
50% 28-35 years	16	44.4%	>50% married/partnered	10	27.8%
50% 36-40 years	14	38.9%	>50% single	3	8.3%
50% 41- 48 years	6	16.6%	>50% divorced separated	1	2.8%
			All less than 50%	14	38.9%
Gender			DK	8	22.2%
100% women	31	86.1%			
50%+ women	3	8.3%	*Index trauma/abuse*		
50%+ men	1	2.8%	Child Sexual Abuse	12	33.3%
100% men	1	2.8%	Adult Sexual Abuse	3	8.3%
			Intimate Partner Violence	6	16.6%
Ethnicity			War (non-combat)	1	2.8%
>50% white/Caucasian	22	61.1%	Mixed traumas	14	38.9%
>50% African American	2	5.6%			
>50% Hispanic American	1	2.8%	*Frequency of abuse events*		
All (above) less than 50%	7	19.4%	> 50% less than 10 child abuse events only	2	5.6%
100% African	4	11.1%	> 50% more than 10 child abuse events only	6	16.6%
			Child and Adult abuse events	3	8.3%
Education			Adult abuse only	3	8.3%
>50% less than high school	2	5.6%	> 50% more than 10 adult abuse events only	2	5.6%
>50% some high school	19	52.8%	*DK*	20	55.6%
>50% some tertiary	4	11.1%			
Mixed (all less than 50%)	6	16.7%	*Mean age of onset of abuse*		
DK	5	13.9%	12 years or less	13	36.1%
			Older than 12 years	-	-
Annual income			18 years plus	-	-
All 'low income'	27	75.0%	DK	23	63.9%

Table 4 continued

	Frequency	Percentage	Frequency	Percentage
Mixed levels of income	1	2.8%		
DK	8	22.2%		

Note: DK: Don't know, indicates missing data (i.e. not reported).

2.3.2. Methodological quality of included studies

Numerous 'unclear' ratings of study quality were made due to unexplained or insufficient detail. Primary publications often contained very little detail about randomisation or concealment, and there was a lack of clarity in reporting primary outcome data. Additional criteria adapted from Ehring et al (2014) and those from the Risk of Bias (RoB) tool are summarised in Table 5. The RoB and GRADE analysis and notes explaining these ratings can be accessed in Appendix 1.4. Similarly, Table 6 includes the outcome quality ratings for the main analyses.

Only 3 studies were low risk on all RoB criteria (Ford et al, 2013; Hollifield et al 2007 and Kaslow et al, 2010). Most studies used a treatment manual (k = 34, 91.9%), however, fewer studies used a structured clinical interview to diagnose PTSD (k = 16, 43.2%). Similarly, in the other domains few studies used diagnostic procedures as part of their post symptom measurements. In addition, studies did not consistently report data on treatment integrity in respect of quality assurance/fidelity measures (k= 16, 43.2%). Approximately, half of the studies clearly reported the use of follow-up assessments, intent-to-treat analyses, or ensured that assessors were blinded (k= 19, 51.4%; k= 21, 56.7%; k= 19, 51.4% respectively).

GRADE quality ratings of each outcome were predominantly either low or very low. This partly reflected the variety of comparators included within the analyses and heterogeneity in methodological approach, including the use of different outcome measures, particularly in the Substance Misuse and Psychological Distress outcome domains. Similarly, it was also noted in the Substance Misuse domain that 'post' treatment data collection time points varied widely from 1.5 weeks to 24 weeks. Appendix 1.4. presents further detail of studies within each analysis and the mean 'post' data collection time frame. Quality was also reduced by inconsistency, in which an unclear direction of effect was observed, in addition to wide confidence intervals.

Table 5: Cochrane risk of bias ratings and methodological quality for included studies

Study	Random sequence generation: selection bias	Performance bias: masking of participants and personnel	Detection bias: masking of assessments	Incomplete outcome data: attrition bias	Selective reporting: reporting bias	Other bias	Manualised Treatment	Data on[1] Treatment Integrity	ITT Analysis
Alexander et al (1989)	Unclear	Unclear	Unclear	Unclear	No	No	No	No	No
Bass et al (2013)	No	Unclear	No	No	No	No	Yes	Yes	No
Bohus et al (2013)	Unclear	No	No	No	Unclear	No	Yes	No	Yes
Bradley et al (2003)	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear	Yes	No	No
Chard (2005)	Unclear	No	No	Yes	Unclear	No	Yes	Yes	Yes
Classen (2001)	Unclear	No	No	Yes	Yes	Unclear	Yes	No	No
Classen (2011)	No	No	No	Unclear	Unclear	No	Yes	Yes	Yes
Cole (2007)	Unclear	Unclear	Unclear	Yes	Yes	Yes	Yes	No	No
Constantino et al (2005)	No	Unclear	Yes	Yes	No	Unclear	Yes	No	No
Crespo & Arinero (2010)	No	Unclear	Unclear	Yes	No	Yes	Yes	Yes	No
Dorrepaal et al (2012)	No	No	No	No	No	Unclear	Yes	Yes	Yes
Falsetti et al (2008)	Unclear	No	No	Unclear	Unclear	Unclear	Yes	Yes	Yes
Ford et al (2013)	No	No	No	No	No	No	Yes	Yes	No
Frisman et al (2008)	Unclear	Yes	Unclear	Yes	Unclear	Yes	Yes	Unclear	Yes
Garland et al (2016)	No	No	No	No	No	Unclear	Yes	Yes	Yes
Ghee et al (2009)	Unclear	Unclear	Unclear	Unclear	No	Unclear	Yes	No	No
Graham-Bermann (2013)	Unclear	Unclear	Unclear	Unclear	Unclear	No	Yes	No	No
Hien et al (2009)	No	No	No	No	No	No	Yes	Yes	Yes
Hinton et al (2011)	Unclear	Unclear	Unclear	No	Unclear	Unclear	Yes	Unclear	No
Hollifield et al (2007)	No	No	No	No	No	No	Yes	Yes	Yes
Kaslow et al (2010)	No	No	No	No	No	No	Yes	Yes	No
Kelly et al (2016)	Unclear	No	No	Unclear	No	No	Yes	No	Yes
Krakow et al (2001)	No	Yes	No	No	No	No	Yes	No	Yes

Table 5 continued

Study	Random sequence generation: selection bias	Performance bias: masking of participants and personnel	Detection bias: masking of assessments	Incomplete outcome data: attrition bias	Selective reporting: reporting bias	Other bias	Manualised Treatment	Data on[1] Treatment Integrity	ITT Analysis
Krupnick et al 2008)	Unclear	Unclear	Unclear	No	No	Unclear	Yes	Unclear	Yes
Lau et al (2007)	No	No	No	No	No	Unclear	Yes	No	Yes
McWhirter (2011)	No	No	No	No	No	Unclear	Unclear	No	Yes
Messina et al 2010	Yes	No	Unclear	Yes	Unclear	No	Yes	No	Yes
Messina et al 2012	Unclear	Unclear	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes
Mitchell et al (2014)	No	Unclear	Unclear	No	No	Unclear	Yes	No	Yes
Rieckert & Moller (2000)	Yes	Unclear	Yes	Yes	Yes	Unclear	Yes	No	No
Sikkema (2007/2013)	Unclear	Unclear	Unclear	Unclear	Yes	Yes	Yes	Yes	Yes
Triado-Munoz (2015)	No	Unclear	Yes	Yes	Yes	Unclear	Yes	Yes	Yes
van der Kolk 2014	Unclear	No	No	No	Unclear	Unclear	Yes	No	Yes
Yeomans et al (2010)	No	No	No	No	No	No	Yes	No	No
Zlotnick et al (1997)	Unclear	Unclear	Unclear	Yes	No	Unclear	Yes	Yes	No
Zlotnick et al (2009)	No	Unclear	Unclear	No	No	No	Yes	Yes	No

Notes: [1]Data on treatment integrity specifically refers to the reporting of quality assurance/fidelity measures as opposed to patient therapy ratings.
ITT= Intent to Treat

2.3.3. Treatment effects

Group based trauma interventions compared to usual care

Studies that had two active group treatment arms were combined (Alexander et al, 1989; Classen et al, 2001; Classen et al, 2011; Garland et al, 2016; Sikkema et al, 2007; Yeomans, 2010) and compared to usual care. Medium to large significant effect sizes favouring group based trauma interventions were found for four of the outcome domains with only Substance Misuse resulting in a small non-significant effect size (see Table 6). The I^2 statistic indicated significantly high levels of heterogeneity; apart from the Dissociation domain. Inconsistency and imprecision resulted in low to very low GRADE quality ratings apart from the Dissociation domain.

Group-based trauma interventions compared to non-trauma group-based treatments

In this set of analyses, trauma informed group treatments were compared to non-trauma group interventions, such as support groups. Non-significant effect sizes were computed with significantly high levels of heterogeneity (Table 6). In the PTSD domain it was apparent that Garland et al (2016) was a considerable outlier with an effect size lying almost 2 standard deviations beyond the adjacent effect size value (see Figure 3). In this instance, a condensed version of Seeking Safety (psychoeducation) was compared against Mindfulness-Oriented Recovery Enhancement (MORE), which was classified as a non-trauma treatment. Removing this study resulted in a non-significant effect size (k= 4, g= -0.15, SE 0.26 [95%CI, -0.67 to 0.37] p= 0.571; I^2 = 20%, p= 0.288). Although this study was not an outlier in other domains, similar non-significant effect sizes were apparent when removed.

Psychoeducational group treatments compared to usual care

For the three outcome domains with the largest number of studies contributing pre-post data (PTSD, Depression, Psychological Distress), *Psychoeducation* interventions computed various medium to large effect sizes in favour of treatment when compared to usual care; although only the first domain was statistically significant (see Table 6 and also Figure 3). The effect size for the Dissociation domain was also statistically significant in favour of *Psychoeducation* interventions but consisted of outcomes from only one study.

Garland et al (2016) was noted to be a considerable outlier in the Depression domain. Removal of this study from the *Psychoeducation* subgrouping reduced the effect size to *k*= 6, g= -0.28, SE 0.32 [95%CI, -0.91 to 0.35], *p*= 0.383; I^2= 5%, *p*= 0.380. Similarly, whilst not an outlier in the Psychological Distress domain the removal of Garland et al (2016) reduced the effect size to *k*= 5, g= -0.28, SE 0.30 [95%CI, -0.87 to 0.32], *p*= 0.361; I^2= 0%, *p*= 0.456. This was also the case in the PTSD domain *k*= 7, g= -0.25, SE 0.21, [95%CI, -0.66 to 0.16], *p*= 0.225; I^2= 0%, *p*= 0.453. The I^2 statistic also reported significantly high levels of heterogeneity when Garland et al (2016) was included (as summarised in Table 6). The main quality rating, regardless of treatment type, across the domains was low or very low except in the PTSD domain. Whilst reasons varied, this included the potential deficits involved in the combined measures utilised specifically in the Psychological Distress and Substance Misuse domains.

In the *Psychoeducation Plus* analyses only the PTSD and Dissociation domains reported significant moderate effect sizes (k= 10, g= -60, [95%CI, -1.00 to -0.20]; k= 2, g=-0.79 [95%CI, -1.19 to -0.39] respectively). Depression and Psychological Distress domains reported moderate and small to moderate, but non-significant, effect sizes (k= 4, g= -0.77 [95%CI, -1.92 to 0.39]; k= 7, g= -0.38 [95%CI, -0.91 to 0.15], respectively). Treatments for this category in the Substance Misuse domain, reported small non-significant effect sizes (k=3, g= 0.10 [95%CI, -0.70 to 0.89]).

Table 6: Effect size estimates for different treatment categorisations and comparisons

Outcome Domain	k	N group 1	N group 2	Hedges's g	95% CI, p =	I^2, p=	Quality (GRADE)
Group-based trauma interventions compared to usual care							
PTSD	24	1253	976	-0.66	-0.94, -0.37 (p=0.001)	86% (p=0.001)	Low
Depression	17	667	498	-0.95	-1.43, -0.48 (p=0.001)	93% (p=0.001)	Low
Psychological Distress	20	959	715	-0.60	-0.89, -0.32 (p=0.001)	88% (p=0.001)	Very low
Substance Misuse	7	413	260	-0.03	-0.56, 0.50 (p=0.909)	87% (p=0.001)	Very low
Dissociation	7	227	193	-0.70	-1.05, -0.35 (p=0.001)	11% (p=0.346)	Moderate
Group-based trauma interventions compared to non-trauma group-based treatments							
PTSD	5	433	431	0.36	-0.24, 0.96 (p=0.238)	96% (p=0.001)	Very Low
Depression	3	118	120	0.05	-1.06, 1.16 (p=0.926)	75% (p=0.019)	Very Low
Psychological Distress	3	126	127	0.06	-0.66, 0.78 (p=0.865)	4% (p=0.353)	Very Low
Substance Misuse	4	386	388	0.45	-0.21, 1.12 (p=0.182)	94% (p=0.001)	Very Low
Dissociation	2	61	62	0.18	-0.43, 0.80 (p=0.563)	92% (p=0.001)	Very Low
Psychoeducation group treatments compared to usual care							
PTSD	8	379	252	-0.49	-0.94; -0.03 (p=0.037)	89% (p=0.001)	Moderate
Depression	6	315	192	-0.90	-1.85; 0.05 (p=0.064)	97%(p=0.001)	Low
Psychological Distress	6	321	196	-0.51	-1.09; 0.08 (p=0.091)	88% (p=0.001)	Very low
Substance Misuse	4	267	161	-0.14	-0.86; 0.59 (p=0.714)	92%(p=0.001)	Very low
Dissociation	1	16	17	-0.82	-1.60; -0.04 (p=0.041)	0% (p=1.000)	Very low
Psychoeducation Plus group treatments compared to usual care							
PTSD	10	707	472	-0.60	-1.00; -0.20 (p=0.003)	86% (p=0.001)	Moderate
Depression	4	198	125	-0.77	-1.92; 0.39 (p=0.192)	91%(p=0.001)	Very low
Psychological Distress	7	484	338	-0.38	-0.91; 0.15 (p=0.161)	93% (p=0.001)	Very low
Substance Misuse	3	146	99	0.10	-0.70; 0.89 (p=0.813)	0% (p=0.609)	Very low

Table 6 continued

Outcome Domain	k	N group 1	N group 2	Hedges's g	95% CI, p =	I^2, p=	Quality (GRADE)
Dissociation	2	130	88	-0.79	-1.19; -0.39 (p=0.001)	0% (p=0.331)	Moderate
TMP group treatments compared to usual care							
PTSD	6	167	256	-0.98	-1.53; -0.43 (p=0.001)	85% (p=0.001)	Moderate
Depression	7	154	181	-1.12	-2.01;-0.23 (p=0.014)	86% (p=0.001)	Low
Psychological Distress	7	154	181	-0.98	-1.66; -0.40 (p=0.001)	77%(p=0.001)	Very low
Substance Misuse[1]	-	-	-	-	-	-	
Dissociation	4	81	88	-0.61	-0.97; -0.24 (p=0.001)	34% (p=0.205)	Moderate
TMP group treatments compared to Psychoeducational group treatments							
PTSD	4	132	131	-0.34	-1.05; 0.36 (p=0.337)	85% (p=0.001)	Very Low
Depression	3	103	97	0.29	-0.83; 1.4 (p=0.607)	88% (p=0.001)	Low
Psychological Distress	6	204	201	0.19	-0.34; 0.71 (p=0.491)	83% (p=0.001)	Very Low
Substance Misuse	1	30	33	1.10	-0.28; 2.48 (p=0.118)	0% (p=1.000)	Very Low
Dissociation	1	55	56	-0.12	-0.92; 0.67 (p=0.759)	0% (p=1.000)	Very Low

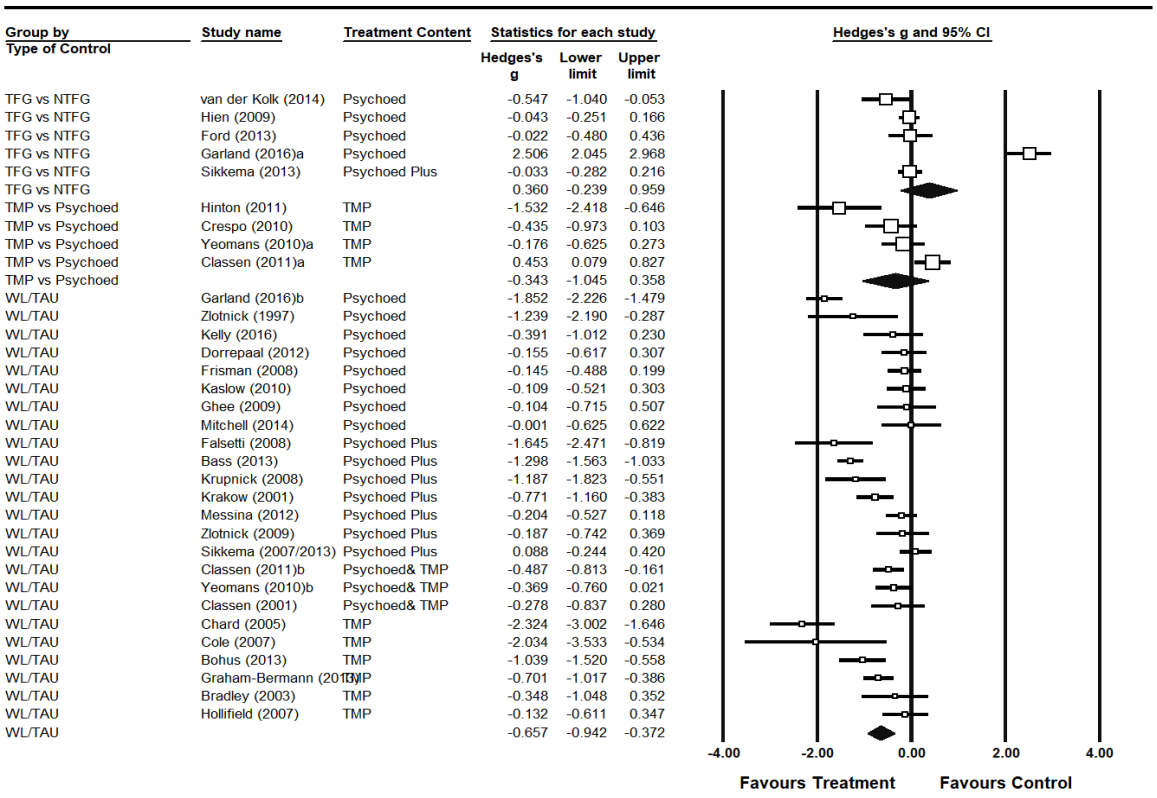
Notes: CI = confidence interval, I^2= I statistic. 1. No included TMP studies measured substance misuse reduction

TMP group treatments compared to usual care

When compared to usual care, significant large effect sizes for *TMP* based treatments were evident in all outcome domains apart from Substance Misuse where there were no available studies. As noted in Table 6 the I^2 statistic indicated significant levels of heterogeneity for the three domains with the largest number of studies. It was also noted that in the Depression domain that if Rieckert & Moller (2000), which was considered to have a particularly high risk of bias, was removed there would be a reduction in the effect size obtained in the *TMP* category (*k*= 6, g= -0.89, SE 0.48 [95%CI, -1.84 to -0.06], *p*= 0.07, I^2= 81%, *p*= 0.001).

Comparing TMP with Psychoeducation group treatments

Few studies directly compared treatment arms categorised as *TMP* interventions with psychoeducation interventions and only one study reported data for the Substance Misuse and Dissociation domains (Classen et al, 2011). A small non-significant effect size in favour of *TMP* treatments was observed for PTSD (k=4, g= -0.34 [95%CI, -1.05 to 0.36), whereas non-significant effect sizes were computed for *Psychoeducation* treatments for Depression (k=3, g= 0.29 [95%CI, -0.83 to 1.40]) and Psychological Distress (k= 6, g= 0.19, [95%CI, -0.34 to 0.71]; see Table 6). As with the other subgroup analyses for treatment arm comparisons, the small number of available studies and considerable heterogeneity issues contributed to the very low GRADE quality ratings.



Notes. Psychoed= Psychoeducation, TMP= Trauma Memory Processing, TFG= Trauma Focused Group, NTFG= Non-Trauma Focused Group (Active Control). Classen (2011)a. TMP arm compared against Psychoeducation arm. Classen (2011)b. Combined TMP and Psychoeducation arm compared against WL arm. Garland (2016)a. CBT Psychoeducation arm compared against NTFG Psychoeducation: Mindfulness-Oriented Recovery Enhancement (MORE). Garland (2016)b. Combined CBT and MORE arms compared to TAU. Sikkema (2007) 'Preliminary' outcome study utilised for WL/TAU comparator. Sikkema (2013) *Psychoed Plus* compared against NTFG. Yeomans (2010)a. Psychoeducation compared to a TMP arm. Yeomans (2010)b. Combined Psychoeducation and TMP arms compared to WL arm. [See supplementary online material for categorisation decisions]

Figure. 3. Forest plot for PTSD symptoms post treatment effect sizes estimates as grouped by comparators

2.3.4. Heterogeneity

As noted in Table 6, the I^2 statistic often reported significantly large amounts of heterogeneity. Subgroup procedures have been used in an attempt to manage heterogeneity; according to comparator and treatment type. It was however apparent that most I^2 analyses were over 40% and that reductions in heterogeneity appeared to reflect analyses with smaller numbers of available studies.

2.3.5. Publication bias

Inspection of funnel plots (Appendix 1.3.) for studies compared to usual care indicated that there tended to be a wider dispersal of studies to the left of the mean; apart from

the Dissociation domain, which showed the greatest symmetry. This indicates the potential presence of publication bias. Egger's test and the rank correlation test were not significant for all of the domains indicating that overall, smaller n studies did not seem to report higher effect size than the larger studies. Orwin's failsafe N indicated that additional studies (n= 60 to 45) would be needed to reach an effect size with a 'trivial' value (i.e., -0.20) for the domains with the larger number of included studies. Given that only 36 studies were identified this is unlikely in the near future. As such the current heterogeneity of treatment approaches, the diversity of psychological effects of complex interpersonal trauma, which of course may or may not be targeted in treatment, are likely to impact on the effect sizes computed. In the Substance Misuse domain, Orwins's failsafe N indicated that 7 missing studies would be needed to reach a modest 'trivial' value of -0.10 and in the Dissociation domain, 19 studies would be needed to reach a 'trivial' effect size at -0.20. The trim and fill method suggested that for only two domains would additional studies be added to the left of the mean to give an adjusted effect size (Depression, n=2; Substance Misuse, n=2). The adjusted effect sizes would be g= -1.09 and g= -0.30. With respect to the Depression domain this would represent only a small improvement to the current outcomes in favour of trauma focused group treatments but these additional studies would constitute a potentially greater improvement to the outcomes currently evidenced in the Substance Misuse domain.

2.3.6. Moderator analysis

Given the small number of studies involved in the Dissociation and Substance Misuse domains, it was difficult to meet the criteria as described by previous authors to identify potential moderators (Hedges, Tipton & Johnson, 2009). Scatterplot analysis also indicated that the significant moderators in the Substance Misuse domain were the result of the large effect size reported by Garland et al (2016). There were also no significant moderators in either the Psychological Distress or Dissociation domains.

Participants' mean age in the PTSD domain indicated that trials with older participants reported significantly lower effect sizes than trials with younger participants (*k*=24, *r*= 0.10, SE 0.04; [95%CI, 0.02 to 0.17], p= 0.012). Inspection of the relevant scatterplot noted that this was particularly robust finding. However, a similar inspection of the scatterplot for gender and treatment setting in the Depression domain again noted

that Garland et al (2016) was a particularly influential study, for example, it was the sole study to include only male participants. When removed from the analysis these variables were no longer significant. Appendix 1.4. presents a summary of the variables computed in the moderator analysis across the various domains.

2.4. Discussion

2.4.1. Summary of findings

The results from the three outcome domains with the largest number of studies (PTSD, Depression and Psychological Distress) indicate that *TMP* interventions had large significant effect sizes when compared to usual care comparators. Medium to large effect sizes were also found in these domains for *Psychoeducation Plus* and *Psychoeducation* interventions against usual care comparators; although these were non-significant for Depression and Psychological Distress. However, when outliers were taken into account the effect sizes in favour of the *Psychoeducation* category were substantially reduced, and whilst still non-significant, heterogeneity was largely accounted for.

Therefore, when outliers were accounted for, indirect comparisons between treatment categories (i.e. *TMP*, *Psychoeducation Plus* and *Psychoeducation*) suggest incremental increases in effect sizes when compared to usual care. This would suggest that treatment efficacy should be defined by the inclusion of protocols that assist with the processing of the traumatic memories. However, when *TMP* and *Psychoeducational* interventions were directly compared there was no clear effect. Important between-study differences, for example baseline symptom severity, may have resulted in TMP studies having a larger effect size when compared to usual care. Given that direct comparisons are empirically more robust, greater weight should be placed on these analyses when considering the relative efficacy of TMP and psychoeducational treatments. Similarly, it is also important to note that whilst the results for group trauma treatments are favourable when compared to usual care this was not the case when other active non-trauma group comparators were used; although this may reflect the limited number of available studies (k= 2 to 5). In addition, the difficulty of treating comorbid substance misuse requires specific consideration, as no intervention was clearly effective at reducing this.

2.4.2. Are TMP treatments more effective than psychoeducational treatments?

The findings of this meta-analysis when interventions are compared to usual care correspond to a theoretical perspective that promotes the integration of TMP in treatments for symptoms and psychopathology associated with complex interpersonal trauma. However, a small number of studies have directly compared TMP with psychoeducational treatments and these findings present a mixed set of results. These direct comparison studies in the PTSD domain returned a small non-significant effect size in favour of TMP interventions. Comparisons in the Depression and Psychological Distress domains resulted in small non-significant effect sizes. This suggests that TMP interventions, as designed, may have more impact on symptoms associated with PTSD but not the wider psychopathology associated with complex interpersonal trauma. Psychoeducation interventions appear to have equal if not more benefits in ameliorating these symptoms. Indeed, psychoeducational treatments might be useful for the treatment of general distress that survivors of interpersonal violence often report. This corresponds to the results, as already noted, when psychoeducational treatments were compared to usual care controls. Overall, results suggest that TMP interventions may be useful for traumatic stress whereas non-TMP interventions may be useful for symptoms of general distress (e.g. anxiety and depression).

2.4.3. Comparison of findings with other meta-analysis

The results of our review with respect to usual care comparators concur with previous meta-analysis where there has been a greater effect size for trauma memory focused interventions (Bisson et al, 2007, 2013). Previous reviews have not, however, investigated the efficacy of group treatments from a phase-based perspective (Barrera et al, 2013; Cloitre, et al, 2012; Mendelsohn et al, 2011; Sloan et al, 2013).

The results from this meta-analysis suggest that TMP group treatments are particularly effective for PTSD symptoms, when compared to usual care. These findings support results from meta-analyses on the effectiveness of individual treatment protocols; particularly where single-arm comparisons have produced larger effect sizes for TMPs (Bisson et al, 2007, 2013; Lenz, Haktanir & Callender, 2018; Roberts et al, 2015; Taylor & Harvey, 2010; Watts et al, 2013). Similarly, results also

support the creation of a *Psychoeducation Plus* category to help explain the potential efficaciousness of a phase 1 'stabilising' intervention.

Importantly, the results from this meta-analysis also concur with studies reporting the relative efficacy of psychoeducational groups, and indeed non-trauma group interventions. In this respect, previous meta-analyses have highlighted that all psychotherapeutic responses generally promote recovery in PTSD symptoms (Erford et al 2016; Lenz, 2018). Similarly, this meta-analysis would seem to concur with previous findings that non-specific interventions are equally efficacious particularly for individuals with complex clinical presentations (Greger et al, 2014). Foa et al (2018) recent large scale RCT, comparing the effectiveness of individually delivered TMP and 'present centred' interventions, with active duty military personnel also indicated that there was no significant difference between these arms. The synthesis of high quality RCTs, for both individual and group based treatment modalities, remains an important endeavour in psychological trauma reviews.

2.4.4. Implications for clinical practice

Arguments have been recently advanced questioning the potential impact of delaying essential trauma processing treatments (de Jongh et al, 2016). If considering this meta-analysis with respect to comparisons against usual care the results add weight to the superior effectiveness of TMP interventions particularly for PTSD, Depression and Psychological Distress symptoms. Similarly, as noted in the direct TMP and psychoeducation comparisons the former interventions were still demonstrated to be equally as effective for PTSD symptoms and therefore may still be the intervention of choice for treatment providers. However, the timing, nature and intensity of such processing elements should be subject to further research. As only post treatment effect sizes have been used, this review does not provide a complete analysis of participants' treatment journeys including any potential for temporary symptom exacerbation (Crawford, Thana, Farquharson, Palmer, Hancock, Bassett, Clarke & Parryvan, 2016; Mott, Sutherland, Williams, Lanier, Ready, & Teng, 2013; Resick, Galovski, Uhlmansick, Scher, Clum, & Young-Xu, 2008; van den Berg, de Bont, van der Vleugel, de Roos, de Jongh, van Minnen, van der Gaag, 2016). There is clearly a need

to investigate the long-term outcomes of both psychoeducational and TMP interventions.

Whilst the specific benefits of delivering interventions consisting solely of psychoeducational material have been questioned, it should also be noted that medium to large aggregated effect sizes were computed for these interventions. However, it is important to consider the impact that outliers had on these effect sizes. The inclusion of both CBT and mindfulness/interoceptive interventions into this category should also be considered. When outliers were removed, small, non-significant effect sizes across all outcome domains, except substance misuse, were noted in favour of psychoeducational treatments. This may be more than acceptable with respect to their public health utility particularly if large-scale programmes, with high degrees of treatment integrity, can be more easily delivered (Brookes, Barrett, Netten & Knapp, 2013). The accessibility of psychoeducational interventions, particularly for populations that have often been regarded as too chaotic or unstable for TMP interventions, also makes this an attractive option (Corrigan & Hull, 2015). Indeed, the dearth of TMP group treatments for comorbid substance misuse would suggest that such exclusion criteria already has an impact on how viable such treatment options are considered.

This meta-analysis suggests that early interventions that are offered as part of a phased approach or are either symptom specific or more intensive (i.e. *Psychoeducation Plus*) are potentially more effective than usual care in ameliorating PTSD (including Dissociative) symptoms. As such, although having the opportunity to safely process trauma based memories is important so too is ensuring that survivors have the specific skills with which to cope with their symptoms. It is therefore important that psychoeducational interventions are matched to an individual's treatment needs or of sufficient intensity that ensures substantial progress is achieved and maintained (Ali, Rhodes Moreea, McMillan, Gilbody, Leach, Lucock, Lutz & Delgadillo, 2017).

The only moderator of significance, after outliers were accounted for, was age within the PTSD domain. Whilst this requires further replication, it suggests that older participants, perhaps with greater histories of repeat traumatisation or symptom

accommodation, represent an increased challenge to treatment programmes. This may raise important clinical considerations relating to age and its link to possible treatment resistance and whether current treatments are sufficiently responsive to older participants (Clapp & Beck, 2012; Pietrzak, Goldstein, Southwick, & Grant, 2012).

2.4.5. Implications for future research

Few studies in this review measured the impact of motivational, normative or empowerment processes that are commonly associated within a psychoeducational group treatment (Burlingame et al, 2003; DiClemente, Schlundt, & Gemmell, 2004; Herman, 1992; McCrone et al, 2005; Mendelsohn et al, 2011). Similarly, the disparity between the different measures used, particularly within the *Substance Misuse* domain, presents a challenge to research in addressing the heterogeneity within this area. It is, however, reasonable to conclude that survivors with co-morbid substance misuse difficulties present substantial clinical and research challenges. Najavits & Hien (2013) and this meta-analysis highlights the need for more high quality trials of *full* dose substantive interventions such as *Seeking Safety*; which has currently only been undertaken once in a prison setting (Zlotnick et al, 2009). It should also be noted that the range of mean post treatment reporting times in the studies synthesised to answer the research questions for this outcome domain was from 8 weeks to 24 weeks with a number of the studies only providing data at 6 or 12 months after treatment (Classen et al, 2011; Meade et al 2010; Messina et al, 2010). This delayed reporting of treatment outcomes of course has a considerable impact on the understanding of post treatment efficacy. Further research is required before group treatments can be considered effective for comorbid substance misuse.

One of the strengths of this meta-analysis is the extensive consideration given to the categorisation of included treatment arms. It is also possible that on occasions, decisions may prove to be somewhat controversial. Whilst evidence of these decisions has been provided within the supplementary material the potential for subjectivity should be considered and analyses replicated. Another controversial aspect of this review is including studies based on experiences of interpersonal violence rather than diagnostically established PTSD symptoms. As such it can be difficult to conclude whether the reduction in symptoms is the result of the interventions attended.

Another potential limitation and complicating factor that needs to be clearly understood by the reader is the effect that various comorbidities may have had upon the results of studies included in this review. Further research on drop-out rates and comparisons with treatment completers for different types of phase based group interventions could also be usefully undertaken.

It should also be noted that there is a small number of RCT studies that have been conducted within prisons and forensic populations and it should not be assumed that trauma focused interventions are equally as effective when conducted in these challenging settings (Ball, Karatzias, Mahoney, Ferguson & Pate, 2013; Wolff et al, 2015). The complexity of these settings and indeed the complexity of often co-occurring presentations and other psychopathologies may limit the effectiveness of otherwise efficacious interventions (Bowen, Jarrett, Stahl, Forrester & Valmaggia, 2018; Cislo & Trestman, 2013). Until the evidence base develops further the inclusion of repeat measure studies within future meta-analyses for such populations may be inescapable. A similar meta-analysis could also be usefully conducted examining the effectiveness of group-based treatments on military and combat based trauma (Barrera et al, 2013; Bradley et al, 2005).

The use of the intragroup correlation (IGC) measure has been advocated by some authors to account for the extent to which group membership has created a dependency between observations. This is intended to avoid Type 1 errors by taking into account that additional variables might impact on group treatment outcomes (Baldwin, Murray, & Shadish, 2005). Criticisms have therefore been raised about using the individual participant as the unit of analysis (Shea et al, 2009; Sloan, 2013). Although this meta-analysis focused on group level data across all included studies thereby to some extent negating some of these concerns such statistical procedures could be explored further (Barrera et al, 2013).

An important finding of this review is that there are currently far too few high quality RCT studies. Indeed, many studies have inadequately reported details of randomisation and blinding procedures (see Appendix 1.4. for GRADE and Risk of Bias summaries). This very poor quality research literature necessarily impacts on the quality of this, and indeed any meta-analytic review, undertaken in this area of clinical

research. As seen in this meta-analytic review the inclusion of poor quality studies creates uncertainty around meta-analytical estimates (Nelson, Simmons & Simonsohn, 2017). It is therefore increasingly incumbent on authors to conduct trials that improve on the quality of what is currently available. Trials that are single-blind, report their randomisation sequence and how they concealed it, are adequately powered, pre-registered, using valid/reliable measures, use ITT analysis and keep drop-out / missing data to a minimum (i.e., below 20%) are imperative. A clear improvement to this situation would be the accessibility of raw data within the public domain (Nelson et al, 2017).

Whilst heterogeneity issues were taken into account within the GRADE analysis the relatively few high-quality, randomised studies available is perhaps particularly evident in respect to direct TMP versus psychoeducational comparisons. It is crucial that multifaceted and dismantling studies are undertaken in which such treatments are directly evaluated as distinct arms utilising both group and individual treatment modalities. As noted in this review, we have only located one study that compared the same TMP based treatment in both group and individual treatment arms (Stalker & Fry, 1999). Comparing such treatment modalities to each other and to non-trauma focused skills based interventions and usual care should help further develop the evidence base. It should also help to ascertain whether psychoeducational interventions are a useful or indeed necessary step to enable survivors' readiness for TMP interventions. Similarly, pragmatic and clinically important questions remain as to which treatment model is likely to be the most effective for brief group based psychoeducational interventions; those based on CBT approaches or those based on mindfulness and interoceptive awareness. The intensity and sequencing of such interventions should also be explored in further research.

2.4.6. Conclusions

There is increasing recognition of the profound impact that experiences of complex interpersonal trauma can have on the developmental trajectory and lives of survivors. Similarly, the debilitating role that symptoms of complex traumatisation such as dissociation and other comorbidities, for example substance misuse, can have are also being increasingly recognised (UKPTS, 2017; Karatzias et al, 2016). Few high quality

RCTs have examined the efficacy of treatments to ameliorate these important symptoms and they may require very different treatment approaches than have been seen to be effective with other symptoms. It is important to reflect on the impact that outliers had on the results of meta-analysis with usual care comparators. With this in mind, although psychoeducational approaches hold some promise for symptoms of general distress, it is also apparent that TMP interventions, which have been recognised as phase 2 interventions, hold the most promise particularly for symptoms of traumatic stress. Whether a phase based conceptualisation of treatment for complex trauma is actually preventing some clients from recovering as quickly as they could is difficult to ascertain from this review. Certainly, even small treatment effect sizes particularly in respect of general distress from large-scale trauma focused programmes may be welcome. However, further work is required to consider these issues in more depth.

3. Methods: Randomised control trial

3.1. Introduction to the Survive & Thrive (Prison Version) randomised control trial

3.1.1. Developing an evidence base for stabilising female prisoners with interpersonal trauma

The current evidence base highlights the importance of investigating the efficacy of psychoeducational interventions designed to help ameliorate interpersonal trauma for female offenders. As noted by several authors who have promoted a phased based approach psychoeducational interventions are considered useful in promoting emotional stabilisation through the provision of coping skills and establishing a sense of safety (Herman, 1992; Courtois & Ford, 2013). Such interventions may be particularly important for forensic services, who cater for potentially large populations of survivors and who are responsible for providing suitable levels of care and responsive interventions whilst individuals are in custody.

To date there are a limited number of interpersonal trauma interventions that can be defined as psychoeducational. As noted in the previous chapters there is a dearth of well researched randomised control trials (RCTs) that consider the efficacy of such interventions for female offenders. Similarly, in the meta-analysis undertaken for this thesis small to medium effect sizes were computed in favour of psychoeducational interventions when compared to usual care. Although these effect sizes were non-significant it suggests that psychoeducational treatments may have an advantage over usual care however this needs further investigation within a prison setting.

3.1.2. Intervention resource: selection and justification

As discussed in the Introduction chapter there is debate as to what might constitute as an appropriate psychoeducational or adjunctive intervention (Dorrepaal, Thomaes, Smit, van Balkom, Veltman, Hoogendorn & Draijer, 2012; van der Kolk, Stone, West, Rhodes, Emerson, Suvak & Spinazzola, 2014). Survive & Thrive (Ferguson, 2013) has been increasingly considered as a useful intervention for promoting emotional and behavioural stabilisation and delivered in multiple community mental health settings across Scotland (UKPTS, 2017). There was therefore an empirical and ethical

imperative to adequately establish the efficacy of this intervention and its appropriateness for use in a prison setting and the mental health difficulties inherent to that environment (Bowen et al, 2018; Hansen, Birmingham, Harty et al 2011). Similarly, as a 'pure' psychoeducation intervention it provided an important platform to understand the stabilisation and other therapeutic benefits that could be achieved from other such interventions. Survive & Thrive was therefore selected as the intervention of choice to use in this study particularly as it has a positive emerging community evidence base (Karatzias, Ferguson, Chourliara, Gullone, Gosgrove & Douglas, 2014). Support and consent for use of this intervention in the study was obtained from the author (Ferguson, 2013) and the Survive & Thrive National Reference Group who were keen to further establish the efficacy of this intervention with high-risk groups of incarcerated female offenders.

The use of an RCT design was also considered appropriate as there have already been preliminary pilot studies on the efficacy of Survive & Thrive in community settings (Ball, Karatzias, Mahoney, Ferguson & Pate, 2013; Karatzias et al, 2014). Various findings have been evident in these pilot studies. However, high dropout rates had also been reported (46% and 43% respectively) indicating that a more careful selection of participants is potentially required. As discussed in the Introduction chapter these pilot studies have reported mixed results and there was a need both for replication and for an adequately powered RCT, including follow-up assessments, to investigate the efficacy of Survive & Thrive particularly for use within a prison setting.

RCTs are considered the most rigorous way of determining cause-effect between treatment and outcome. It was hoped to establish what the clinical outcomes of Survive & Thrive are from both self-report and staff observations so as to ensure the highest standards of research and objectivity. In addition, it was also important to understand the impact that standard care (i.e. usual care within a prison setting) could have on participants' recovery (Crawford, Thana, Farquharson, Palmer, Hancock, Bassett, Clarke & Parry, 2016; Duggan, Parry, McMurran, Davidson & Dennis, 2014).

3.2. Survive & Thrive Prison Version

3.2.1. Mutualisation and responsivity of the intervention

Survive & Thrive (Ferguson, 2013) as previously described is an established manualised community intervention designed to promote emotional and behavioural stabilisation in survivors of complex interpersonal trauma (Karatzias et al, 2014). This manualised intervention, as delivered in community settings, consists of 8-10 weekly sessions.

In order to increase the acceptability and responsivity of the intervention to a prison setting amendments were made to the community version. These changes to the Survive & Thrive protocol were piloted (n=11 participants) prior to the commencement of the trial to help ensure responsivity and acceptability of the material as well as to consider any relevant organisational difficulties. These changes included material that reflected a greater gender responsivity as well as information relevant to prison and forensic resources (Covington & Bloom, 2004). These changes were also intended to make the Survive & Thrive more accessible to a range of cognitive and educational abilities that are often present in prison populations (Stewart, Wilton & Sapers, 2016). On the basis of positive participant feedback and a reduction in self-reported symptomatology, as seen in the measures used in the pilot, approval was also granted by the Scottish Prison Service's *Approved Interventions Panel* to deliver this Survive & Thrive as part of the interventions offered in the female estate. As such the adjustments made to the community protocols were considered to have met the standards necessary for Survive & Thrive's subsequent delivery in the trial. These changes are further described in Table 7 below.

3.2.2. Treatment theory and psychoeducational techniques

The prison version continued to have the same sessions and treatment aims as the community version. This included increasing the stabilisation of specific behavioural and mental health difficulties, helping participants make links between past traumas and current symptoms and preparing for further trauma focused or other relevant therapeutic work. The prison version of Survive & Thrive ensured consistency with other prison based treatment interventions particularly with respect to anger management. Other changes included a greater development of the session focusing

on shame and guilt and a greater emphasis on compassion and mindfulness based approaches. The motivational statements used at the end of every session were also replaced with more concrete versions accompanied by inspiring images depicting positive and healthy women within the PowerPoint slides.

The psychoeducational basis of Survive & Thrive was emphasised to participants both prior and during the intervention. For example, participants were informed that Survive & Thrive was a 'course' and not 'therapy'. The delivery rooms and environment were also arranged to emphasise this with participants sitting at tables as was the use of PowerPoint slides and a participant course booklet.

.

3.2.3. Treatment targets

An outline of the sessions and their contents are included below in Table 7.

Table 7. Components of the Survive & Thrive Prison Version intervention

Weeks	Session	Contents	Changes made for Prison Version
1	**Session 1: Introduction**	Introduction to the course and metaphors linked to recovery process. Emphasises on establishing safety as an important first step. Breathing exercises introduced.	Information on prevalence of abuse and trauma in prisons. Establishing safety in prisons. Additional instructions on breathing exercises and how to manage any disconcerting experiences associated with this.
	Session 2: Effects of abuse & trauma	Introduction to different forms of abuse and trauma as well as the impact on self, relationships and others. Explanation of the phased based approach to recovery. Brief mindfulness awareness exercise also introduced	Additional material to help explain the importance of mindfulness and interoceptive awareness.
2	**Session 3: Keeping safe & getting started**	Further safety/re-victimisation material including the cycle of relational abuse, the impact of substance misuse and other harmful ways of coping, promoting self-care behaviours and plans.	Additional information on power, control and abusive connections and intimate relationships. Further material and resources on coping with domestic abuse in prisons. Information and exercises promoting healthy relationships, attachments and boundaries. Compassionate based other exercise.
	Session 4: 'Surviving the surviving'	Further instruction of safe ways of coping. The 'pain paradox' introduced and coping positively with emotional difficulties and self-harm explored.	Further information on timeout exercises and resources to manage self-harm in prison.
3	**Session 5: Anxiety & coping with anxious feelings**	The brain – body connection explained including the neurological impact of trauma and avoidant behaviours. Awareness and practice of challenging negative cognitions, visualisation/guided imagery exercises and other skills to promote relaxation and stress control.	Additional images and material to help explain pre-existing content.

Table 7 continued

Weeks	Session	Contents	Changes made for Prison Version
	Session 6: Anger & coping with angry feelings	Normalising anger as well as too little vs. too much anger. Awareness of and challenging cognitions that promote anger. Practice of grounding and calming visualisation skills	Further development of exercises and skills to manage anger and to ensuring consistency with other prison based interventions and the custodial environment. Additional CBT information and exercises connected to relevant beliefs and thoughts to help participant's express their anger appropriately. Compassion based material and exercises as well as the neurological explanation of changing response systems.
4	**Session 7: Depression, what it is and how to cope**	Recognising signs of depression/ low mood, understanding the link to trauma/abuse. Cognitive and behavioural and somatic patterns (inc posture – breath work) associated with depression explained. The development of support networks and self-esteem.	Further information on introjection, parenting and childbirth difficulties for survivors who are also mothers. Skills emphasising observation, non-judgement and the connection between posture and emotions. Further motivational messages and additional material concerning coping with suicidal feelings and support networks in prison.
	Session 8: Understanding & coping with shame and guilt	Normalising the prevalence of shame and guilt, identifying relevant cognitive and behavioural patterns. Self-care, mindfulness, visualisation and self- soothing exercises revisited.	Additional material linked to individual responses to shame and guilt and common thoughts that survivors have. Further material on compassionate versus shame based self-attacking and revisiting compassionate based other exercise.
5	**Session 9: Understanding flashbacks, nightmares and disassociation**	Understanding neurological mechanisms including triggers, how to 'switch on' as well as coping skills, grounding etc.	Further neurological material to help explain dissociation as well as mindfulness and grounding excises to help manage these difficulties.
	Session 10: Assertiveness, looking back and looking forwards	Compassion focused, self-esteem, patterns of communication, assertiveness training, rights and responsibilities, Stage 2 planning.	Further images and information on assertiveness. Additional inspirational material and explanation of prison based resources for further treatment.

3.2.4. Dose of treatment and method of delivery

From the initial pilot and from a population review at the trial sites it was concluded that in order to ensure adequate numbers of participants the intervention needed to be delivered over a condensed number of weeks. This was also to ensure that short term sentenced prisoners and other prisoners with imminent release dates accessed the intervention. As such the intervention was delivered bi-weekly rather than weekly in order to accommodate this situation.

The Survive & Thrive Prison Version involved greater contact time than might be available in the community. This was as a result of initial facilitator introductions, increased availability of contact within the prison as well as treatment 'exit' interviews where facilitators reviewed the experience of the intervention with participants and collaboratively constructed further treatment plans and options.

Participants were provided with a booklet which was adapted where appropriate for use in a prison setting. Each chapter of the *Participant Booklet* mirrored the 10 sessions as described above.

The trial consisted of 9 Survive & Thrive intervention groups, each consisting of between 4 and 10 participants.

3.3. Research questions

The efficacy of Survive & Thrive was investigated with respect to a female prison population and based on an intention-to-treat (ITT) analysis as well as adequate dose completer (7 sessions) analysis. The specific objective of this study was to answer the research question: *Are group based psychoeducational interventions effective for the stabilisation of trauma symptomatology in a prison setting for female offenders?* With respect to Survive & Thrive (S&T), the intervention being used to investigate this study's specific research objective, the following questions were considered:

1. Will S&T be an efficacious intervention for promoting behavioural and emotional stability as associated with survivors of interpersonal trauma compared to a wait list control group (i.e. usual care) in a prison setting?
2. Will S&T be an efficacious intervention for stabilising symptoms associated with PTSD compared to a wait list control group in a prison setting?

3. Will S&T be an efficacious treatment for stabilising general symptoms of psychopathology compared to a wait list control group in a prison setting?
4. Will S&T be a more efficacious treatment for those participants who receive an 'adequate dose' compared to a waitlist control group in a prison setting?

3.3.1. Hypothesis

The central hypothesis for this trial originates from seminal works by authors such as Herman (1992) where a phased based approach to the recovery of interpersonal trauma was advocated. As such it was hypothesised that participants receiving the intervention would, compared to the control group, have an improvement in behavioural stabilisation reflecting a greater improvement in affect and symptom management. These were conceptualised as the trials *main outcomes* in preference to a focus on primary outcome measures which are explained in the following section. From the meta-analysis conducted for this thesis it was predicted that there would be a small effect size in favour of the intervention arm. Whilst the full amelioration of PTSD symptomatology is only expected from phase 2 Treatment Memory Processing (TMP) interventions improvements in general psychological distress, i.e. anxiety and depression, described as *other outcomes* (previously referred to as secondary outcomes), were also expected.

3.4. Design and rationale of the randomised control trial

This study was a single-blinded randomised control trial (RCT). The trial was registered and conducted in compliance with CONSORT (Consolidated Standards of Reporting Trials; http://www.consort-statement.org) 2010 Statement guidelines to ensure high levels of methodological integrity (Schulz, Altman & Moher, 2010). The recently released extensions to CONSORT for reporting Social and Psychological Intervention (CONSORT-SPI) trials were also consulted and amendments made to the methodology to ensure compliance with these guidelines (Grant, Mayo-Wilson, Montgomery, Macdonald, Michie, Hopewell, Moher, 2018; Montgomery, Grant, Mayo-Wilson, Macdonald, Michie, Hopewell, Moher, 2018). Central to this study design was an ITT analysis where all participants were entered into the main analysis regardless of whether they received or completed the intervention (Hollis & Campbell, 1999).

RCTs are the gold standard approach to clinical research. The random allocation of participants to either an intervention or control arm helps minimise differences between the arms so that differences at post assessment can only be explained by the intervention received. This helps ensure high levels of internal validity as well as high levels of external validity in the case of more pragmatic designs (Patsopoulos, 2011). However, as McDougal, Clarbour, Perry, & Bowle (2009) highlights this type of evaluation has infrequently been conducted in UK prisons, largely due to ethical and risk management concerns about withholding treatment from a control group as well as other practical considerations about random allocation. The RCT that McDougal et al (2009) conducted sought to provide a framework for addressing these concerns which are considered particularly relevant to large scale psychotherapy trials within forensic settings.

A pragmatic randomised control trial

Central to this study was the use of a waiting list control design where by all eligible prisoners ultimately received the intervention. However, in emulating the precedence set by McDougal et al (2009) the pragmatic design of this trial will also prioritise prisoners with an overriding need to attend the intervention on the basis of imminent release or interpersonal trauma being formulated as central to their risk management plan. These prisoners were assigned to a separate cohort group prior to the random allocation rather than being included in the RCT. The target population is described in more detail below in section 4.8 (*Sample and Selection*). As this was the first RCT conducted in the Scottish Prison Service (SPS) for any psychological intervention a pragmatic approach was regarded as having several benefits including the replication of outcomes in routine settings (Dunn, 2013; Patsopoulos, 2011). Indeed, the Medical Research Council (MRC) framework recommends that RCTs are used at an early stage in the research process for developing and testing complex interventions (Craig, Dieppe, Macintyre, Michie, Nazareth & Petticrew, 2013).

Participants in the control arm were offered the intervention after their follow up assessment. Although there were no other standardised trauma interventions being offered at the research sites, at the time of the study, participants were not excluded from any other psychotherapy or pharmacological treatments.

Trial equipoise

This trial was design to provide information regarding the efficacy of both the specific intervention used in the treatment arm (Survive & Thrive) and how responsive female prisoners, with histories of complex interpersonal trauma, are to trauma focused psychoeducation as delivered in a group modality. As reviewed in the previous chapters, there is a lack of information as to the efficacy of such interventions in comparison to usual/standard prison care. Resolving this uncertainty was an important part of this trial (Hey, Weijer, Taljaard & Kesselheim, 2018).

3.5. Trial registration

International Standard Randomised Controlled Trial Number (ISRCTN) Registry

ISRCTN35772940 https://doi.org/10.1186/ISRCTN35772940

Study Title: *Complex trauma psychoeducational intervention for female offenders*

Condition category: Mental and Behavioural Disorders

Date applied: 07/10/2013 Date assigned: 15/01/2014

3.6. Ethical approval

Ethics approval was granted by the NHS Research & Ethics Committee, Edinburgh Napier University and the Scottish Prison Services Research & Ethics Committee.

1. Scottish Prison Service, 13/06/2013

2. East of Scotland Research Ethics Service (EoSRES) REC 1, 15/11/2013, REC ref: 13/ES/0111

3. Edinburgh Napier University: January 2014

3.7. Study setting: delivery teams and locations

As this was a pragmatic trial existing staff at the study sites were trained in delivering Survive & Thrive (S&T). This included four operational members of prison staff previously trained in delivering group based treatment programmes and working in a pre-existing department dedicated to the delivery of high intensity offending behaviour interventions as well as low intensity mental health and addiction psychoeducation courses. Prison staff had a variety of experience and had been working in this capacity for between 12 months to 120 months. All prison staff had extensive previous training in CBT and group based interventions approaches and worked under supervision from the Scottish Prison Service's Psychological Service. In addition 3 forensic psychologists in training were involved in delivery as were four mental health nurses. The psychologist in training had between 48 and 60 months of experience in post and nursing staff had on average 120 months experience.

The author, i.e. Principal Investigator (PI), provided supervision to all members of staff delivering the S&T intervention and was a consistent member of staff throughout all deliveries of the intervention. The PI was a Chartered Forensic Psychologist with 180 months professional experience in delivering psychological treatments and interventions and was a Psychology Manager based at Site 1.

In order to ensure that a suitable level of mental health expertise was available for intervention delivery facilitator selection was based on a multi-disciplinary approach where at least one psychologist / trainee psychologist or mental health nurse was included at any one time. In addition, as the facilitation team consisted of both male and female members of staff a gender balance or preference towards female facilitators was ensured. All facilitators had extensive previous contact with numerous participants in addition to the pre-intervention treatment planning and preparation undertaken.

The two trial sites included:

Site 1: This national facility was the only female specific prison establishment in the SPS estate. The majority of interventions and treatment for female offending and mental health concerns

have previously been delivered in this establishment. Female offenders with difficulty coping in custody and other mental health concerns have historically been transferred to this establishment.
7 intervention groups were delivered at this trial site.

Site 2: This establishment has one residential female unit which houses approximately a quarter of the SPS women's estate. Most female prisoners within this establishment originate from that region. A number of mental health interventions for female offenders have been delivered at this establishment in recent years. There were no operational prison staff involved in the delivery at S&T at this site; facilitators were either psychologists in training or mental health nurses.
2 intervention groups were delivered at this trial site.

3.8. Standardisation and training

Initial training prior to the trial was provided by Dr S. Ferguson, NHS Lothian Consultant Clinical Psychologist the author of Survive & Thrive (2013) and NHS Lothian lead clinician for complex interpersonal trauma and the Survive & Thrive Reference Group. The PI was trained as a trainer in Survive & Thrive and then delivered a further 2 training events to ensure facilitator succession during the trial period.

Both trial sites used the same intervention materials. The SPS's Approved Interventions Panel certificated Survive & Thrive for delivery within the prison estate during this trial period further ensuring adherence to standardisation.

3.8.1. Treatment supervision

Supervision was provided during all stages of the intervention including 2 formal supervised sessions during each delivery of the intervention. All sessions were video recorded in order that supervision could be based on any session and that facilitation or security concerns could be addressed at any time. Supervision included quality

assurance monitoring which occurred either during session delivery or via video recordings.

3.9. Sample and selection

Target population and justification

Recent and past international research has indicated that there is an extremely high prevalence of interpersonal violence and abuse within female offender populations (Moloney, van den Bergh & Moller, 2009). This includes those in custody within the Scottish female prison estate (Loucks, 1997; Mahoney & Karatzias, 2012; Karatzias, Power, Woolston, Apurva, Begley et al, 2018). In addition, female offenders in prison have also demonstrated high rates of symptoms and behaviours associated with interpersonal trauma such as affect management difficulties, suicidal and non-suicidal self-injury behaviours and violent behaviours directed towards others (Howard, Karatzias, Power & Mahoney, 2017; Karatzias et al, 2017). It is therefore concluded that there will be a high prevalence of women at the trial sites presenting with a range of symptoms or symptom clusters associated with PTSD or CPTSD symptoms (Hien, Cohen, Miele, Litt & Capstick, 2004; Lynch, Heath, Mathews & Capeda, 2012).

As a result of this prevalence it was hypothesised that potentially large numbers of women at the trial sites would benefit from implementing a psychoeducational intervention where previously there had been none. However, an appropriate understanding and appreciation of which individuals might benefit from Survive & Thrive and under what circumstances was unknown. Therefore, a broad approach to including individuals with a range of presenting trauma symptoms was taken whilst also upholding relevant clinical governance and organisational concerns so as to ensure treatment was offered appropriately.

Inclusion criteria

All convicted women at the trial sites who indicated in previous assessments a history of interpersonal violence and complex trauma and who were over the age of 18 years were invited to participate in the study. It was emphasised that participation was entirely voluntarily and separate from other mandated offending behaviour interventions. Participants were also required to be able to give written consent and

able to cope with the demands of baseline screening interviews. The Self-Report Instrument for Disorders of Extreme Stress (SIDES-SR) and Trauma Antecedents Questionnaire (TAQ, van der Kolk, 2002; 2010) was used to identify the presence of complex trauma and relevant symptomatology.

Routinely administered prison mental health and offending behaviour assessment processes, such as the Generic Assessment and the Core Screen (the later undertaken at reception into prison), were utilised to ascertain which individuals may have a history of complex interpersonal trauma and may have been suitable to participate. The selection process included the following inclusion criteria:

- Willingness to voluntarily participate in the intervention for therapeutic and rehabilitative reasons and to be able to given written consent.
- An adequate length of time in prison, as per expected liberation date, that would enable randomisation into either arm of the study.
- An adequate level of mental and physical health and substance misuse stability so as to cope with the requirements of the intervention as determined by the local Multi-disciplinary Mental Health Team (MDMHT).
- Have a competent use of the English language and literacy skills. Learning disabilities or difficulties were considered on a case by case basis and discussed with the Treatment Manager.
- Located in a mainstream residential location where risk of violence to self and others was within regular security parameters.

Exclusion criteria

It was intended that as few women as possible would be excluded from the study. Therefore, the following exclusion criteria were primarily based on ensuring the safety of the individual and/or other participants and staff as well as the good order and functioning of the prison establishment. The selection process therefore sought to ensure appropriate level of psychological and mental health functioning. Exclusion was based on:

- Those who were unwilling to participate or unwilling to give written consent.

- Unable to cope with the demands of baseline interviews because of mental or physical illness.
- Those who present a high risk of institutional violence and who required segregation from mainstream residential units.
- Those who were on frequent observational procedures (see SPS's suicide and self-harm policy: *ACT 2 Care*, now replaced with *Talk to Me*) at any time in the week prior to the intervention starting were not included.
- Individuals who had enemies or who had formed intimate relationships with others in custody were not included within the randomisation process as it was important to ensure they were in separate groups.

Treatment readiness and receptiveness to intervention

There was no formal assessment of treatment readiness. However, all potential participants were discussed by the research sites Multi-Disciplinary Mental Health Team (MDMHT) at meetings prior to participation in the study which the PI or other senior member of the Psychology Department attended. In this respect, any issues that might occlude a potential participant from participating in the Survive & Thrive course were discussed.

In addition, the RA discussed the behavioural and mental health targets of the course with all potential participants to ensure their full understanding of the reasons for wanting to participate in Survive & Thrive. Trauma symptomatology, as understood from baseline measures, was also discussed by the RA and PI; the latter acting as the supervisor and Treatment Manager for the intervention during the trial period. Any potential participants who were not regarded as ready for treatment were discussed with colleagues from the MDMHT so that appropriate support and treatment could be provided.

3.9.1. Sample size calculations

Although it was anticipated that the majority of women at the trial sites would have a history of traumatic events, it was still uncertain prior to the trial what percentage of those would present with active or current trauma symptomatology. However, a conservative estimate suggested that about 50% of women in prison with a history of

significant negative life events would present with some form of traumatic symptomatology and would benefit from Survive & Thrive (Karatzias et al, 2017).

In order to establish a suitable level of statistical power a priori calculations were undertaken to ensure that the study would be able to detect possible therapeutic benefits as a result of the intervention. Considering the sample size that might be required to establish behavioural change with an adequate effect size presented several challenges. For example, McDougall et al (2009) found that existing behavioural change measures such as the Behavioural Assessment Checklist (BAC), which is based on prison staff observations, have unsatisfactory levels of inter-rater reliability. As such power computations were based on an effect size likely to reflect a clinically important difference in PTSD symptomology. Given this situation the power analysis was conducted on the PCL-C and therefore both this measure and the BAC are considered important main outcomes for this trial.

In acknowledging that only one of the two previous peer reviewed studies that evaluated Survive & Thrive (Karatzias et al, 2014; Ball et al, 2013) involved a forensic population it was concluded that this previous research might help indicate the required sample size. The Cohen's d score as reported from the Total PCL-C scores in Karatzias et al (2014) reporting a small effect size of 0.1. Similarly, Ball et al (2013) reported a medium effect size of 0.5. A midway point from these score was taken supporting the expectation for a small-medium effect (under guidelines from Cohen, 1988). This was converted to an F score of 0.18 to enable the effect size to be calculated using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007). The sample size for this study was based on a Repeated Measures Analysis of Variance (RMANOVA) statistical tests, within-between interactions between 2 groups (experimental and control group). This indicated that a total sample of 70 participants would be needed to detect the expected effect size with power set at 0.95 (1 - β) and α = 05. The correlation between scores at any one level of the repeated factor was set at 0.6 based on 3 measurements.

Intra-class correlation and variance inflation factor

Anticipated intra-class correlational coefficients (ICC) were used to adjust the calculated sample size to account for outcome inflation by the design effect. This was

important to ensure that a reasonable estimate of statistical power required by the study was undertaken. From the existent literature anticipated ICCs of between 0.01 and 0.03 were apparent for similar interventions (Eldridge, Ashby, Feder, Rudnicka, & Ukoumunne, 2004; Murray, Catellier, Hannan, Treut, Stevens, Schmitz, Rice & Conway, 2004). A low ICC was anticipated as being the most relevant for a psychoeducational 'course' such as Survive & Thrive with higher ICC reflecting increased interactions between participants. The anticipated ICC (*p*) was used within the design effect (*DE*) formula to calculate the variance inflation factor. This formula is summarised as: *DE*= 1+ (n-1) *p* (Rutterford, Taljaard, Dixon, Copas, Eldridge, 2015) and resulted in an increase in the standard sample size (i.e. n=70) of between n=118 (*p*= 0.01) to n=145 (*p*= 0.03).

Recruitment from the trial sites for prisoner flow throughout the entire study period (2 years) was based on the following population forecast:

- 200[1] women meet the initial criteria as based on the pre-screening and sentencing demographics.
- 80% (N = 160) women will qualify based on history of repeat interpersonal trauma and associated problematic behaviours as associated with complex interpersonal trauma as established by baseline measures.
- 90% (N = 144) of women with a relevant trauma profile will be eligible for randomisation to study arms; 10% attrition based on various organisational and/or risk circumstances.
- 70% (N = 101) consent to participate in the study and randomised to either study arm (i.e. S&T vs. TAU/WL); 30% (N=43) attrition based on refusal to consent and participate in the research.
- 70% (N = 70) complete all study assessments from pre to post time points; 30% attrition (N= 30) based on various individual circumstances and refusal.

[1] This estimate was based on the available SPS population data in May 2013 for convicted female prisoners at the 2 trial sites. However, this estimate does not take into account sentence length (i.e. 6 months or over) and the flow of newly sentenced and released prisoners.

As a result of the previous prison based pilot delivery of Survive & Thrive it was anticipated that a 25-30% drop out rate was to be expected. Therefore, it was concluded that if the overall sample was to address the anticipated variance inflation factor (VIF), the trial would need to ensure Survive & Thrive was extensively delivered in the trial sites in order to ensure sufficient participants could adequately demonstrate statistical power.

3.10. Control comparator

Crucial to an RCT design, is investigating an intervention's relative efficacy to a control condition therefore preventing threats to a trial's internal validity (Freedland et al, 2011). These threats were described as including influencing outcome expectancies or changing participant's recovery-promoting behaviours, contamination and differential drop out by the control group (Mohr, Spring, Freedland, Beckner, Arean, Hollon, Ockene & Kaplan 2009; Steins Bisschop, Courneya, Velthuis, Monninkhof, Jones, Friedenreich van der Wall, Peeters & May, 2015). The study design utilised a TAU/WL comparator group and it was assumed that randomisation should ensure that existing psychotherapeutic and pharmacological treatments (i.e. 'usual care') would not present as a sophisticated trauma responsive or informed active comparator where dose or intervention superiority would be important influencing factors in respect of the studies internal validity (Stice, Burton, Bearman, & Rohde, 2007).

Due to ethical considerations and the multi-disciplinary and multi-agency working practices at the trial sites it was not possible to control the various personnel and therapeutic resources available to participants in both the control and the intervention arm. As such the trial compared the intervention group with usual care, i.e. treatment as usual (TAU), within the trial sites. Usual care consisted of prison coordinated treatments accessed by participants in both arms according to individual need and consisted of psychiatric and mental health nursing and addiction services. Usual care also consisted of a non-controlled provision by various third sector agencies involving various counselling and bereavement services however none offered a standardised psychoeducational intervention for complex interpersonal trauma. Similarly, these voluntary agencies did not disclose who they were working with and they were not systematically co-ordinated by prison management or mental

health services. It was therefore not possible to collect detailed information about the treatments that participants were accessing or indeed their relative efficacy. These interventions did not constitute an active standardised treatment condition arm and an assumption was therefore made that randomisation would ensure that effects of standard practices and treatment would be accounted for. There is however considerable debate about the difficulties in fully accounting for the efficacy of usual care and the impact this has on RCTs as well as other biases that can influence trial outcomes (Freedland et al, 2011; McCambridge, Kypri & Elbourne, 2014; Mohr, Ho, Hart, Baron, Berendsen et al 2014).

3.11. Outcome measurements

3.11.1. Demographics and sentencing characteristics

A questionnaire was designed to collect data on the demographic, forensic and mental health profiles of participants. This included age, ethnicity as well as other personal and family circumstance such as previous employment experiences and childcare. In addition, basic information about index offences, previous convictions and involvement in other mental health interventions were also collated.

3.11.2. Baseline trauma profile measures

The treatment aims of Survive & Thrive are to provide safety, stabilization, and affect management skills to individuals who have experienced interpersonal trauma (Karatzias et al, 2014). It was therefore important that baseline measures assessed experiences of complex psychological trauma and relevant symptomatology (Herman, 1992; Courtois & Ford, 2013, pg. 46). Symptom chronicity and complexity were of interest at baseline for trial selection (McLean & Gallop, 2003). The following measures were used to help identify appropriate participants for the trial.

Self-Report Instrument for Disorders of Extreme Stress (SIDES-SR: van der Kolk, 2002)

A 45-item measure to assess the presence and/or severity of complex trauma and has behavioural anchors which have demonstrated good internal reliability. The SIDES-SR consists of 6 major scales with related subscales. The major scales include: (1) alteration in regulation of affect and impulses, (2) alterations in attention or

consciousness; (3) alterations in self-perception; (4) alterations in relations with other; (5) somatization, (6) alterations in systems of meaning. There are two scores obtained for each symptom item: lifetime presence (rated as a yes/no dichotomy) and current presence during the past month, also rated on a 4 point scale according to severity. Items which are scored 1, indicate symptom experience at non-pathological levels (for example 'feels quite angry but able to shift to other matters'). Items scored as 1 are combined with scores where participants indicate no difficulties in a particular area (i.e. score 0).

At the time of data collection the SIDES was the only existent clinical measure for the assessment of complex PTSD (Palic, Zerach, Shevlin, Zeligman, Elklit & Solomon, 2016). Whilst the SIDES-SR has demonstrated good psychometric properties in a number of different populations the construct validity of its subscales are less certain (e.g., Dorhay et al, 2009; Ford & Kidd, 1998; Ford, Stockton, Kaltman & Green 2006; Pelcovitz, van der Kolk, Roth, Mandel, Kaplan & Resick, 1997). Luxenberg et al (2001 pg 381) cite an unpublished study by Spinazzola, Blaustein, van der Kolk (2001) and conclude that there are 'acceptable to high rates of internal consistency' for both the full scale (Cronbach alpha = .93) and five subscales (α = .74 to .82) indicating that the SIDES-SR, 'can be reliably interpreted in a continuous fashion'. However, the Somatisation scale had lower observed levels of internal consistency (α = .68) which the authors conclude the scores on this scale should be interpreted with caution. In addition, other studies have found that subscales of the SIDES correlate highly with measures of other constructs such as alterations in the regulation of affect and dissociation (Zlotnick & Pearlstein, 1997).

Given some of the uncertainty surrounding the lack of validity with the SIDES-SR, particularly for use with offender populations, the baseline assessments were used only to indicate levels of complex trauma and relevant symptomatology within the study population.

Traumatic Antecedents Questionnaire (TAQ, van der Kolk; 2002)

The TAQ seeks to assess an individual's 'trauma load' as a result of traumatic experiences incurred over the life course. The TAQ gathers information on the frequency (never (0), rarely (1), occasionally/moderately (2) often/commonly (3) and

don't know) across 11 domains. The TAQ includes both positive and negative experiences, the former including competence and safety. Negative experiences include neglect, separation, emotional, physical, sexual abuse, witnessing, other trauma, alcohol and drugs (witnessing and use). Four developmental periods are separately assessed including early childhood (ages 0–6 years), middle childhood (7–12 years), adolescence (13–18), and adulthood (19 years plus). Scores for never and rarely are combined and scored as 0.

Previous studies have indicated the TAQ to have high test-retest reliability. For example, Garieballa, Schauer, Neuner, Saleptsi, Kluttig, Elbert, Hoffmann & Rockstroh (2006) presented scores across the 11 domains for a small sample (n=31) of German and Sudanese forensic patients and indicated high test-retest reliability across the life span and for more severe events (emotional, physical abuse, other traumas, alcohol and drug abuse: $r = .85–.88$, $p < .05$). The highest reliability was found for the witnessing trauma domain ($r = .95$, $p < .01$). In contrast, reliability of reports of sexual abuse and neglect during the early developmental periods failed to reach significance ($r = .32$ to $.77$).

3.11.3. Main outcome measures (MOM): behavioural change and PTSD symptom stabilisation

The main outcome measures, sought to investigate changes in behavioural stabilisation as well as management of symptomatology associated with PTSD. Pre-treatment assessment also provided an accurate profile of behavioural and symptom stability both at an individual and cohort level. The main outcome measure of most interest was hypothesised to be behavioural stabilisation linked to PTSD symptom management. There is, however, little research concerning the observational assessment of behavioural change and self-regulation linked to PTSD symptom amelioration in either the offender or complex trauma literature. This is discussed in the following sections.

The internal consistency and cut off scores for all outcome measures as used in relevant statistical analyses are noted within the Appendix 2.1.

MOM 1.: Behavioural Assessment Checklist-Revised (BAC-R: as revised for this trial in 2013 from Nugent, Geohagan, & Travers, 2006)

A staff/observer rated measures was chosen due to its potential to be an objective assessment procedure to rigorously test behavioural stabilisation outcomes. In addition, it was concluded that a staff rated measure would address any concerns related to self-deception and impression management prevalent in forensic settings (Cima, 2003; Merckelbach, & Collaris, 2012; Rogers, 2018).

The Behavioural Assessment Checklist (BAC) originated from the Prison Behaviour Rating Scales (Cooke, 1996) and the main purpose of this measure is to consider whether skills learnt within a taught environment are transferred to a participant's day to day life. The BAC contains 54 items and is completed by a member of staff who knows the individual well, for example, a Personal Officer (i.e. key worker). The six subscales that constitute the BAC include: *Belligerence*, *Withdrawal*, *Distress*, *Impulsivity*, *Ego-centricity* and *Problem Solving* and are usually scored on a 3 item scoring procedure: *Never, Sometimes, Always* (Cooke, 1998; McDougall, Clark, & Woodward, 1995). Whilst the BAC was originally designed to measure change with respect to prisoners who had participated in offending behaviour programmes it was concluded that the subscales such as withdrawal and distress were also relevant to interventions seeking to stabilise symptoms associated with interpersonal trauma and that the other subscales would also reflect improved psychological functioning.

Test–retest reliability of this measure over one month ranged between r= 0.46 for withdrawal and r= 0.56 for egocentricity and internal consistency varied between r= 0.64 for withdrawal and r= 0.88 for belligerence and impulsivity (Nugent, Geohagan, & Travers, 2006 cited in Draycott, Kirkpatrick & Askari, 2012). It was noted that, McDougal et al (2009) in their RCT of the Enhanced Thinking Skills intervention found that the BAC had poor to moderate inter-rater reliability and questioned its value and use in their analysis. Indeed, McDougal et al (2009) documents the difficulties that have largely lead to the discontinuation of staff rated observational measures particularly as forms were not being completed conscientiously, especially when repeated over long periods, and a lack of inter-rater reliability. However, McDougal et al (2009 pg. 14 and pg. 50) also concluded that it could not determine whether this

was because of the construction and integrity of the BAC or the circumstances in which the measure was used.

Revisions made to the Behavioural Assessment Checklist

Whilst it was recognised that there were problems with the BAC's psychometric properties there were few other suitable observer rated measures available. Therefore, adjustments were made which it was hypothesised would increase the measure's sensitivity for recording change in participant's behaviour. This included changing from the usual 3 point to a 5 point Likert scale. The new 5 point scale, known as the BAC-R, was accompanied by the follow descriptors for each response on the scale: *Never, Rarely, Sometimes, Often, Always*. These changes required the internal consistence of the adjusted measure to be assessed; the Cronbach's alpha for these computations can be accessed in the Results section.

Given the above reported difficulties in administering the BAC it was also imperative that any administration was accompanied by suitable staff awareness training. Guidance notes also accompanied the distribution of relevant forms and the RA was active in supporting staff completing this measure.

MOM 2.: PTSD Checklist Civilian Version (PCL-C: Blanchard et al., 1996)

The PCL-C is a 17-item self-report measure originally designed with reference to PTSD symptoms as defined by the DSM-IV. The symptoms endorsed may not be specific to just one event, which can be helpful when assessing survivors who have symptoms due to multiple events. In this respect the PCL-C focuses on ascertaining the presence of symptoms in terms of generic 'stressful experiences' that may be evident in any population.

Participants respond on a 5 point scale, ranging from 'not at all' to 'extremely' indicating how frequently the specific symptom was present in the participant's life over the past month. The PCL-C can be scored by providing a total symptom severity score (range = 17-85) with higher scores indicating greater symptom severity. The subscales in this measure are: *Intrusion* (re-experiencing), *Avoidance* (numbing) and *Arousal* (hyper-arousal). The USA VA National Centre for PTSD (2012) guidelines for 'speciality mental health clinics' where the estimated prevalence rate of PTSD is 40%

or above suggest a PCL-C score cut-point of 45-50. Monson, Gradus, Young-Xu, Schnurr, Price, & Schumm, (2008) suggest that 5 point change is a minimum for determining whether an individual has responded to treatment and 10 point change is a minimum for clinically meaningful improvement.

3.11.4. Other outcome measures (OOM): alleviation of psychological distress

OOM 1.: Difficulties in Emotional Regulation Scale (DERS; Gratz & Roemer, 2004).

The DERS is a 36-item self-report measure that has been developed to be a comprehensive measure emotional regulation. With good internal consistency it seeks to measure six dimensions of emotional regulation: (a) *Non-Acceptance* of emotional responses; (b) *Goal* directed behaviour difficulties; (c) *Impulse* control difficulties; (d) *Awareness* of emotional difficulties; (e) *Strategies*, limited access to those involved emotion regulation; and (f) *Clarity*, lack of emotional clarity. All DERS subscales have been demonstrated to be moderately to strongly correlated, showed good internal consistency adequate test-retest reliability for a period of 4-8 weeks (Gratz & Roemer, 2004). There is no official cut-off score. The DERS has excellent internal consistency and good construct validity (Fowler, Carak, Elhai, Allen, Frueh & Oldham, 2014).

OOM 2.: Dissociative Experiences Scale (DES II, Bernstein & Putnam,1986; Carlson & Putnam, 1993)

This 28-item measure assesses the frequency of dissociative experiences. Participants rate, in increments of 10% to 100% of the time, how frequently dissociative experiences happen to them; excluding those that happen under the influence of substances. The total DES score is the mean of the all 28 items with mean total scores also being calculated for each of the three subscales: *Depersonalisation* (including derealisation); *Amnestic* dissociation and *Absorption* (including imaginative involvement). The DES II is designed to be used as a screening measure for dissociative disorders and to help determine the contribution of dissociation to a participant's mental health difficulties and behaviours.

Reliability and validity of the DES has been well established including in offender populations (Mazzotti, Farina, Imperatori, Pruetti, Speranza & Barbaranelli, 2016; Ruiz, Poythress, Lilienfeld & Douglas, 2008). The DES measures both normal and

pathological levels of dissociation. However, criticisms of the DES have questioned whether its clinical utility is somewhat poorly defined and that it is more suited as a screening measuring particularly of non-pathological forms of dissociation (Olsen, Clapp, Parra & Beck, 2013; Ruiz, et al (2008).

OOM 3.: Hospital Anxiety and Depression Scale (HADS: Zigmond & Snaith, 1983; Snaith, 2003)

This brief, 14 item, self-rating measure is frequently used to screen levels of anxiety and depression. Half of the items relate to anxiety symptoms and half to depressive symptoms making two distinct subscales the *Anxiety* and *Depression*. Each item is coded from 0 to 3 and therefore the scores for each subscale can vary from 0 to 21; with increased scores indicating the severity of the symptom.

Since its original publication the HADS has been widely used and researched. A meta confirmatory factor analysis was conducted by Norton, Cosco, Doyle, Done & Sacker (2013) who concluded that general distress factor explained 73% of the covariance between items, with the (autonomic) anxiety and (anhedonic) depression factors explaining 11% and 16%, respectively. As such they recommended that it was used as a measure of general distress as there was not a good separation between anxiety and depression symptoms. This is obviously an important consideration with the interpretation of outcome data.

3.11.5. Theoretical outcome measures (TOM): rehabilitation of criminal thinking styles

Whilst the trial intervention was not designed to address offending behaviour it was hypothesised that trauma based experiences might mediate various criminogenic thinking styles (Howard et al, 2017a). It was, therefore considered important to include a forensic measure that might help account for any relevant changes as a result of participating in the trial.

TOM 1.: The Criminogenic Cognitions Scale (CCS; Tangney, Meyer, Furukawa, & Cosby, 2002)

This 25-item self-report measure is designed to assess five dimensions: (a) notions of *Entitlement*; (b) failure to accept *Responsibility*; (c) *Short-term* orientation; (d)

Insensitivity to the impact of crime; and (e) negative attitudes toward *Authority*. Items are rated on a 4-point scale that range from 'strongly disagree' to 'strongly agree'. Items are averaged across each of the five dimensions as well as a CCS total score.

Tangney, Stuewig, Furukawa, Kopelovich, Meyer & Cosby (2012) investigated the psychometric properties of the CCS with 552 prisoners and the reliability, validity and predictive utility of the measure were supported. CCS scores were also linked to criminal justice system involvement, self-report measures of aggression, impulsivity, and lack of empathy, violent criminal history, antisocial personality, as well as clinicians' ratings of risk for future violence and psychopathy. Similarly, individual's criminogenic thinking at admission to prison also predicted subsequent misconduct reports during imprisonment.

Tangney et al (2012) reported that the CCS total score was reliable with a Cronbach alpha of .81. Whilst the internal consistencies for the dimensions were lower, the authors reasoned that they were reliable given the number of items (i.e. 5 items) in each subscale. The authors also reported that the inter-correlations among the domains were small to moderate, indicating that they tap distinct constructs. CCS scores were modestly negatively correlated with age and with positive impression management reflecting the deviant nature of these cognitions.

Study adaption of the CCS

In correspondence with the author of the CCS (Tangney 1/15/2013) prior to the start of the trial a further set of 8 items were made available (by the author) which specifically introduced a new subscale not included in the original published and psychometrically validated version. This new subscale concerned cognitions associated with *Reparation* (in essence to assuage feelings of guilt, for example, *I owe something to those hurt by my criminal actions*). The items for this subscale were provided separately from the original version and were therefore randomly integrated amongst the existing items to form a new updated version of the CCS for this trial. There was however no valid psychometric data for these additional items or for the new adapted version CCS. Therefore, the internal validity of the new adapted version was tested using Cronbach's alpha, the results of which can be found in the next chapter.

3.11.6. Administration schedule

The planned administration of the psychometrics proposed for use in this study is outlined in Table 8. A further summary of the measures, subscales and psychometric properties can be found in the Appendix 2.1.

Table 8. Administration schedule for measures included in the RCT

Measures (No. Items)	**Baseline**	**Pre** (0 weeks)	**Post** (5 weeks)	**Follow Up 1** (9 weeks)
Demographics	✓			
SIDES-SR (42)	✓			
TAQ (41)	✓			
BAC-R (54)[1]		✓	✓	✓
PCL-C (17)		✓	✓	✓
DERS (36)		✓	✓	✓
DES II (28)		✓	✓	✓
HADS (14)		✓	✓	✓
CCS (25)		✓	✓	✓

[1]BAC-R is a staff completed measure of participant's behaviour all other measures are based on participants self-report.

3.12. Trial procedures and management

The procedure employed in the present study are similar to that used by McDougal et al (2009) thus ensuring the RCT design was adjusted to the needs and procedural requirements of a forensic setting. Central to this design was the use of a waiting list (WL) control design where by all eligible women ultimately received treatment. As previously discussed individuals with an overriding need to attend the intervention due to imminent early release or being considered in some other way a priority were not included in the randomisation procedure but were assigned to a separate cohort group (see Figure 4).

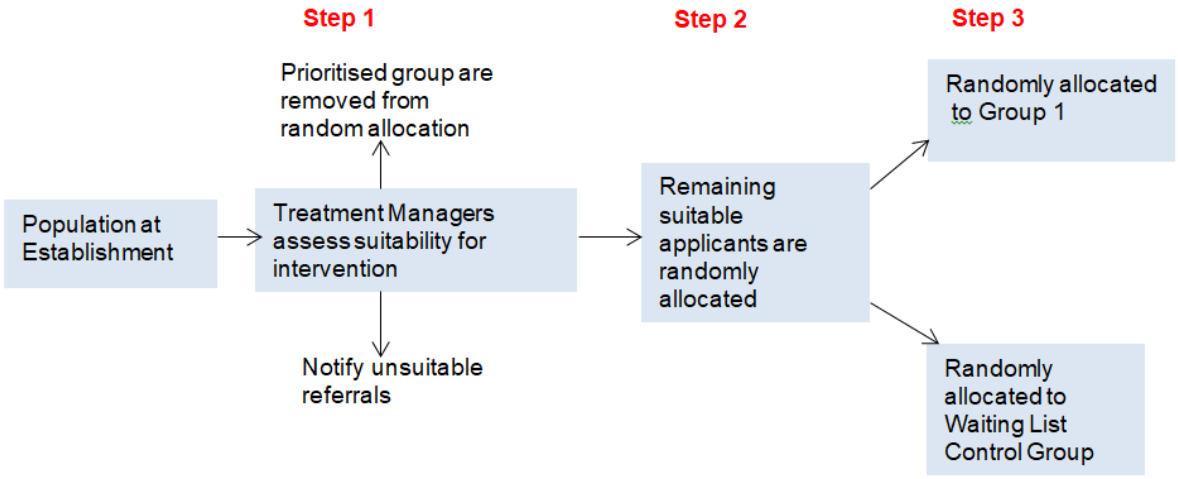


Figure 4. Study enrolment and randomisation procedure

Selected participants received an initial assessment at baseline to establish level and complexity of their interpersonal trauma (TAQ and SIDES-SR, van der Kolk, 2002; 2010) and symptomatology.

As illustrated in Figure 5 it was originally planned that participants in both arms would be assessed at 4 time points (week 0, week 5 at post treatment and follow up periods at weeks 9 and 17). However, it became apparent during the trial that most potential participants would not have been able to undertake the second planned follow up assessment due to serving relatively short sentences. The second follow up assessment was therefore discontinued.

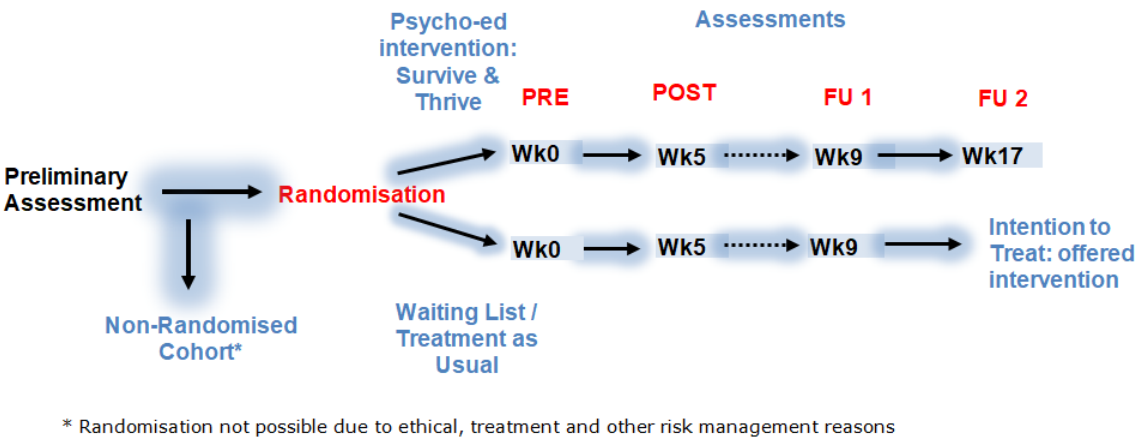


Figure 5. Planned study and design and assessment points

3.12.1. Recruitment

Existing well-established assessment procedures used in the Scottish Prison Service (SPS), *Generic Assessment* and *Core Screen*, which were already in place to refer prisoners to the Interventions and Psychology Departments were utilized and adapted to include information about Survive & Thrive. Those prisoners whose prior assessments indicated a history of childhood or adulthood trauma were initially approached by members of the prison health care and interventions teams who introduced and discussed participating in Survive & Thrive with them. In addition, posters in residential units also advertised the intervention and staff working in these units received awareness training and regular communications about the trial and the introduction of the intervention.

Following their agreement potential participants were invited to attend a one off interview with the Research Assistant (RA), and a mutually convenient time was arranged for an initial appointment during which the aims of the study were explained and any questions in relation to the study's procedure and methodology answered. A letter was then given to all suitable participants to inform them about the study (see Figure 4). Following this, participants were informed by the RA about the aims of the study and were given at least 48 hours to consider whether they wished to participate. All potential participants were given a copy of the Participant Information Sheet by the RA. This was discussed with them prior to obtaining their consent and they were informed that the study design included them being randomly assigned to either a Treatment or Waiting List group. Potential participants were also clearly informed that they could access the intervention without participating in the study and at a subsequent meeting with the RA participants confirmed their willingness to be included in the study.

Semi-structured initial interviews and baseline trauma questionnaires were conducted on an individual basis to ensure confidentiality. A senior member of the Psychology Department at the trial sites acted as Treatment Manager and screened questionnaires for the relevant interpersonal trauma histories and symptomatology and excluded those without interpersonal trauma.

3.12.2. Informed consent

Given ethical concerns about participants perceiving that they were in some way obliged to participate in the intervention consent was not obtained at the initial meeting with the RA. Potential participants were therefore given at least a further 48 hours to consider whether they wished to be involved in the study. If an individual agreed to participate a second appointment was scheduled with the RA during which inclusion and exclusion criteria were explained and a signed consent obtained. After consenting to participate and completion of baseline screening assessments individuals were randomised into either the Survive & Thrive or Waiting List/ Treatment As Usual (WL/TAU) control arm. Participation in the study was voluntary and confidential. All data was treated in accordance with the Data Protection Act (1998). Details were recorded of any prisoner's refusal to participate for research purposes and to ensure that they were not approached again by the RA.

Potential participants were also provided with verbal and written information in the form of a Participant Information Sheet. This document aimed to be accessible to individuals with a wide range of reading and writing abilities. Individuals who were approached to participate in the study were informed that their decision to participate in the study would not affect their care or progression through the prison system. They were also informed of their rights: to decline participation in the study, to withdraw from the study at any stage and to make a complaint.

Premature termination was possible by participants at any time. Participants terminating prematurely were however not included in the completer analysis unless they had undertaken 7 or more sessions. A record of attendance was kept to help ensure that completers were identified.

3.12.3. RCT management and fidelity

The use of an independent RA was regarded as integral to safeguard against awareness by the PI and staff involved in the trial of allocation and randomisation procedures. Thus, the RA helped to mitigate and guard against researcher bias thereby maintaining the empirical efficacy and purpose of the RCT methodology (Kaptchuck, 1998). The RA was also integral to the overall management and

administration of the trial ensuring recruitment and retention of participants in the trial.

Three important features of RCTs are known to impact on the quality of results: randomisation, awareness of allocation, and the choice of control group (Moher et al, 2010; Jüni, Altman & Egger, 2001). It is considered that over 50% of studies fail to adequately address these major sources of bias (Chalmers & Glasziou, 2009). These features are therefore considered and described with specific reference to the design of the current study:

Randomisation and allocation concealment

Randomisation independent of the PI was undertaken to minimise the risk of selection bias to ensure that the outcomes of the randomisation process were clearly attributed to the experimental intervention and not to an alternative explanation (Jüni et al, 2001). Randomisation and allocation concealment is also essential to the validity of subsequent statistical tests of significance (Dunn, 2013). A major source of bias undermining the internal validity of randomisation is inadequate concealment (Viera & Bangdiwala, 2007).

All available participants that met the study criteria prior to a Survive & Thrive delivery being scheduled were randomised and allocated to either the treatment or the WL/TAU control group. The former automatically began the intervention as soon as it was next delivered. The delay in attending courses was reduced by the removal of the planned 2nd follow up assessment; which meant that participants only need 4 months between random allocation and commencing of treatment. Where possible groups where balanced in sets of 8-20 according to available number of participants and the need to include prioritised non-randomised participants accounted for.

Participants were assigned a unique study code after completion of baseline measures and prior to random allocation. To ensure concealment of allocation, codes were stored electronically by the RA who used a member of the main trial site's psychology department who was not involved in the trial to produce a computer generated randomisation list of allocated participants. Whilst the randomisation procedure was planned to be administered independently of the department by a senior researcher

at Edinburgh Napier University organisational circumstances resulted in this task being reallocated. The RA was contracted by Edinburgh Napier University and all other staff involved in the trial, including the PI, were contracted by the Scottish Prison Service. The PI and the supervisory team at Edinburgh Napier University were satisfied that randomisation and appropriate levels of concealment had been undertaken at all stages of this trial. This study made use of a computer generated schedule (randomizer.org) in order to ensure participants were appropriately randomised. This process was conducted by the RA and the process was concealed from the Principle Investigator, facilitators and participants.

Awareness of trial arm allocation

Performance bias (i.e. the preferential receipt of additional treatment in one group, as well as differences in placebo response between groups) is considered to be mitigated against when participants and therapists are 'blinded'. Detection bias on the other hand is ensured against when assessors, PIs and statisticians are not aware of allocation (Juni et al., 2001). This ensures that independent judgements can be made about the outcomes. In psychotherapeutic research it is usually regarded as impossible to blind participants and the therapists as to which intervention they receive or group that they have been assigned to (Freedland, Mohr, Davidson & Schwartz, 2011). Preventing awareness of allocation for PIs and assessors in trials of psychological interventions can require substantial resources (financial, personnel, and organisational). Preventing awareness of allocation can also be difficult to maintain, for example, ensuring that participants do not disclose their allocation on assessment.

Whilst it was intended that the RA was fully blinded due to administrative and organisational constraints, which became apparent during the trial, this was not possible. The PI, intervention facilitators and other staff within the establishment were however blinded throughout the trial. Participants were for obvious reasons not blinded to group allocation and they were informed at recruitment that they were on a waiting list due to randomisation reasons. All outcome measures undertaken at baseline, pre, post and follow up were administered by two consecutively employed RAs.

Facilitators were not informed of whether they were working with participants in the intervention arm or participants in the delayed WL/TAU arm. This approach was thought to mitigate against performance bias. As such after randomisation, the RA sent the facilitators the names of participants for the next scheduled delivery of the intervention without revealing group allocation. This also ensured that the PI and other senior prison based managers were unaware of allocation. As it was not possible to blind the RA, the study should be regarded as a single-blinded randomised controlled trial with procedures in place to ensure the PI was unaware of allocation and to limit the possibility of detection bias. Unbinding of the PI only took place after the trial ended and the scoring of psychometrics completed.

3.12.4. Quality assurance / treatment integrity and fidelity

Compliance to the treatment protocol and to a trauma-informed treatment framework was ensured by evaluation procedures that monitored the efficacy and competency of delivery. Supervision was provided to the facilitators by the PI during the trial. The PI was trained by Dr S. Ferguson, Consultant Psychologist NHS Education for Scotland, the original author of Survive & Thrive, and technical expertise provided accordingly.

A random selection of sessions was monitored for quality assurance and treatment integrity purposes by Dr Ferguson. This accounted for 10% of all sessions delivered. An adapted version of the Video Monitor Form (Shine, 2003) was utilised to provide a quantitative and measureable approach to the overall quality of delivery. The form measured facilitator performance across three domains: adherence to the Survive & Thrive manual (9 items); adherence to treatment style (9 items); appropriate use of therapeutic and psychoeducation skills (6 items). Each facilitator was scored on a 5 point Likert scale and their scores amalgamated to produce an overall score for the 3 domains. Composite quality assurance scores and the corresponding variance associated with post treatment outcomes in the S&T arm were computed.

3.12.5. Data storage and protection

Participant information and outcome data were stored in a confidential manner with procedures to ensure that limits to confidentially were clearly communicated. As such

in the case where the facilitation or research team considered a participant to be at risk to themselves or others an appropriate member of staff involved in their security or care would be informed. This would be discussed with the participant prior to disclosing the information. All participants were informed of this process prior to the completion of the consent form. This procedure adheres with Section 7.1 of the British Psychological Society's *Code of Human Research Ethics* (2014). All data was stored in a locked filing cabinet in the Psychology Department at the main trial site. Only the RA had access to identifiable data during the trial. All collected data was coded with unique identifiers. Data stored on a computer for the purpose of scoring and analysis was anonymised and password protected.

3.12.6. Safety and ethical considerations

All participants in both Survive & Thrive and WL/TAU arm (and those approached but who decline to take part) continued to receive mental health input and treatment as usual, i.e. whatever treatment they had receive as part of normal routine care and treatment. Within previous research there had been no reported adverse effects of Survive & Thrive (Ball et al, 2013) which therefore suggested that Survive & Thrive was a safe and beneficial intervention for psychological trauma.

Further support, if required, was offered by the prison MDMHT (Multi-Disciplinary Mental Health Team) and the trial site's Psychology Department as per service protocol. All facilitators were acquainted with the referral routes to those services. Each individual participant was seen throughout her treatment by the same facilitators to ensure adequate levels of therapeutic rapport and trust.

In addition, all Survive & Thrive sessions were video recorded so that any delivery could be potentially supervised for treatment fidelity and security purposes.

Consideration was given to the potential that measures focusing on past trauma and abuse may result in some individuals experiencing distress. In the possible event that this happened, participants were encouraged to discuss any upsetting issues with the RA and/or clinical staff involved in their routine care. Prior to participation within the study, the RA liaised with prison staff and relied on their judgement as to whether specific individuals were too emotionally or physically unstable to participate. The

topics covered by the psychological measures were routine in research and clinical practice as well as those interventions specifically designed to help with traumatic and abusive experiences. All trauma based questionnaires were administered in a confidential environment with the RA who was a qualified mental health practitioner.

3.13. Analysis plan

Analysis plans for investigating treatment efficacy with respect to the main and secondary outcomes are reported below. After checking for the assumption of normality, parametric methods were used where appropriate to investigate the linearity of correlations and equality of variance. All statistical analysis presented in this thesis were carried out by the author. Data was analysed using Statistical Package for Social Science (SPSS) version 23.0, a full description of the statistical analysis are given in the corresponding sections below.

3.13.1. Data checking and cleaning

In order to ensure a high standard of data quality, making any subsequent analyses trustworthy, prescribed data verification and validation procedures were followed (Osborne, 2013).

Data verification: assessing whether the responses given by participants were accurately represented in the electronic data file (i.e. avoiding transcription errors). Double entry and checking by departmental colleagues at the data storage site helped prevent errors. This involved colleagues from the trial sites Psychology Department and the RA working in pairs to manually check the electronic data set against the original paper questionnaires.

Data validation: ensuring the internal consistency of data based on pre-existing knowledge about the expected characteristics of participants. This was done by the RA as a sense-check for data quality and to detect defective data. Defective data were either accounted for according to the published scoring guidelines pertaining to each psychometric or removed from the data set. Suspected defective demographic and offence related data were cross referenced by the RA with prison records which were then given preference as this was assumed to be more objective.

3.13.2. Missing data

Consideration was given to ensuring statistical procedures appropriately accounted for missing participant data at the various assessment time points. Multiple Imputation (MI) procedures were used in ITT analyses and pooled results considered along with those from the raw data. This procedure was undertaken for reporting outcome data, for example, Means (SD). In procedures where other imputation methods are conventionally selected, i.e. Full Information Maximum Likelihood (FIML), these were selected so that only one missing data procedure was used per computation (Sterne, White, Carlin, Spratt, Royston, Kenward, Wood, & Carpenter, 2009; van Ginkel et al, 2014). This was the preferred option for the analysis of treatment effect as described below using Linear Mixed Models. It was assumed that imputation methods would produce more conservative analyses and therefore only results from these approaches were considered reliable. See Appendix 2.2. for further details.

3.13.3. Primary data analysis – Intention-to-treat (ITT) analysis

A primary ITT analysis was conducted in which all randomised participants were retained in the arms to which they were allocated and no individuals removed. Thus, all participants were retained regardless of their adherence to the protocol (i.e. attendance of sessions and completion of all measures).

Means (SDs) were calculated for all continuous variables and frequencies (%) for all categorical variables. Comparisons between arms were made for 'pre', i.e. time point 1 (T1), and 'post', i.e. time point 2 (T2) outcome scores and between arms for pre and 'follow up', i.e. time point 3 (T3) scores were made by utilising repeated measures analyses. Treatment effect sizes between S&T versus WL/TAU at T2 and T3 for all outcome measures were calculated using Cohen's d formula (Cohen, 1988). As noted in the power analysis this study's design is based on establishing whether there was a significance difference between the intervention and WL/TAU arms following random allocation.

Primary analysis of treatment effect – linear mixed models (LMM)

LMM offers increased power and flexibility in outcome data processing and also computes more precise estimates of the differences between two randomised arms. LMM is considered superior to General Linear Models (GLM) where repeated measures only operate on a complete case analysis (CCA) basis. Therefore, participants who have a single missing measurement were discarded from a GLM analysis. LMM allows all data to be used (if data meets missing-at-random criteria). Similarly, LMM is a more appropriate statistical procedure where assessment time points are unevenly spaced and as T3 assessment occurred after the intervention (Seltman, 2018). Another advantage of using LMM is that it can account for the proportion that is explained by the grouping structure or intra-class correlation (ICC) which is an important concern when evaluating group treatments (Heck, Thomas & Tabata, 2014 pg. 8).

The variables of time, treatment arm and the interaction of time x treatment arm were included within the LMM analysis as fixed effects and considered within a hierarchy of levels, with the interaction being a nested term (Seltman, 2018). Model selection was also determined on the basis of being the most parsimonious (fewest variables) as well as explaining the most variance; and which had the best model fit (see below). LMM also provides an option for simultaneous analyses of experimental effects and associated individual differences which were included as random effects (Matuschek, Kligl, Vasishth, Baayen & Bates, 2017). The dependent variable in each computation was the total or subscale score from the outcome measures across all three time points, as coded within a 'long' data format. The estimates of fixed effects and confidence intervals (CI) were used to help determine the difference between intervention and WL/TAU control arm at follow-up, adjusted for pre-treatment. Results demonstrating the 95% CI to cross the point of no effect and/or a value representing a clinically relevant effect were also used to understand the efficacy of the Survive & Thrive as compared to the WL/TAU arm.

Exploratory testing was undertaken to confirm whether a linear or non-linear trajectory was suited to the outcome data. Data collection time points were coded according to a number of different non-linear trajectories when investigating the best

model fit. Individual participant trajectories were graphed to show a visual representation of reported symptom change in the measures used (see Appendix 3.5.1.). This showed that a non-linear approach accounted for idiosyncratic outcomes within the results but also resulted in an increasing distortion of the time trajectory (i.e. non-monotonic) associated with the grand mean. The construction of the different time coding models were accounted for by observing the Akaike's Information Criterion (AIC) score as an indication of model fit. The lower an AIC score is, the better 'fit' of a particular statistical model. Different models were constructed and the changing values of the AIC score were noted as indicative of optimal model construction and fit. Whilst not every subscale indicated that a linear time coding was optimal, a standardised approach was taken for consistency across all outcomes.

As a result a conservative approach using a linear scale of T1=0, T2=1 and T3=2 was utilised to assess change across the time points as non-linearity was not deemed appropriate to warrant model manipulation. An alternative approach to examining change over time through utilising the repeated measures as separate outcomes to determine the difference between the arms at 'post' (i.e. T2) and 'follow-up' (i.e. T3). This model specified time as a categorical variable rather than assuming a polynomial growth curve over a single time sequence (see Heck, et al, 2014, pg. 237). As time was treated as covariates within the model separate estimates were computed for both T2 and T3 (see Appendix 2.3.2. for further details).

A Diagonal covariance matrix used for repeated measurement for the estimate of fixed effects and a Variance Components covariance matrix was used for the random effects within the LMM used. Fixed effects being defined as the variables of primary interest that are fixed in time (e.g. treatment site or age) and random effects being defined as those derived from a larger 'random' set of variables (Seltman, 2015). Individual participant effects being almost always regarded as random effects. Exploratory procedures were undertaken where the AIC scores of several different covariance matrix types were compared. Selection of the Diagonal covariance matrix also concurs with the model used in Heck et al (2014) for their model comparing outcomes from a randomised control trial. See Appendix 2.3. for further details on the model design and selection.

3.13.4. Secondary data analysis – adequate dose (as per completer protocol)

A pre-planned secondary analysis was also conducted using a 'per-protocol' approach and involved participants who attended at least 70% of scheduled sessions (i.e. ≥7 sessions) defined by the clinical team as Adequate Does (AD) 'completers'. Participants who attended <7 of the 10 sessions attended were referred to as 'non completers'. This analysis was performed for outcome scores and included only completers in order to test whether adherence to the treatment programme (as indicated by attendance) had any greater impact on these outcomes.

The same analyses as used in the ITT analysis were used in the AD analysis.

3.13.5. Reliable and clinically significant change analysis

To further investigate individual change *post hoc* reliable change index/clinically significant change (RCI/CSC) analyses were undertaken with AD treatment participants and those assigned to the WL/TAU arm. This methodological approach sought to identify whether individuals in either arm had made a big enough change for it to be considered important or clinically significant. Data from T1 and T2 were used in this analysis as this was considered to be where the impact of the intervention would be most robust.

The criteria specified by Jacobson, Roberts, Berns & McGlinchey (1999) on which to determine the CSC analysis varied for each outcome measure depending upon available clinical and non-clinical (comparison) normative data. The criteria as established by Jacobson & Truax (1991) enabled statistically validated cut off to be applied to the level of scores which was applied to T2 scores for all outcome measures according to the following criteria:

a. a reduction in scores in the range of the comparison reference group of more than 1.96 standard deviations
b. a reduction in scores in the range of the comparison reference group within 1.96 standard deviations of the mean of the reference group
c. a reduction is closer to the mean of the comparison group than the clinical group.

Selection of the above criteria depended on the norms that were available. Where data for a comparison group was available criteria b or c were chosen. It the scores from the groups overlapped, then c was chosen. When they did not overlap then b was chosen. If norms for a comparison non-clinical group were not available then criterion a was the only criterion that could be chosen (Morley & Dowzer, 2014 pg. 3).

Data from externally validated criteria from published and validated cut off scores can also be applied as pre-existing cut off points rather than those established by Jacobson & Traux (1991) see Morley & Dowzer (2014 pg. 3).

Published normative and clinical psychometric data as well as clinically relevant cut off scores were established for each outcome measure. These were used to meet the requirements for calculating the RCI/CSC as undertaken using the Leeds Reliable Change Indicator (Morley & Dowzer, 2014) a Microsoft Excel based application. The psychometric data for each outcome measure is presented in the Appendix 2.1.

3.14. Summary

The methodology as described in this chapter seeks to fulfil the 'gold standard' guidance for RCT administration as articulated in the CONSORT-SPI protocols (Grant et al, 2018; Montgomery et al, 2018. As such it describes the randomisation and awareness of allocation procedures as well as relevant issues of implementing an RCT in a prison setting. To our knowledge, this is the first RCT to be attempted in the Scottish Prison Service and indeed the first trial of a complex trauma intervention in any UK prison. It is also the first RCT utilising an inclusive psychoeducational trauma informed protocol and ITT analysis including follow up assessments with female prisoners.

The relative high intensity delivery of the intervention, i.e. delivering the 10 sessions over 5 weeks as opposed to 10 weeks, was designed to enable a maximum level of inclusivity to a relatively transient short term prison population. Other pragmatic issues concerning the RCT design have also been discussed in this chapter which should help to ensure replicability and generalizability of findings.

As this study furthered the scientific enquiry into the potential behavioural outcomes achieved from Survive & Thrive, and more generally from psychoeducational

interventions, measurement issues had to be considered. The inclusion of the BAC-R, a staff rated measure, originally designed to ensure objectivity within a battery of otherwise self-report measures. Adaptations were made to form the BAC-R which sought to minimise previously documented difficulties with this measure (McDougal et al, 2009).

Treatment fidelity, quality assurance and supervision issues were all considered as an integral aspect of this study. The statistical investigation and modelling used to consider the research questions have also been described. As Gambl, Krishan, Stocken et al (2017) note the transparency of such statistical procedures are essential for clinical trials given concerns over given research reproducibility. Therefore, a more detailed summary of the statistical procedures used, particularly for the LMM analysis, are contained within the Appendix should be referred to 2.3.

4. Results: Randomised control trial

4.1. Introduction

The primary objective of this thesis is to contribute towards answering the wider research question: *Are group based psychoeducational interventions effective for the stabilisation of trauma symptomatology in a prison setting for female offenders?* With respect to the randomised control trial (RCT) and the analysis of outcomes in this chapter the efficacy of Survive & Thrive (S&T), a prototypical psychoeducational intervention, was considered. The following research questions were therefore deemed relevant:

1. Will S&T be an efficacious intervention for promoting behavioural and emotional stability as associated with survivors of interpersonal trauma compared to a wait list control group (i.e. usual care) in a prison setting?
2. Will S&T be an efficacious intervention for stabilising symptoms associated with PTSD compared to a wait list control group in a prison setting?
3. Will S&T be an efficacious treatment for stabilising general symptoms of psychopathology compared to a wait list control group in a prison setting?
4. Will S&T be a more efficacious treatment for those participants who receive an 'adequate dose' compared to a waitlist control group in a prison setting?

The main outcome measures used in this trial were considered specifically relevant to demonstrating the stabilisation of trauma symptomatology. These included the PTSD Checklist- Civilian Version (PCL-C) and the Behavioural Assessment Checklist-Revised (BAC-R).

4.2. Study participants

A total of 139 women prisoners were referred to the study and were eligible for assessment. Eligibility was a two stage process. The first stage involved women consenting to be referred to the study after indicating, though prison-based generic assessment procedures, a history of interpersonal trauma. Some of the women referred to the study, n=17 (12.2%), declined to participate in subsequent baseline

assessments and others withdrew after completing baseline assessments, n=18 (13.0%). In addition, n=15 (10.8%) women were deemed by joint prison and research treatment management procedures as requiring immediate assistance due to difficulties with emotional distress or imminent release from custody and were prioritised as a separate non-randomised cohort and therefore were not eligible to be part of the trial. 86 women (61.9%) of the women who had been referred agreed to participate in baseline assessments and were randomised to either the wait list / treatment as usual (WL/TAU) control arm or the S&T arm in a 1:1 allocation.

The CONSORT diagram in Figure 6 details participant flow and trial involvement. Due to prison based procedures as well as refusal after randomisation there was a considerable degree of attrition (19.1%) involving participants in the WL/TAU control arm. This was due to trial administration procedures and the short delay between randomisation, after consent and administration of baseline trauma screening measures, and 'pre' assessment (T1) at the beginning of the intervention. As noted this resulted in WL/TAU control participants being particularly vulnerable to attrition. Further participant loss was also experienced at the 'post' (T2) and 'follow up' (T3) assessments.

4.2.1. Demographic and sentencing characteristics

Participant characteristics are presented in Table 9. Participants considered themselves to be predominantly of white ethnicity, Scottish, and not in a relationship and most reported that they had children 61.9% (WL/TAU arm) and 63.9% (S&T arm). However, relatively few mothers had their children placed in care indicating the use of family support structures given their custodial status.

The mean age that participants left education highlighted that most left school prior to the statutory leaving age and did not obtain further qualifications. In addition, only 64.3% and 63.6% of the WL/TAU and the S&T arms respectively, indicated that at some point in their lives had they been in formal employment. These demographic markers are indicators of social and economic marginalisation. Histories of self-harm and current prescription of psychotropic medication were also common amongst participants, as indicated in Table 9.

Most participants were undertaking sentences for violent offending (71.4% and 65.9%, WL/TAU and S&T arms respectively). In addition, participants in both arms reported histories of previous offending including violent offending. The median sentence length for the WL/TAU arm was 44.5 months (IQR= 102.3) and 25.5 months (IQR= 24.0) for the S&T arm. The difference in sentence length between the two arms was partly explained by an additional 3.5% of Murder/Culpable Homicide category participants in the WL/TAU arm. Sentence length (in months) within that category, WL/TAU arm: Mdn= 144.0, IQR= 66.0; S&T arm: Mdn= 120.0 IQR= 45.0) as well as participants in the WL/TAU arm with Misuse of Drug Act (MDA) offences having longer sentences (in months), WL/TAU arm: Mdn= 30.5, IQR= 14.5; S&T arm: Mdn= 24.0, IQR= 14.5. Accumulatively, this indicates a bias in sentence length with the S&T arm having shorter sentences. There is only limited evidence to suggest that different sentence lengths and reoffending rates may lead to different treatment needs and outcomes for female offenders (Mahoney & Karatzias, 2012; de Vogal & Nichols, 2016).

4.2.2. Trauma characteristics at baseline

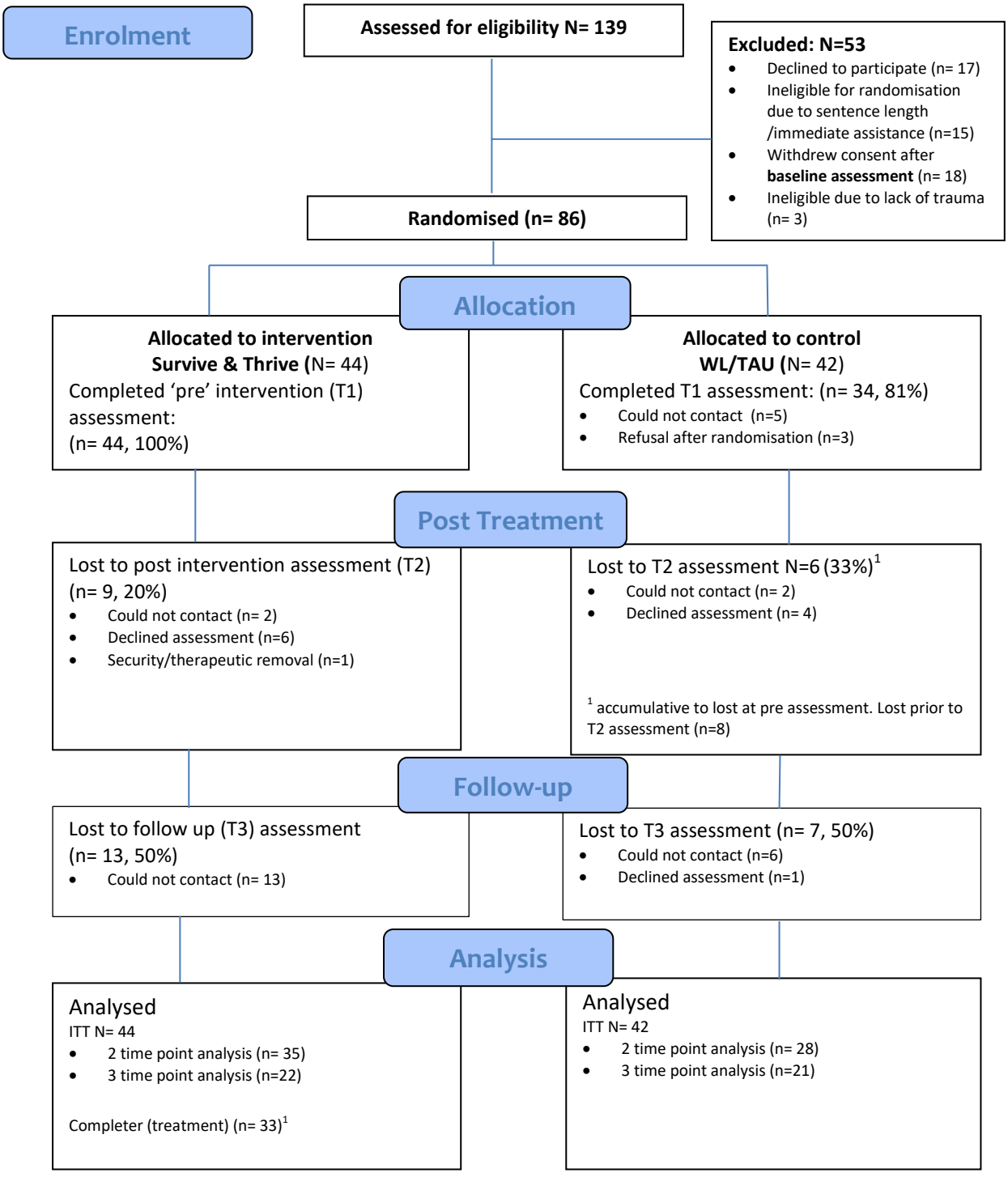
Baseline assessment data for symptoms associated with complex interpersonal trauma (SIDES-SR) and for histories of traumatic events (TAQ) are presented in Table 10 and 11. These are summary tables of symptom clusters and adverse life events, for prevalence of specific symptoms see the Appendix 3.1. The TAQ indicated that the majority of individuals (over 80.0% in both arms) had lifetime experiences of neglect, separation, emotional abuse, witnessing abuse, other traumas, drug and alcohol abuse. Lifetime experiences of physical abuse were reported by 71.4% of individuals in the WL/TAU arm and 84.1% of individuals in the S&T arm. The category of trauma least frequently reported by participants in both arms was sexual abuse. A history of sexual abuse at some point in an individual's life was reported by 64.1% in the WL/TAU arm and 54.5% in the S&T arm. On average participants in both arms more frequently reported the occurrence of sexual abuse in childhood than adulthood. The TAQ contains eight trauma categories; the mean number of categories experienced, during an individual's lifetime, was Mdn= 7.0 (IQR= 2.0) and Mdn=7.0 (IQR= 1.0) for the WL/TAU and S&T arm respectively.

A composite score across the childhood age ranges in the TAQ (0-6, 7-12, 13-18 years) is provided for each domain in Table 9. These composites scores indicate the prevalence of traumatic events (occurring in at least one age range) during childhood. It is also noted that the majority of participants reported positive experiences (i.e. Safety and Competence) during at least one point in their childhood. Appendix 3.1. provides further details for the prevalence of traumatic events in each childhood age range.

It was noted that participants indicated a greater lifetime occurrence than current presence for each of the SIDES-SR symptom domains. This was particularly noted in symptom domains such as alterations in affect regulation, consciousness, relationships and meaning. Thus, the mean scores for current severity ratings was frequently more reduced that expected. Within the SIDES-SR symptom domains the most prevalent for current difficulties was alterations in relationships (59.5% and 61.4% in the WL/TAU and S&T treatment arms respectively). High levels of lifetime difficulties in this domain were also noted. Symptoms in the somatisation domain were the least reported suggesting that these experiences were either least experienced or recognised (11.9% and 25.0% respectively).

Luxenberg, Spinazzola, van der Kolk, 2001, pg. 374) suggest that disturbances in all six SIDES-SR domains of functioning are required for their proposed 'diagnostic' criteria of complex interpersonal trauma. However, this diagnostic criterion was very rarely met and was therefore used as an indicator of trauma severity rather than as a precondition for inclusion in the trial with a focus instead being placed on the experiences (one or more) of interpersonal trauma. A graphical summary of the number of symptoms domains participants reported using the SIDES-SR diagnostic criteria in each arm is presented in Figure 7.

Overall, the demographic, sentencing and trauma characteristics indicate an ethnically uniform population of socio-economically marginalised women with poor relationship histories and current interpersonal difficulties. Whilst multiple types of trauma and relevant symptoms were reported by participants as occurring during their lifetime their current symptoms and severity of those symptoms, as self-reported, was less pronounced.



Notes: *Could not contact*= transfer to another prison or release to the community; *Completer (treatment)* = 7+ sessions.

Figure 6: CONSORT diagram

Table 9. Sociodemographic, forensic and clinical characteristics of participants

	WL/TAU (n=42) M(SD) / N(%)	S&T (n=44) M(SD) / N(%)
Age	33.17 (10.32)	33.53 (10.34)
Ethnicity		
White Scottish	38 (90.5%)	41 (93.2%)
Other	4 (9.5%)	3 (6.8%)
Relationship Status		
Single (inc. Divorced / Separated)	35 (83.3%)	31 (70.5%)
Married / In Partnership	7 (16.7%)	13 (29.5%)
Age left school (years old)	15.42 (1.03)	15.32 (1.32)
Previous employment (ever)?	27 (64.3%)	28 (63.6%)
Parent?	26 (61.9%)	28 (63.6)
How many children?[2]	3 (0-5)	2 (0-4)
Any children in care? (if a parent)	9 (34.6)	6 (21.4)
Sentence Length (months)	70.64 (65.61)	49.82 (53.43)
Index offence		
Murder / Culpable Homicide (including Attempted Murder)	14 (33.3)	11 (25.0)
Violence (including assault, fire arm, Sch 1[3], fire raising)	16 (38.1)	18 (40.9)
MDA[1]	6 (14.3)	9 (20.5)
Other offending (inc shoplifting, theft, public order offence)	6 (14.3)	6 (13.7)
No. of previous convictions	8.90 (16.97)	5.95 (13.87)
Age at first offence (years old)	23.36 (10.66)	24.63 (11.28)
Previous violence? (conviction)	31 (73.8%)	32 (72.7%)
Previous MDA?[1] (conviction)	11 (26.2%)	12 (27.3%)
History of self-harm?	22 (52.4%)	15 (34.1%)
Psychotropic medication? (current)	31 (73.8%)	26 (59.1%)

Notes: WL/TAU= Wait Listed/ Treatment as Usual control arm; S&T= Survive & Thrive experimental intervention arm. All N% presented in the affirmative (i.e. 'Yes'). As baseline measures were collected prior to randomisation statistical comparisons between the arms were not considered appropriate. 1) MDA= Misuse of Drugs Act. 2) Median and minim – maximum reported. 3) Sch 1.= Schedule 1 Offence of physical/emotional offence against a child under the Criminal Procedure (Scotland) Act 1995. There were no sexual offences included within this sample. As baseline measures were collected prior to randomisation statistical comparisons between the arms were not considered appropriate.

Table 10. Frequency and type of traumatic events at baseline

	WL/TAU (n=42) N (%)	S&T (n=44) N (%)
TAQ: Traumatic Events		
Neglect		
Child (0-18 yrs)	32 (76.2)	34 (77.3)
Adult (≥ 19 yrs)	29 (69.0)	37 (84.1)
Lifetime	33 (78.6)	41 (93.2)
Separation		
Child (0-18 yrs)	37 (88.1)	38 (86.4)
Adult (≥ 19 yrs)	38 (90.5)	37 (84.1)
Lifetime	38 (90.5)	42 (95.5)
Emotional abuse		
Child (0-18 yrs)	30 (71.4)	31 (70.5)
Adult (≥ 19 yrs)	32 (76.2)	33 (75.0)
Lifetime	34 (80.1)	36 (81.8)
Physical abuse		
Child (0-18 yrs)	26 (61.9)	30 (68.2)
Adult (≥ 19 yrs)	25 (59.5)	28 (63.6)
Lifetime	30 (71.4)	38 (86.4)
Sexual abuse		
Child (0-18 yrs)	24 (57.1)	21 (47.7)
Adult (≥ 19 yrs)	15 (35.7)	8 (13.2)
Lifetime	27 (64.1)	24 (54.5)
Witnessing		
Child (0-18 yrs)	35 (83.3)	32 (72.7)
Adult (≥ 19 yrs)	31 (73.8)	27 (61.4)
Lifetime	35 (83.3)	37 (84.1)
Other trauma		
Child (0-18 yrs)	33 (78.6)	32 (72.7)
Adult (≥ 19 yrs)	36 (85.7)	28 (63.6)
Lifetime	38 (90.5)	37 (84.1)
Alcohol / drugs		
Child (0-18 yrs)	34 (81.0)	34 (77.3)
Adult (≥ 19 yrs)	35 (83.3)	33 (75.0)
Lifetime	38(90.5)	35 (79.5)
Mean (SD) Trauma categories		
Child (0-18 yrs)	5.97(1.86)	5.74(2.16)
Adult (≥ 19 yrs)	5.70(1.70)	5.21 (1.76)
Lifetime	6.46(1.53)	6.48(1.21)
Safety		
Child (0-18 yrs)	33 (78.6)	37 (84.1)
Adult (≥ 19 yrs)	35 (83.3)	35 (79.5)
Lifetime	38 (90.5)	40 (90.1)
Competence		
Child (0-18 yrs)	33 (78.6)	37 (84.1)
Adult (≥ 19 yrs)	33 (78.6)	32 (88.1)
Lifetime	37 (88.1)	40 (91.0)

Notes: WL/TAU= Wait Listed/ Treatment as Usual control arm; S&T= Survive & Thrive experimental intervention arm. All N% presented in the affirmative (i.e. 'Yes'). TAQ: Traumatic Antecedents Questionnaire. See Appendix 3.2. for full outcome of SIDES within symptom clusters. As baseline measures were collected prior to randomisation statistical comparisons between the arms were not considered appropriate.

Table 11. Summary of lifetime occurrence, current presence and current severity of trauma symptom clusters at baseline

	WL/TAU (n=42) N (%) / M (SD)	S&T (n=44) N (%) / M (SD)
SIDES Symptom Domains		
Alterations in affect regulation		
Lifetime[1]	28 (66.7%)	32 (72.7%)
Current	22 (52.4%)	20 (45.5%)
Current Severity	.80 (.74)	.71 (.64)
Alterations in consciousness		
Lifetime	38 (90.5%)	39 (88.6%)
Current	25 (59.7%)	23 (52.3%)
Current Severity	1.02 (.96)	.97 (97)
Alterations in self-perception		
Lifetime	35 (83.3%)	40 (90.9%)
Current	18 (42.9%)	19 (43.2%)
Current Severity	.68 (.74)	. 65 (.58)
Alterations in relationships		
Lifetime	39 (92.9%)	43 (97.7%)
Current	25 (59.5%)	27 (61.4%)
Current Severity	.82 (.81)	.79 (.70)
Somatisation		
Lifetime	21 (50.0%)	15 (34.1%)
Current	5 (11.9%)	11 (25.0%)
Current Severity	.27(.48)	.36 (.51)
Alterations in meaning		
Lifetime	35 (83.3%)	37 (84.1%)
Current	13 (31.0%)	20 (45.5%)
Current Severity	.47 (.76)	.64 (59)
Total SIDES (No. of symptom domains)		
Lifetime	4.67 (1.41)	4.68 (1.03)
Current	2.26 (1.90)	2.35 (1.73)

Notes: WL/TAU= Wait Listed/ Treatment as Usual control arm; S&T= Survive & Thrive experimental intervention arm. *Lifetime*= lifetime presence; *Current*= Meets current diagnostic criterial; *Severity*= current severity ratings scale 0-3.
All N% presented in the affirmative (i.e. 'Yes'). SIDES: Structured Interview for Disorders of Extreme Stress.
See Appendix 3.2. for full outcome of SIDES within symptom clusters. As baseline measures were collected prior to randomisation statistical comparisons between the arms were not considered.

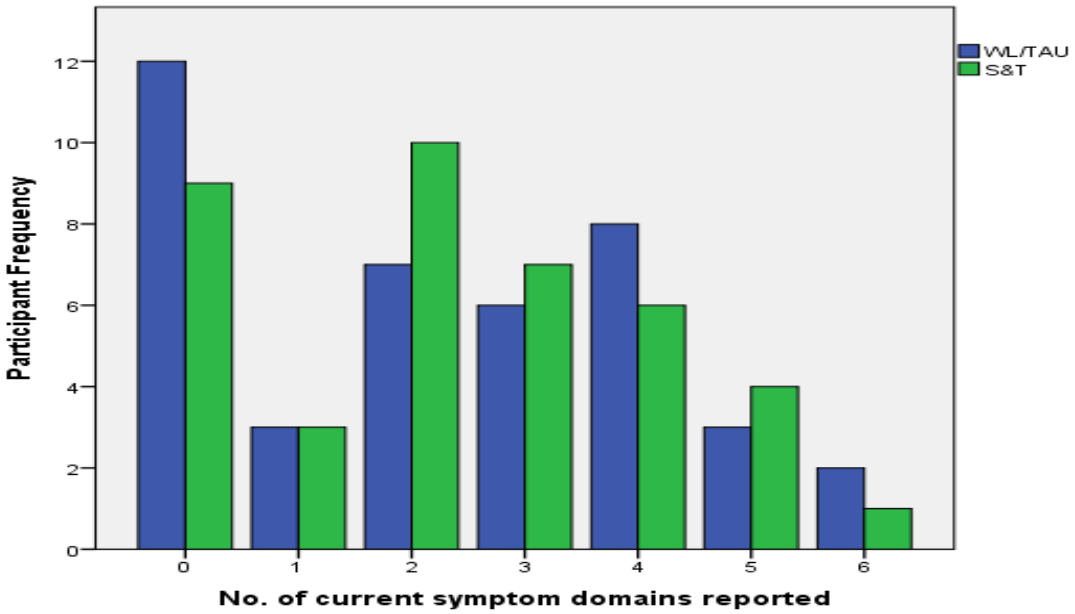


Figure 7: Frequency of participants meeting the diagnostic criteria for SIDES-SR symptom domains

4.3. Missing data analysis

Missing participant data (MPD) was defined as a missing outcome measure for a participant. The amount of MPD was most extensive for the BAC-R which was the only measure that was completed by staff rather than participant self-report. In respect of self-report measures, including the PCL-C as a main outcome measure, a total of 41 women (48%) completed all three assessments; 71 (82.5%) completed the T1 'pre' assessment and at least one follow-up assessment (T2 or T3). Further detail of the number of missing assessments is given below in Table 12.

Table 12: Missing data count

Measure	**WL / TAU** (N=42) **S&T** (N=44)	**List Count (Missing)** N (%)			**Complete Case Count (Missing)** N (%)	
		Pre (T1)	Post (T2)	FU1 (T3)	T1 - T2	T1 – T3
PCL-C	WL/TAU	8 (19.0)	13 (31.0)	20(47.6)	14 (33.3)	21 (50.0)
	S&T	1 (2.3)	9(20.5)	21 (47.7)	10(22.7)	22 (50.0)
BAC-R	WL/TAU	14(33.3)	20(47.6)	24(57.1)	20 (47.6)	26 (61.9)
	S&T	9(20.5)	17(38.6)	27(61.4)	20(45.5)	27 (61.4)
DERS	WL/TAU	8 (19.0)	13 (31.0)	20(47.6)	14 (33.3)	21 (50.0)
	S&T	1 (2.3)	9(20.5)	20(45.5)	10(22.7)	21 (47.1)
DES	WL/TAU	8 (19.0)	14 (33.0)	21(50.0)	14 (33.3)	21 (50.0)
	S&T	0 (0.0)	9(20.5)	22(50.0)	9(20.5)	22 (50.0)
HADS	WL/TAU	8 (19.0)	13 (31.0)	21(50.0)	14 (33.3)	22 (52.4)
	S&T	2(4.5)	9(20.5)	21(47.7)	11 (25.0)	23(52.3)
CCS	WL/TAU	8 (19.0)	13 (31.0)	20(47.6)	14 (33.3)	21 (50.0)
	S&T	2(4.5)	9(20.5)	21(47.7)	11 (25.0)	23(52.3)

Notes: WL/TAU= Wait Listed/ Treatment as Usual control arm; S&T= Survive & Thrive experimental intervention arm.

Given the substantial amount of missing data for the BAC-R it was not considered appropriate to use imputation methods from which to generate adjusted values for participants without initial 'pre' (i.e. T1) scores (Jakobsen, Gluud, Wetterslev & Winkel, 2017).

4.3.1. Comparisons between outcome completers and missing data participants

An analysis of MPD characteristics in respect of demographic and baseline characteristics (as measured by the TAQ and the SIDES) was undertaken with both the main self-report outcome measure (PCL-C) and the staff report measure (BAC-R). The differences in missing data between the two arms are presented in Appendix 3.3. This analysis involved exploratory testing to ascertain the possibility of otherwise unknown variables that might have led to MPD (Elliott, Cheruvelil, Montgomery & Soranno, 2016). These variables were selected on an intuitive and practice based perspective as there were no previous explanations within the existing population specific literature as to what might influence MPD. Aki, Shawwa, Kahale, Agoristsas et al (2015) also specify the importance of reporting the baseline characteristics of participants with and without missing data for each arm. Participants with missing data at T2 and T3 were 'dummy coded' with respect to dichotomous binary categories (MPD= 0 and

complete participant data= 1, following the method outlined by Heck, Thomas & Tabata, 2014 pg. 23). Both between and within group analyses were undertaken.

Differences between means and frequency of distributions for MPD were examined using Chi-square and ANOVA tests which indicated that age and sentence length were statistically significant variables (see Appendix 3.3.). This included a statistically significant interaction of study arm by sentence length for the PCL-C measure at T2 and T3 with WL/TAU arm 'completers' having longer sentences (T2: F (1,82)= 6.67, p=.012). There was also a statistically significant interaction of study arm by participants' age at T3 for the PCL-C (F (1,82)= 5.77, p=.019) and BAC-R outcome measures at T2 (F (1,82)= 4.02, p=.048. These results indicated that completers in the WL/TAU arm were significantly older compared to those in the S&T arm.

4.3.2. Missing completely at random (MCAR) / Missing at random (MAR) analysis

Factors that were associated with MPD at T2 were: 'could not contact' (i.e. transfer to another prison or released to the community) n= 9 (10.5%); refused/declined assessment n= 14 (16.3%). At T3 these factors accumulatively accounted for 'could not contact' n=28 (32.6%) and declined assessment n=15 (17.4%). A count of missing participant data can be accessed in Table 12.

To examine the assumption of the data being missing completely at random Little's test was used. Table 13 summarises the results for both the PCL-C and BAC-R measures. Whilst, results from Little's test indicate that data were missing at random (MAR) this is not a definitive demonstration of the reasons for missing data (Dziura, Post, Zhao, Fu & Peduzzi, 2013). Missing data within this study can best be understood as due to unforeseen circumstances connected to population management within the female prison estate and involving other criminal justice agencies as summarised in 'could not contact' (see Figure 6). In this respect data for an individual participant at a time point was missing rather than items within a measure. MAR was not assumed for those who 'refused' an assessment. Previous authors have concluded that such missingness can be a feature of an intervention, for example, participants finding the intervention to be unacceptable or the intervention making them feel worse (see Brunton-Smith, Carpenter & Kenward, 2014). However, consideration was also given to other population-specific (i.e. forensic) issues where refusal to undertake

assessment might indicate offence paralleled behaviours and attitudes (Jones, 2010). Participants who refused to engage in assessments were a relatively small group and as residential staff largely report 'refusals' without further explanation the actual reason for this was hard to gauge. Indeed, refusal could be connected to illness or an individual's decision to prioritise other activities at the time of assessment. It was therefore assumed that refusal to undertake assessments was a reflection of individual motivation and availability at a particular moment. The lack of differences in baseline measures between completer and missing participant data suggests that the predominant variable to explain MPD was sentence length. As discussed above this was particularly a factor for the WL/TAU arm as prison management structures would frequently conclude that participants on shorter sentences were eligible for release from custody under the Home Detention Curfew order (i.e. electronic tagging) as they were not involved in treatment. This would therefore suggest that there was a systematic bias to MPD. It was also recorded that problems with security based issues reduced the number of follow-up assessments for only 1 participant.

Table 13. Results from Little's missing completely at random (MCAR) test

Measure	Chi-Square (df), p value
PCL-C (Pre-Post)	4.44 (6) .617
PCL-C (Pre, Post, FU1)	10.63 (15).778
BAC-R (Pre-Post)	16.78 (12) .158
BAC-R (Pre, Post, FU1)	39.99 (36) .298

Notes: based on all subscales. Little's MCAR test should be non-significant ($p > 0.05$)

4.4. Psychometric analysis

Analyses were undertaken for two of the outcomes measures used in this study. In the case of the Behavioural Assessment Checklist-Revised (BAC-R) the internal consistency was investigated after adjustments were made in order to increase this measures sensitivity. With respect to the Criminal Cognitions Scale (CCS) the internal consistency was assessed due to the inclusion of a new subscale with no previous psychometric testing.

Behavioural Assessment Checklist-Revised

This measure was revised from the usual 3 point to a 5 point Likert scale in order to improve on the previously described deficits in reliability and validity (McDougal et al, 2009, pg. 14 and 50). These changes required the internal consistency of the adjusted measure to be assessed; the Cronbach's alpha for these computations are included in Table 14. It was hypothesised that this adjustment would increase the measure's sensitivity for recording any change in participant's behaviour. Table 14 also includes data both prior to and after corrections from exploratory Cronbach's Alpha procedures.

Table 14. Cronbach's alpha for the Behavioural Assessment Checklist-Revised and subscales

Subscale	Item Orientation	No. of Items	Cronbach's Alpha	Alpha after correction
Belligerence	Positive Indicators	8	.94	
	Opposite Indicators	1	-	
	All Items	9	.95	
Distress	Positive Indicators	6	.89	
Withdrawal	Positive Indicators	4	.43	
	Opposite Indicators	3	.49	
	All items	7	.54	.63 (5 items)
Impulsivity	Positive Indicators	12	.91	
	Opposite Indicators	1	-	
	All items	13	.91	
Egocentricity	Positive Indicators	6	.90	
	Opposite Indicators	2	.71	
	All items	8	.91	
Problem solving	Positive Indicators	7	.75	
	Opposite Indicators	4	.39	
	All items	11	.66	.80 (10 items)
BAC-R Total	All items	54	.91	.92 (51 items)

Table 14 continued

Note: 'Opposite' indicators are reversed scored items and indicate improvement or more pro-social behaviours. All tests based on participants' raw T1 'pre' intervention scores.

The minimum conventional level Cronbach's alpha is considered acceptable is .70 with the minimum reliability for research considered to be .80 and .90 for important decision making (Nunnaly & Bernstein, 1994). As such the subscales *Withdrawal* (all indicators) and *Problem Solving* (opposite indicators), as summarised in Table 14, were considered unacceptable and results suggested that either the sample size was not large enough or these subscales required modification to improve their reliability. Specific items within these subscales appeared to impact on the reliability of these scales. Therefore, changes were made to improve the scale as detailed below.

Withdrawal subscale items: Item 38 was negatively correlated with items 18 (α = -.05) and 32 (α =-.03) and was therefore deleted resulting in α= .60. Item 38 is written *No interests/hobbies* from considering participants responses it was hypothesised that on a number of occasions respondents seemed to struggle with the double negative involved in the *Never* option on the Likert scale; and at the other end of the scale the *Always* option. It was also noted that the reversed *Opposite Indicator* Item 8 *Seeks staff support* also had a very low correlation with other items. When this item was also deleted α= .63 was obtained.

Problem Solving subscale items: It was noted in the reverse item scored that Item 29 *Sticks to rigid routines* was negatively correlated with the other items and that its removal resulted in α= .80 within those 4 items and within all the items for that subscale.

Given this adverse situation the reliability of *Withdrawal* and *Problem Solving* subscales were 'corrected' with the removal of items 38 and 29. This correction was subsequently extended to the Total BAC-R score as a means of increasing the reliability of the use of this measure in the statistical models subsequently used. The corrected means, as obtained after removal of these items, were used in all subsequent analyses.

Criminal Cognitions Scale

Results from the Cronbach's alpha test indicated that removing any of the items within subscales would not have positively adjusted (i.e. improved) the overall Cronbach's score. Therefore, no corrections were made. The results in Table 14 below, excluding the new *Reparation* subscale, are comparable with other published psychometric data for the CCS (Tangney et al, 2012).

Table 15. Cronbach's alpha for the Criminal Cognitions Scale and subscales

Subscale	No. of Items	Cronbach's Alpha
Short Term Orientation	5	.44
Notions of Entitlement	5	.70
Failure to Accept Responsibility	5	.64
Negative Attitude Towards Authority	5	.71
Insensitivity to Impact of Crime	5	.67
Reparation and Making Amends	8	.70
Total	33	.74

It has been argued that low alpha scores are often a feature of scales that contain fewer items, however, this may also be a feature of criminal cognition scales based on self-report (Burgoyne & Tyson, 2013). Similar, alpha scores, for example, have been reported for the CRIME PICS II (Frude, Honess & Maguire, 2013). Nevertheless, consideration should be given to the weak internal reliability of the CCS, particularly when interpreting results based on its subscales.

4.5. Fidelity analysis: quality assurance of treatment facilitation

An overall composite score was calculated from the quality assurance form used to assess the fidelity and integrity of S&T treatment delivery. The subscales of this form included: adherence to the treatment protocol, adherence to treatment style and adherence to facilitation skills. A total treatment quality rating (range 1 - 5) was then calculated from these subscales based on random checks undertaken by the treatment author (Dr Sandra Ferguson). These random checks included monitoring one session from each group treatment delivered. The frequency of the scores associated with each delivery of S&T is presented in Figure 8.

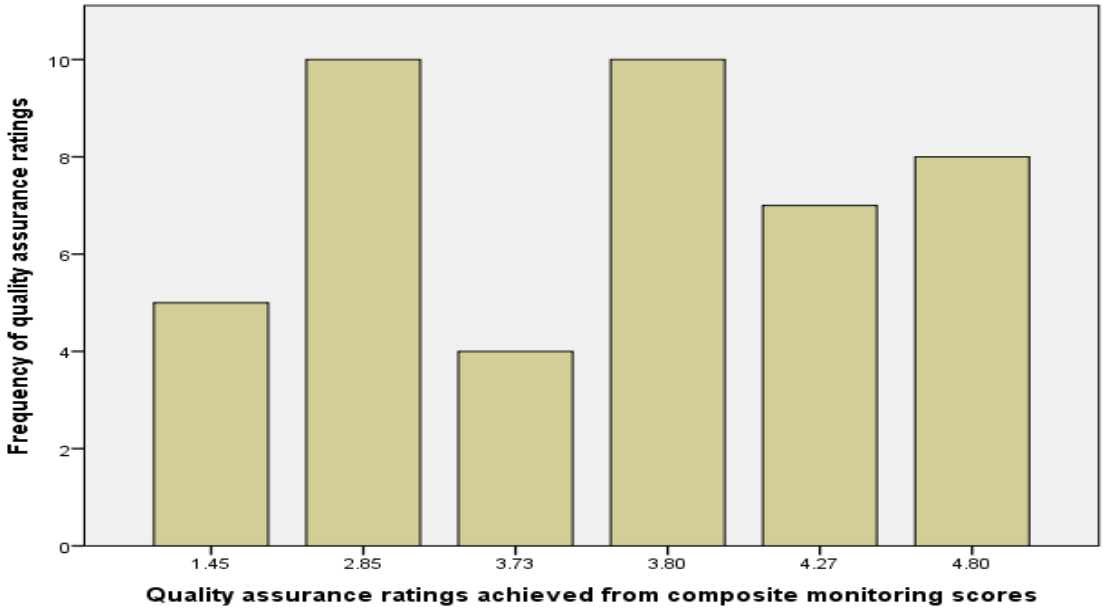


Figure 8. Frequency of the composite treatment quality assurance ratings achieved from monitoring the delivery of the Survive & Thrive intervention

Treatment quality scores varied from: 1.45 to 4.80, M= 3.57 (SD= 1.01), Mdn= 3.80 (IQR= 1.42). 56.8% (n=25) participants received treatment from S&T interventions which was facilitated at the median or above.

A simple linear regression analysis was completed on T2 (post treatment) scores across all measures (summary scores for each psychometric used) for the S&T arm. T2 was considered to be the time point where the most immediate impact of treatment quality would be apparent. Results suggested that treatment quality explained only a very small amount of the variance R^2 0% - 9% and did not significantly predict outcome (see Appendix 3.4.).

Of those individuals who completed 7 or more sessions (i.e. adequate dose participants n=33) only 48.5% (n=16) attended a delivery of S&T with a treatment quality rating above the median. 81.8% (n=9) of participants who did not receive an adequate dose attended S&T groups with quality assurance ratings above the median. Chi-square analysis indicated that the association between quality assurance and dose was not significant χ^2= 3.74(1), p= .053. There was also no observed difference between those who had received an adequate dose and those who had not, in the treatment quality mean scores (F (1,42)= 2.75, p= .105).

In summary, statistical analyses did not indicate that treatment quality (as assessed by the procedures used in this study) was associated with variations in treatment outcome or number of sessions attended.

4.6. Efficacy of Survive & Thrive: Intent-to-treat analysis

All MPD for the ITT analysis was accounted for using Multiple Imputation (MI) which is presented in Table 16. Appendix 3.5. also contains both raw and MI pooled data so that these can be compared. As previously noted there was a considerable amount of MPD for the BAC-R staff rated measure. As such MI was only used for those participants who had T1 measures.

Linear mixed modelling (LMM) was used as this provided greater power and flexibility to the analysis of data particularly in comparison to General Linear Modelling (GLM). In this respect LMM enabled individual change to be investigated. Modelling time as a linear variable produced a better model fit, as seen in the Akaike's Information Criterion (AIC) score, and enabled change to be investigated as an entire temporal sequence. However, modelling time as a categorical variable enabled change at specific assessment time points to be better understood in terms of treatment impact and sustainability. Another advantage of using LMM, particularly in the case of MPD for the BAC-R, was that it accounted for uneven sample sizes in the study's arms. See the Methods chapter for further detail.

Full Information Maximum Likelihood (FIML), as available in SPSS (version 23.0), was used in LMM procedures to account for MPD. LMM was therefore used to examine whether there was a significant difference between the two arms in the outcomes measured, which was interpreted as indicative of either symptom amelioration or increased stability, over time. Graphical representations of the change that occurred for each outcome across the time points can also be accessed in the Appendix 3.5.

RQ 1. *Will S&T be an efficacious intervention for promoting behavioural and emotional stability as associated with survivors of interpersonal trauma compared to a wait list control group (i.e. usual care) in a prison setting?*

Table 17 includes a summary of the interaction of the study's arms by time as computed using the LMM models outlined above. This was undertaken for the

subscales and total scores for each outcome measure. As noted in Table 17 there were few statistically significant differences across the majority of measures.

Symptom increases for the S&T arm were evident in BAC-R outcomes. For example, there was a medium effect size between the study's arms on the BAC-R *Impulsivity* subscale at T2 of *d*= .50 which reduced to *d*= .27 at T3. As summarised by the BAC-R total score this symptom increase was predominantly observed in the S&T arm between T1 (M= 69.94, SD= 23.72) and T2 (M= 73.18, SD= 19.84). There were however no statistically significant differences between the study's arms suggested a wide variance in participant's outcomes (β= 2.99 [95%CI, -10.97 to 16.96], p= .668).

RQ 2. *Will S&T be an efficacious intervention for stabilising symptoms associated with PTSD compared to a wait list control group in a prison setting?*

Symptom reduction was observed for both the S&T and WL/TAU arms in the majority of measures and subscales. These small, non-statistically significant, reductions were reflected in the mean differences and effect sizes presented in Table 16. This is summarised in the somewhat larger, although not statistically significant, mean difference from T1 to T3 for the S&T arm in the PCL-C total score (M= 5.34, SD= 10.24) compared to the WL/TAU arm (M= 2.23, SD= 10.57).

The PCL-C and BAC-R total scores across the time points are illustrated below in Figures 9 and 10. This highlights that reductions in the self-reported scores in the PCL-C do not appear to be associated with staff reported outcomes in the BAC-R.

RQ 3. *Will S&T be an efficacious treatment for stabilising general symptoms of psychopathology compared to a wait list control group in a prison setting?*

An exception to the non-significant differences in outcomes between the arms was observed in the HADS *Depression* subscale (β= .95, 95% CI .11 to 1.79, p= .027). This reflected a decrease in the level of depression in the WL/TAU arm from T1 (M= 12.12, SD= 2.50) to T3 (M= 11.87, SD=3.42) and a corresponding increase from T1 (M= 11.11, SD= 3.31) to T3 (M= 12.15, SD= 3.13) in the S&T arm. Whilst, there were differences in the direction of clinical outcomes for this subscales these differences were not substantive and were represented by small between arm effect sizes (T2: *d*= .13; T3: *d*=.09).

Another statistically significant finding was evident in the time by study arm interaction for the DERS *Non-Acceptance* subscale (β= -1.65 [95%CI, -3.22 to -.07] p= .041). This difference between the arms indicated a decrease in the S&T arm from T1 (M= 18.73, SD= 7.15) to T3 (M= 16.08, SD= 7.30) compared to the WL/TAU arm (T1: M=17.78, SD=6.22; T3: M=17.46, SD=6.87). The largest difference between the two study arms in Non-Acceptance occurred after T2 (β= 3.16 [95%CI, -.03 to 6.35] p= .052).

It should also be considered that in situations where multiple comparisons are made, statistically significant differences, as described below, can represent spurious findings and therefore should be interpreted with caution. The impact of using a more conservative probability threshold (p< 0.01) and Bonferroni confidence interval adjustments within the LMM used were also checked. The estimates and significance levels as reported in Tables 17 and 19 remained the same.

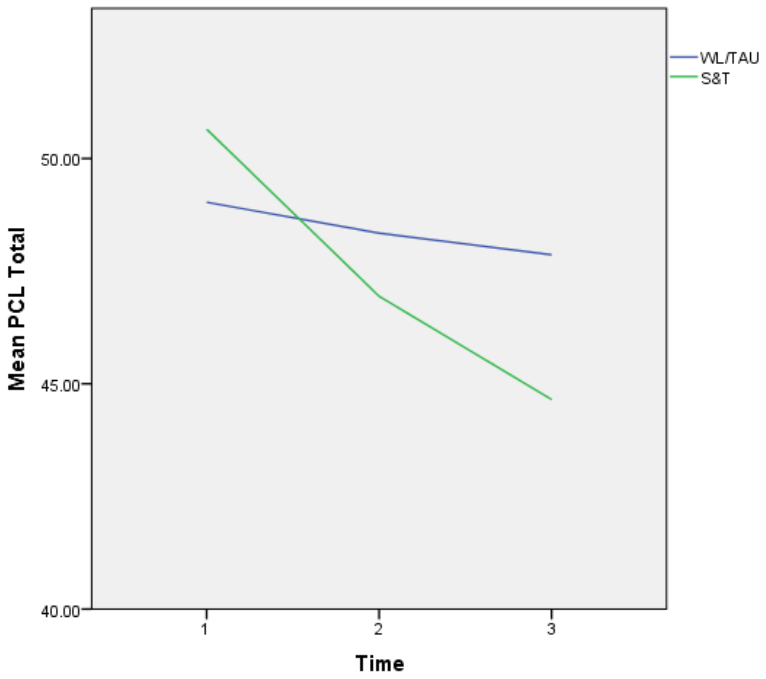


Figure 9: Comparison between WL/TAU control and S&T intervention outcome data for PCL-C Total scores.

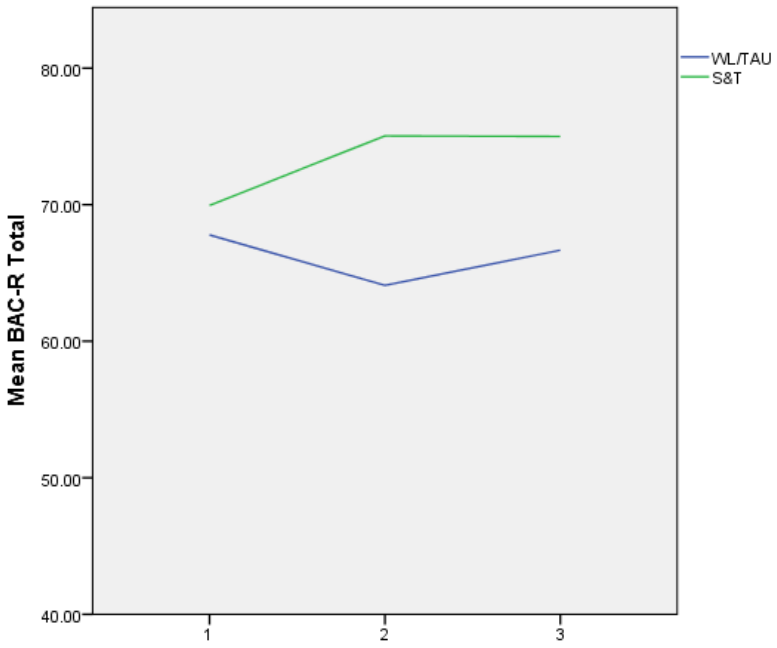


Figure 10: Comparison between WL/TAU control and S&T intervention outcome data for BAC-R Total scores.

Table 16: ITT outcomes for all measures across time points and study arms

		T1	**T2**	**T3**	**Mean Difference**		**Between arm ES**	
		Pre treatment	*Post treatment*	*Follow Up*	T1 – T2	T1 – T3	T2	T3
Measure / Study Arm		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Cohen's *d*	
PTSD Checklist								
Intrusion	WL/TAU	15.93(5.48)	15.33(5.62)	13.47(5.83)	.59 (2.81)	2.46 (3.53)	.23	.18
	S&T	14.82(6.00)	14.04(5.75)	12.57(4.29)	.78 (4.73)	2.25 (4.28)		
Avoidance	WL/TAU	18.80(6.67)	18.32(7.26)	19.68(7.94)	.47 (5.27)	-.88 (6.48)	.02	.18
	S&T	20.10(6.86)	18.66(7.26)	18.34(6.61)	1.44 (5.15)	1.76 (4.50)		
Arousal	WL/TAU	14.62(5.74)	14.58(5.74)	13.97(6.20)	.04 (4.06)	.65 (3.55)	.01	.07
	S&T	15.68(5.44)	14.54(6.36)	14.35(5.01)	1.14 (4.36)	1.32 (4.79)		
Total	WL/TAU	49.35(16.19)	48.24(16.84)	47.11(18.04)	1.10 (9.41)	2.23 (10.57)	.06	.11
	S&T	50.60(16.56)	47.24(16.62)	45.26(14.41)	3.36 (11.74)	5.34 (10.24)		
Behavioural Assessment Checklist-Revised								
Belligerence	WL/TAU	6.61 (5.62)	7.00 (5.36)	7.01 (5.00)	-.4 (3.52)	-.41 (6.76)	.27	.14
	S&T	8.45 (6.96)	8.72 (7.17)	7.99 (8.96)	-.27 (4.61)	.46 (5.7)		
Distress	WL/TAU	7.96 (4.32)	6.83 (4.15)	7.14 (3.08)	1.13 (3.49)	.82 (5.24)	.22	.35
	S&T	7.34 (4.76)	7.68 (3.65)	8.38 (4.01)	-.33 (5.02)	-1.04 (5.51)		
Withdrawal	WL/TAU	6.89 (2.92)	6.35 (3.66)	6.51 (2.67)	.55 (4.03)	.39 (3.49)	.08	.01
	S&T	6.69 (3.14)	6.57 (1.78)	6.56 (3.90)	.12 (3.24)	.12 (3.73)		
Impulsivity	WL/TAU	14.25(7.92)	13.05(7.69)	13.66 (5.80)	1.2 (4.49)	.59 (9.21)	.50	.27
	S&T	14.71 (8.47)	17.17(9.20)	15.45 (7.24)	-2.46 (7.35)	-.74 (7.29)		
Egocentricity	WL/TAU	8.54 (5.37)	8.45 (5.98)	8.32 (4.69)	.09 (3.08)	.22 (6.78)	.29	.28
	S&T	8.71 (5.50)	10.31(6.88)	9.74 (5.42)	-1.61 (5.15)	-1.02 (4.93)		
Problem solving	WL/TAU	23.53 (6.12)	23.87 (5.47)	25.61 (4.93)	-.33 (3.95)	-2.08 (7.01)	.20	.21
	S&T	24.03 (6.84)	22.71(6.16)	24.52 (5.47)	1.31 (5.60)	-.49 (5.87)		
Total BAC-R	WL/TAU	67.79 (20.23)	65.54 (20.81)	68.26 (15.31)	2.25 (14.05)	-.48 (24.59)	.38	.26
	S&T	69.94 (23.72)	73.18(19.84)	72.64 (18.55)	-3.24 (16.34)	-2.7 (16.54)		

Table 16 continued

		T1	**T2**	**T3**	**Mean Difference**		**Between arm ES**	
		Pre treatment	*Post treatment*	*Follow Up*	T1 – T2	T1 – T3	T2	T3
Measure / Study Arm		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Cohen's *d*	
Difficulties in Emotional Regulation Scale								
Non-Acceptance	WL/TAU	17.78(6.22)	18.16 (6.88)	17.46 (6.87)	-.37 (5.35)	.32 (5.16)	.22	.20
	S&T	18.73(7.15)	16.65 (6.53)	16.08 (7.30)	2.07 (6.06)	2.64 (6.17)		
Goals	WL/TAU	16.55(5.49)	14.60 (5.53)	14.05 (5.03)	1.96 (3.98)	2.50 (4.41)	.17	.07
	S&T	16.43(4.71)	15.49 (5.12)	14.42 (5.98)	.94 (3.80)	2.02 (5.43)		
Impulsive	WL/TAU	15.58(7.99)	14.57 (7.15)	13.62 (6.28)	1.02 (3.17)	1.96 3.69)	.06	.07
	S&T	16.65(6.30)	14.88 (6.01)	14.06 (6.65)	1.77 (4.59)	2.59 (6.76)		
Aware	WL/TAU	19.73(6.10)	18.49 (5.94)	17.34 (5.86)	1.24 (4.31)	2.39 (4.28)	.11	.05
	S&T	17.77 (5.98)	19.34 (8.62)	17.06 (6.23)	-1.55 (8.27)	.73 (4.28)		
Strategies	WL/TAU	23.24 (8.05)	21.36 (8.13)	19.96 (8.34)	1.88 (1.29)	3.28 (5.94)	.14	-.03
	S&T	23.37 (7.82)	22.59 (8.25)	20.21 (8.86)	.77 (6.66)	3.16 (7.51)		
Clarity	WL/TAU	13.44 (4.21)	12.78 (3.61)	11.69 (4.56)	.66 (2.94)	1.74 (4.49)	.28	.06
	S&T	14.00 (4.60)	13.97 (4.87)	11.44 (4.46)	-02 (3.63)	2.51 (4.31)		
DERS Total	WL/TAU	106.33 (29.27)	99.95 (27.24)	94.14 (29.49)	6.39 (16.87)	12.20 (18.78)	.11	.03
	S&T	106.86(27.67)	103.83(28.47)	93.27 (31.23)	3.99 (22.10)	13.65 (22.26)		
Dissociative Experiences Scale								
Amnesia	WL/TAU	23.84(19.57)	24.05(19.87)	24.76(19.36)	-.20 (11.05)	-.92 (13.24)	.15	.18
	S&T	22.52(18.73)	21.49(15.09)	21.66(15.03)	1.03 (11.87)	.86 (12.95)		
Absorption	WL/TAU	31.57(20.86)	33.58(22.62)	30.02(20.11)	-2.01 (12.63)	1.55 (9.75)	.05	.05
	S&T	32.82(20.11)	32.43(19.79)	29.12(14.78)	.40 (14.13)	3.71 (9.51)		
Depersonalisation	WL/TAU	18.67(18.06)	20.27(18.50)	20.85(19.61)	-1.61 (10.55)	-2.18 (11.16)	.05	.07
	S&T	20.80(19.82)	21.23(18.13)	22.03(15.99)	-.43 (14.09)	-1.23 (12.91)		
DES Total	WL/TAU	28.27(20.43)	29.68(21.59)	28.56(21.24)	-1.41 (11.31)	-.28 (10.97)	.02	.05
	S&T	29.41(20.21)	29.24(19.41)	27.63(16.21)	.17 (13.03)	1.78 (10.60)		

Table 16 continued

		T1	**T2**	**T3**	**Mean Difference**		**Between arm ES**	
		Pre treatment	*Post treatment*	*Follow Up*	T1 – T2	T1 – T3	T2	T3
Measure / Study Arm		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Cohen's *d*	
Hospital Anxiety Depression Scale								
Anxiety	WL/TAU	11.97(3.73)	11.95(3.47)	11.50(4.21)	.02 (2.02)	.46 (3.54)	.11	.04
	S&T	12.61(2.68)	11.58(3.32)	11.64(2.69)	1.02 (3.03)	.97(3.19)		
Depression	WL/TAU	12.12(2.50)	11.92(3.08)	11.87(3.42)	.19 (2.61)	.24 (1.70)	.13	.09
	S&T	11.11(3.31)	12.30(2.70)	12.15(3.13)	-1.20 (2.53)	-1.05 (3.92)		
HADS Total	WL/TAU	24.09(2.42)	23.87(3.42)	23.38(3.35)	.21 (3.14)	.71 (2.54)	.01	.12
	S&T	23.72(2.88)	23.88(3.70)	23.78(3.40)	-.18 (2.89)	-.08 (4.19)		
Criminal Cognitions Scale								
Short-term	WL/TAU	9.56(2.41)	9.66 (2.89)	9.79 (2.46)	-.09 (2.17)	-.22 (1.92)	.45	.17
	S&T	10.40(2.74)	10.79 (2.01)	10.20 (2.50)	-.39 (2.99)	.20 (3.13)		
Entitlement	WL/TAU	8.88 (2.43)	8.92 (3.05)	9.23 (3.39)	-.03 (2.47)	-.35 (2.22)	.18	.08
	S&T	9.19 (2.97)	9.40 (2.13)	9.04 (2.19)	-.21 (2.61)	.15 (2.95)		
Responsibility	WL/TAU	11.47 (2.95)	11.33 (3.00)	11.29 (3.39)	.14 (1.92)	.18 (1.94)	.11	.01
	S&T	11.81 (2.86)	11.66 (2.77)	11.33 (2.19)	.15 (3.20)	.48 (2.79)		
Authority	WL/TAU	10.61 (2.69)	10.93 (2.95)	11.29 (3.12)	-.32 (2.13)	-.69 (1.98)	.10	.05
	S&T	11.51 (2.93)	11.19 (1.98)	11.15 (2.19)	.32 (2.86)	.36 (3.06)		
Intensivity	WL/TAU	8.88 (2.91)	8.47 (3.46)	8.52 (3.02)	.42 (2.06)	.36 (1.3)	.09	.17
	S&T	9.38 (2.89)	8.74 (2.28)	9.00 (2.56)	.64 (2.34)	.39 (2.14)		
Reparation	WL/TAU	21.60 (5.29)	22.06 (5.52)	23.18 (4.68)	-.46 (4.39)	-1.59 (3.22)	.01	.21
	S&T	20.69 (4.44)	22.02 (3.79)	22.30 (4.07)	-1.34 (4.00)	-1.61 (3.86)		
CCS Total	WL/TAU	70.71 (12.18)	71.19 (13.92)	73.13 (13.27)	-.48 (9.02)	-2.42 (5.66)	.23	.01
	S&T	72.89 (12.18)	73.81 (7.98)	73.02 (8.71)	-.92 (12.00)	-.13 (10.23)		

Note: ITT data based on MI: WL/TAU control arm (n= 42); S&T: Survive & Thrive experimental arm (n=44) except for BAC-R: Behavioural Assessment Checklist: (n= 28 and n=35 respectively). Between-group effect sizes (Cohen's d) based on post assessment and also at follow up assessment.

Table 17: ITT analysis of the interaction between study arm and time

	T2 *(Post)* **x WL/TAU vs S&T**			**T3** *(Follow Up)* **x WL/TAU vs S&T**			**Time x WL/TAU vs S&T**		
	β	(95% CI)	p	β	(95% CI)	p	β	(95% CI)	p
PTSD Checklist Civilian Population									
Intrusion	.01	(-2.08 to 2.08)	.998	-.80	(-2.79 to 1.19)	.423	-.01	(-1.04 to 1.04)	.999
Avoidance	2.11	(-.75 to 4.97)	.144	2.55	(-.43 to 5.53)	.093	-1.26	(-2.74 to .23)	.097
Arousal	-.99	(-3.08 to 1.10)	.345	.28	(-1.94 to 2.51)	.800	-.18	(-1.29 to .938)	.751
PCL Total	2.76	(-2.88 to 8.40)	.331	.25	(-5.15 to 5.65)	.927	-1.47	(-4.30 to 1.36)	.303
Behavioural Assessment Checklist- Revised									
Belligerence	1.92	(-2.41 to 6.24)	.377	-.26	(-4.43 to 3.90)	.869	.97	(-.80 to 2.74)	.278
Distress	-1.33	(-4.22 to 1.56)	.360	-2.84	(-5.86 to .176)	.064	1.39	(-.11 to 2.89)	.069
Withdrawal	-.77	(-3.16 to 1.63)	.522	-.86	(-3.24 to 1.51)	.469	.36	(-.79 to 1.51)	.531
Impulsivity	2.74	(-2.66 to 8.13)	.313	-2.12	(-7.28 to 3.05)	.412	1.94	(-.47 to 4.36)	.113
Egocentricity	.59	(-3.50 to 4.67)	.774	-1.76	(-5.49 to 1.96)	.363	1.37	(-.29 to 3.03)	.104
Problem solving	.20	(-4.05 to 4.46)	.923	2.20	(-1.98 to 6.37)	.294	-1.38	(-3.40 to .63)	.175
Total BAC-R	2.99	(-10.97 to 16.96)	.668	-6.10	(-19.77 to 7.57)	.372	4.60	(-1.68 to 10.88)	.148
Difficulties in Emotional Regulation Scale									
Non-Acceptance	.93	(-2.29 to 4.15)	.566	3.16	(-.03 to 6.35)	.052	-1.65	(-3.22 to -.07)	.041
Goals	1.16	(-.92 to -3.24)	.266	-.25	(-2.65 to 2.14)	.833	.23	(.-97 to 1.42)	.706
Impulsive	.416	(-1.54 to 2.37)	.670	1.67	(-.93 to 4.27)	.206	-.83	(-2.12 to .47)	.209
Aware	1.63	(-2.03 to 5.29)	.379	-1.21	(-3.71 to 1.30)	.338	.72	(-.52 to 1.96)	1.96
Strategies	1.78	(-1.27 to 4.84)	.247	.82	(-2.64 to 4.27)	.639	-.34	(-2.06 to 1.39)	.696
Clarity	1.87	(-.14 to 3.88)	.068	1.12	(-1.07 to 3.32)	.310	-.34	(-1.42 to .74)	.538
DERS Total	8.14	(-2.89 to 19.18)	.145	6.19	(-5.19 to 17.57)	.281	-2.60	(-8.22 to 3.02)	.359
Dissociative Experiences Scale									
Amnesia	-1.50	(-7.86 to 4.87)	.638	.70	(-6.85 to 8.25)	.853	-.52	(-4.27 to 3.23)	.783
Absorption	1.21	(-4.56 to 6.99)	.675	4.14	(-1.35 to 9.62)	.137	-1.93	(-4.63 to .76)	.156
Depersonalisation	-1.24	(-7.30 to 4.81)	.681	.31	(-6.85 to 7.47)	.932	-.10	(-3.69 to 3.48)	.954
DERS Total	-.41	(-6.40 to 5.58)	.892	1.93	(-4.47 to 8.34)	.548	-.92	(-4.13 to 2.28)	.566

Table 17 continued

	T2 *(Post)* **x WL/TAU vs S&T**			**T3** *(Follow Up)* **x WL/TAU vs S&T**			**Time x WL/TAU vs S&T**		
	β	(95% CI)	p	β	(95% CI)	p	β	(95% CI)	p
Hospital Anxiety Depression Scale									
Anxiety	-.35	(-2.00 to 1.29)	.687	1.02	(-.67 to 2.72)	.223	-.67	(-1.50 to .16)	.113
Depression	-.24	(-1.94 to 1.47)	.783	-1.64	(-3.39 to -.11)	.066	.95	(.11 to 1.79)	.027
HADS Total	-.58	(-2.68 to 1.52)	.586	-.62	(-2.50 to 1.27)	.515	.25	(-.62 to 1.13)	.565
Criminal Cognitions Scale									
Short-term	.81	(-.37 to 1.98)	.174	.08	(-1.22 to 1.37)	.909	-.08	(-.78 to .61)	.812
Entitlement	.35	(-.92 to 1.62)	.585	.04	(-1.32 to 1.40)	.949	-.02	(-.71 to .66)	.947
Responsibility	.25	(-1.05 to 1.55)	.701	.03	(-1.33 to 1.40)	.961	-.02	(-.70 to .67)	.958
Authority	.72	(-.60 to 2.04)	.280	.93	(-.57 to 2.42)	.220	-.45	(-1.21 to .31)	.243
Intensivity	-.05	(-1.23 to 1.12)	.927	-.14	(-1.14 to .87)	.782	.06	(-.47 to .59)	.826
Reparation	1.13	(-.94 to 3.19)	.278	-.12	(-2.25 to 2.00)	.907	.05	(-1.01 to 1.12)	.921
CCS Total	3.05	(-1.50 to 7.59)	.185	.59	(-4.21 to 5.38)	.808	-.62	(-3.19 to 1.95)	.630

Note: WL/TAU= Wait Listed/ Treatment as Usual control arm; S&T= Survive & Thrive experimental intervention arm. Linear mixed model estimates of the treatment effects: β (Estimate of Fixed Effects) at post (T2) and follow up, 1 month after treatment, (T3). Full Information Maximum Likelihood used to account for missing data. The separate analysis for T2 and T3 used time since randomisation as a categorical variable, with time, participant and slope random effects, treatment, treatment by time interaction, and time as fixed effects and treatment group specified as a baseline covariate. All results presented utilising a linear time trajectory.

4.7. Efficacy of Survive & Thrive: Adequate dose analysis

4.7.1. Comparison of adequate dose and non-adequate dose participants

In order to consider whether treatment efficacy was influenced by adequate participation or 'dose' (as defined by attending a minimum of 7 sessions) an initial comparison was undertaken between adequate dose (AD) and non-adequate dose (Non-AD) participants. This was undertaken for only the main outcome measures the PCL-C and the BAC-R. The number of participants that were categorised as AD participants was n= 33 versus n= 11 participants who were classified as Non-AD.

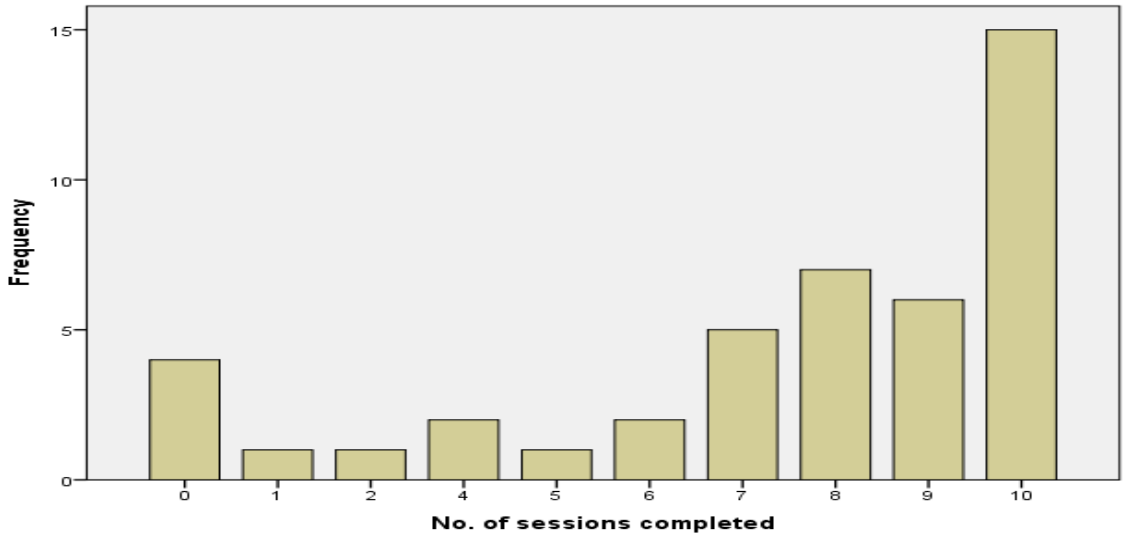


Figure 11: Frequency of sessions completed by S&T arm participants

Figure 11 presents a graphical illustration of the frequency of sessions completed for the S&T intervention arm. Sessions completed: M= 7.3 (SD 3.2), Mdn= 8 (IQR= 6). n=33 participants completed 7 or more sessions. This included 15 participants who completed all 10 sessions. 11 participants did not complete 7 sessions, including 4 participants who completed no sessions.

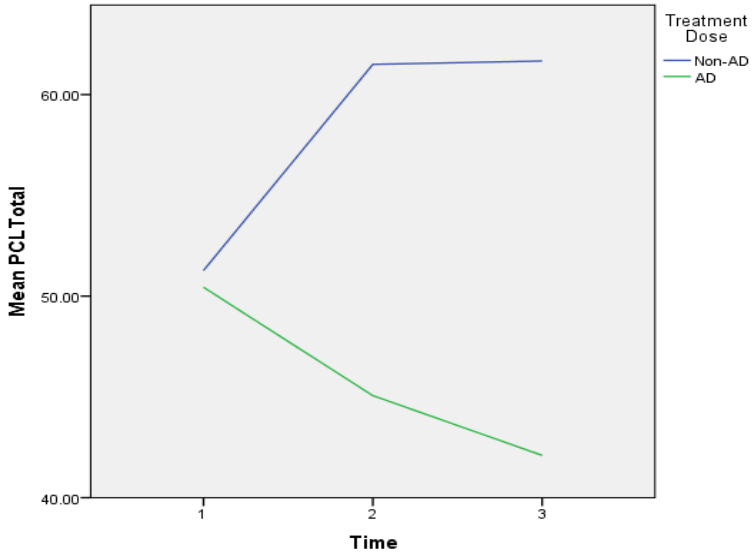


Figure 12: Comparison between AD and non-AD S&T intervention participants outcome data for PCL-C total scores

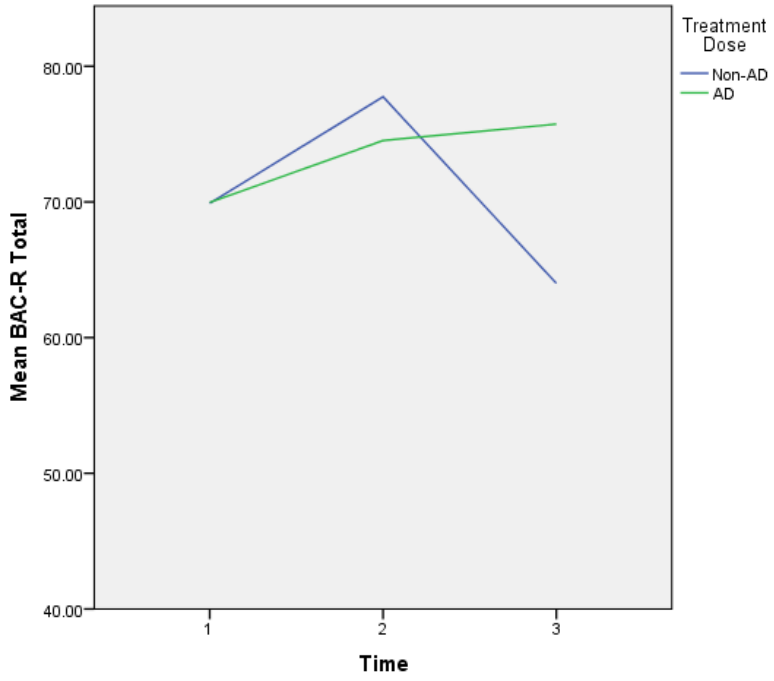


Figure 13: Comparison between AD and non-AD S&T intervention participants outcome data for BAC-R total scores

Differences between AD and Non-AD participants were investigated in respect of the PCL-C and BAC-R outcomes (see Figures 12 and 13). As there was considerable MPD within the Non-AD group ITT (MI) was not used. This was particularly the case for BAC-R as only n=1 Non-AD participant's data was available at T3 and therefore no analyses at this time point were attempted. Given this situation Table 18 presents the raw means for AD and Non-AD participants.

Only the LMM model computing time as a categorical variable was used as it was noticed that this produced a consistently better fit for this analysis. For example, when

computing the PCL-C total the AIC score for the categorical use of time was 800.136 compared to 802.404 for the linear coding of time. Therefore, the F statistic produced by the categorical use of time in the LMM statistics (tests for fixed effects) was referred to for the overall AD vs Non-AD by time interaction. In this respect the AD participants showed a statically significant greater reduction on the PCL-C total score over time (F (29,2)= 3.36, p= .049) and on the *Avoidance* subscale (F (30,2)= 4.11, p= .026) but not for the *Intrusion* (F (31,2)= 3.03, p= .063) or *Arousal* (F (31,2)= 1.06, p= .359) subscales. These results also corresponded with the substantive between arm effect sizes computed at T2 and T3 as presented in Table 18 and as graphically presented in Figure 12. A statistically significant result is also presented in Table 18 for the PCL-C *Intrusion* subscale occurring after T2.

Only small non-statistically significant differences were evident between the AD and Non-AD participants in the BAC-R measure. Results from the BAC-R measure appear to be in contrast to the differences computed for the PCL-C measure.

Table 18: PCL-C and BAC-R outcome data and analysis of the interaction between study arm and time based on adequate dose and non-adequate dose participants

		T1 *Pre treatment* M (SD)	**T2** *Post treatment* M (SD)	**T3** *Follow Up* M (SD)	**Between arm ES** T2 Cohen's *d*	T3	**Time x** WL/TAU **vs S&T** T2 x WL/TAU vs S&T β	(95% CI)	p	T3 x WL/TAU vs S&T β	(95% CI)	p
Measure / Study Arm												
PTSD Checklist												
Intrusion	Non-AD	14.82 (6.63)	18.25 (5.85)	19.33 (3.21)	.93	2.39	.15	(-4.06 to 4.36)	.942	-4.73	(-8.81 to .66)	.024
	AD	14.81 (5.87)	12.94 (5.54)	11.40 (3.41)								
Avoidance	Non-AD	20.27 (7.86)	25.25 (5.50)	23.67 (7.64)	1.13	0.96	.4.54	(-.92 to 10.00)	.100	-2.64	(-7.35 to 2.88)	.378
	AD	20.06 (6.61)	18.06 (7.12)	17.00 (6.19)								
Arousal	Non-AD	16.18 (6.49)	18.00 (5.77)	18.67 (4.51)	0.71	1.06	2.24	(-2.15 to 6.63)	.303	-.54	(-5.27 to 4.20)	.820
	AD	15.56 (5.13)	14.06 (5.34)	13.70 (4.85)								
Total	Non-AD	51.27 (19.32)	61.50 (16.78)	61.67 (15.18)	1.01	1.39	7.62	(-3.25 to 18.48)	.162	-6.56	(17.61 to -4.50)	.236
	AD	50.44 (15.84)	45.06 (15.91)	42.10 (12.79)								
Behavioural Assessment Checklist- Revised												
Belligerence	Non-AD	8.60 (7.47)	11.75 (5.50)	7.00 (-)	.35	-	-2.11	(-14.85 to 10.62)	.896	-.77	(-13.04 to 11.50)	.896
	AD	8.40 (6.91)	9.43 (7.50)	8.80 (6.16)								
Distress	Non-AD	8.30 (5.66)	8.25 (5.50)	6.00 (-)	.15	-	-4.99	(-13.66 to 3.69)	.245	-6.27	(-15.01 to 2.46)	.150
	AD	6.96 (4.42)	7.57 (3.36)	9.13 (4.07)								
Withdrawal	Non-AD	6.50 (3.63)	6.00 (2.00)	3.00 (-)	.30	-	-2.99	(-11.09 to 5.10)	.446	-3.52	(-11.76 to 4.71)	.381
	AD	6.76 (3.00)	6.57 (1.78)	7.13 (3.89)								
Impulsivity	Non-AD	17.78 (9.60)	18.25 (5.56)	17.00 (-)	.07	-	-4.60	(-9.50 to 18.71)	.506	-9.37	(-3.86 to 22.61)	.153
	AD	14.20 (8.02)	17.71 (9.84)	16.20 (7.49)								
Egocentricity	Non-AD	9.20 (5.90)	13.00 (8.28)	9.00 (-)	.33	-	-6.09	(-16.05 to 3.86)	.218	-5.12	(-14.31 to 4.07)	.256
	AD	8.52 (5.45)	10.48 (6.74)	10.53 (5.59)								

Table 18 continued

		T1 *Pre treatment* M (SD)	**T2** *Post treatment* M (SD)	**T3** *Follow Up* M (SD)	**Between arm ES** T2 Cohen's *d*	T3	**Time x** WL/TAU **vs S&T** T2 x WL/TAU vs S&T β	(95% CI)	p	T3 x WL/TAU vs S&T β	(95% CI)	p
Measure / Study Arm												
Problem solving	Non-AD	21.30 (7.07)	20.50 (6.14)	22.00 (-)	.36	-	3.10	(-8.86 to 15.07)	.595	4.35	(-7.34 to 16.03)	.442
	AD	25.12 (6.56)	22.76 (6.25)	23.93 (5.64)								
Total BAC-R	Non-AD	69.90 (33.11)	77.75 (19.99)	64.00 (-)	.16	-	-17.01	(-52.79 to 18.76)	.335	-23.54	(-57.91 to 10.82)	.165
	AD	69.96 (19.64)	74.52 (20.26)	75.73 (18.96)								

Notes: Non-AD= Non-Adequate Dose participants (< 7 sessions), AD= Adequate Dose participants (≥ 7 sessions). All Means based on raw data except due to extent of missing data for Non-AD participants

4.7.2. Comparison of adequate dose and WL/TAU arm participants

Given the differences computed between AD and Non-AD participants in the S&T arm for the PCL-C measure a further analysis was undertaken between AD S&T and WL/TAU arm participants. This was undertaken using the same LMM modelling procedures as described above for the ITT analysis including accounting for MPD using FIML. The outcomes for AD S&T participants are presented in the Appendix 3.6. as well as in Table 19 below. Table 20 presents a summary of the time by study arm interaction at both T2 and T3 using time as a categorical variable as well as for the overall temporal sequence using a linear time coding.

RQ 4. *Will S&T be a more efficacious treatment for those participants who receive an 'adequate dose' compared to a waitlist control group in a prison setting?*

The main outcomes indicated small effect sizes in favour of AD S&T participants for most of the PCL-C subscales (*d*= .08 to .39) except for the *Arousal* subscale at T3. The interaction analysis also notes that there were few statistically significant differences between the study's arms. The BAC-R outcomes also noted small to medium effect sizes (*d*= .08 to .55) in favour of the WL/TAU arm. This was also statistically significant for the BAC-R *Distress* subscale (β= 1.75 [95%CI, .24 to 3.26] p= .023) where the most substantive increase was noted in the S&T arm between T1 (M=6.96, SD=4.42) and T3 (M=9.13, SD=4.07). The categorical modelling of time indicated that this statistically significant difference in the Distress subscale occurred after T2. There was no statistically significant difference between the arms for the BAC-R total score (β= 5.15, [95%CI, -1.15 to 11.42] p= .107).

With respect to general symptoms of psychopathology the outcome of the HADS *Depression* subscale were statistically significant with increases in depression noted for the AD S&T arm (β= 1.14 [95%CI, .27 to 2.00] p= .011). Conversely, the DERS subscale *Non-Acceptance* computed a statistically significant time by study arm interaction with reductions for the AD S&T arm (β= -1.92 [95%CI, -3.52 to -.31] p= .020). Both this increase in depression and decrease in non-acceptance for the AD-S&T arm occurred after T2.

The results presented in this section, as summarised in Table 20, used statistical modelling from LMM procedures that represent the best overall fit across the outcome data (lower AIC score and most parsimonious construction). Despite this and given the number of analyses undertaken caution should be applied to the statistically significant results discussed in this section.

Table 19: ITT outcomes for all measures across time points and study arms as adjusted for AD S&T participants

		T1	**T2**	**T3**	**Mean Difference**		**Between arm ES**	
		Pre treatment	*Post treatment*	*Follow Up*	T1 – T2	T1 – T2	T2	T3
Measure / Study Arm		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Cohen's *d*	
PTSD Checklist								
Intrusion	WL/TAU	15.93(5.48)	15.33(5.62)	13.47(5.83)	.59 (2.81)	2.46 (3.53)	.39	.37
	AD S&T	14.82 (5.87)	13.13 (5.54)	11.69 (3.41)	1.69 (4.62)	3.14 (4.45)		
Avoidance	WL/TAU	18.80(6.67)	18.32(7.26)	19.68(7.94)	.47 (5.27)	-.88 (6.48)	.04	.26
	AD S&T	20.04 (6.61)	18.05 (7.12)	17.84 (6.19)	2.00 (4.34)	2.20 (4.47)		
Arousal	WL/TAU	14.62(5.74)	14.58(5.74)	13.97(6.20)	.04 (4.06)	.65 (3.55)	.08	.01
	AD S&T	15.51 (5.13)	14.13 (5.34)	14.01 (4.85)	1.38 (3.91)	1.51 (4.68)		
Total	WL/TAU	49.35(16.19)	48.24(16.84)	47.11(18.04)	1.10 (9.41)	2.23 (10.57)	.18	.23
	AD S&T	50.38 (15.84)	45.31 (15.91)	43.54 (12.79)	5.07 (9.91)	6.85 (10.08)		
Behavioural Assessment Checklist-Revised								
Belligerence	WL/TAU	6.61 (5.62)	7.00 (5.36)	7.01 (5.00)	-.40 (3.52)	-.41 (6.76)	.37	.32
	S&T AD	8.40 (6.91)	9.43 (7.50)	8.80 (6.16)	-.66 (4.59)	-.02 (5.90)		
Distress	WL/TAU	7.96 (4.32)	6.83 (4.15)	7.14 (3.08)	1.13 (3.49)	.82 (5.24)	.20	.55
	AD S&T	6.96 (4.42)	7.57 (3.36)	9.13 (4.07)	-.58 (4.54)	-1.78 (5.32)		
Withdrawal	WL/TAU	6.89 (2.92)	6.35 (3.66)	6.51 (2.67)	.55 (4.03)	.39 (3.49)	.08	.19
	AD S&T	6.76 (3.00)	6.57 (1.78)	7.13 (3.89)	.15 (3.12)	-.09 (3.82)		
Impulsivity	WL/TAU	14.25(7.92)	13.05(7.69)	13.66 (5.80)	1.2 (4.49)	.59 (9.21)	.53	.38
	AD S&T	14.20 (8.02)	17.71 (9.84)	16.20 (7.49)	-3.14 (7.70)	-1.26 (6.99)		
Egocentricity	WL/TAU	8.54 (5.37)	8.45 (5.98)	8.32 (4.69)	.09 (3.08)	.22 (6.78)	.32	.43
	AD S&T	8.52 (5.45)	10.48 (6.74)	10.53 (5.59)	-1.73 (5.15)	-1.41 (5.00)		
Problem Solving	WL/TAU	23.53 (6.12)	23.87 (5.47)	25.61 (4.93)	-.33 (3.95)	-2.08 (7.01)	.19	.32
	AD S&T	25.12 (6.56)	22.76 (6.25)	23.93 (5.64)	2.21 (5.87)	.81 (6.01)		
Total BAC-R	WL/TAU	67.79 (20.23)	65.54 (20.81)	68.26 (15.31)	2.25 (14.05)	-.48 (24.59)	.44	.43
	AD S&T	69.96 (19.64)	74.52 (20.26)	75.73 (18.96)	-3.76 (16.68)	-3.75 (16.26)		

Table 19 continued

		T1	**T2**	**T3**	**Mean Difference**		**Between arm ES**	
		Pre treatment	*Post treatment*	*Follow Up*	T1 – T2	T1 – T2	T2	T3
Measure / Study Arm		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Cohen's *d*	
Difficulties in Emotional Regulation Scale								
Non-Acceptance	WL/TAU	17.78(6.22)	18.16 (6.88)	17.46 (6.87)	-.37 (5.35)	.32 (5.16)	.29	.28
	AD S&T	19.06 (6.96)	16.23 (6.29)	15.46 (7.19)	2.82 (6.12)	3.59 (5.85)		
Goals	WL/TAU	16.55(5.49)	14.60 (5.53)	14.05 (5.03)	1.96 (3.98)	2.50 (4.41)	.13	.01
	AD S&T	16.43 (4.92)	15.29 (5.13)	14.06 (5.89)	1.14 (3.72)	2.37 (5.21)		
Impulsive	WL/TAU	15.58 (7.99)	14.57 (7.15)	13.62 (6.28)	1.02 (3.17)	1.96 3.69)	.02	.01
	AD S&T	16.47 (6.54)	14.44 (5.83)	13.57 (6.29)	2.03 (4.16)	2.9 (6.41)		
Aware	WL/TAU	19.73(6.10)	18.49 (5.94)	17.34 (5.86)	1.24 (4.31)	2.39 (4.28)	.12	.15
	AD S&T	17.37 (5.86)	19.36 (9.13)	16.47 (6.06)	-2.00 (8.63)	.89 (4.29)		
Strategies	WL/TAU	23.24 (8.05)	21.36 (8.13)	19.96 (8.34)	1.88 (1.29)	3.28 (5.94)	.05	.05
	AD S&T	23.49 (8.09)	21.76 (7.90)	19.58 (8.28)	1.73 (5.89)	3.9 (6.73)		
Clarity	WL/TAU	13.44 (4.21)	12.78 (3.61)	11.69 (4.56)	.66 (2.94)	1.74 (4.49)	.15	.19
	AD S&T	14.03 (4.82)	13.37 (4.42	10.88 (3.84)	.65 (3.52)	3.14 (4.37)		
DERS Total	WL/TAU	106.33 (29.27)	99.95 (27.24)	94.14 (29.49)	6.39 (16.87)	12.20 (18.78)	.02	.14
	AD S&T	106.83 (29.12)	100.46 (27.31)	90.04 (28.67	6.37 (21.3)	16.8 (20.37)		
Dissociative Experiences Scale								
Amnesia	WL/TAU	23.84(19.57)	24.05(19.87)	24.76(19.36)	-.20 (11.05)	-.92 (13.24)	.22	.31
	AD S&T	22.75 (19.28)	20.10 (15.84	19.58 (13.76	2.65 (11.74	3.18 (9.72		
Absorption	WL/TAU	31.57(20.86)	33.58(22.62)	30.02(20.11)	-2.01 (12.63)	1.55 (9.75)	.17	.11
	AD S&T	33.48 (20.7)	29.67 (19.87	28.09 (14.40)	3.82 (13.39	5.39 (9.64		
Depersonalisation	WL/TAU	18.67(18.06)	20.27(18.50)	20.85(19.61)	-1.61 (10.55)	-2.18 (11.16)	.09	.46
	AD S&T	22.26 (20.94)	18.61 (17.09)	20.09 (12.43)	3.65 (12.35)	2.17 (10.36)		
DES Total	WL/TAU	28.27(20.43)	29.68(21.59)	28.56(21.24)	-1.41 (11.31)	-.28 (10.97)	.17	.16
	AD S&T	29.84 (20.94)	26.27 (19.11)	25.64 (14.36)	3.57 (12.05)	4.2 (9.16)		

Table 19 continued

		T1	**T2**	**T3**	**Mean Difference**		**Between arm ES**	
		Pre treatment	*Post treatment*	*Follow Up*	T1 – T2	T1 – T2	T2	T3
Measure / Study Arm		M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	Cohen's *d*	
Hospital Anxiety Depression Scale								
Anxiety	WL/TAU	11.97(3.73)	11.95(3.47)	11.50(4.21)	.02 (2.02)	.46 (3.53)	.19	.04
	AD S&T	12.53 (2.62)	11.32 (3.25)	11.38 (2.33)	1.21 (3.02)	1.15 (3.08)		
Depression	WL/TAU	12.12(2.50)	11.92(3.08)	11.87(3.42)	.19 (2.61)	.24 (1.70)	.18	.17
	AD S&T	11.07 (3.34)	12.47 (2.76)	12.41 (3.00)	-1.40 (2.52)	-1.33 (3.99)		
HADS Total	WL/TAU	24.09(2.42)	23.87(3.42)	23.38(3.35)	.21 (3.14)	.71 (2.54)	.02	.12
	AD S&T	23.61 (2.81)	23.79 (3.71)	23.79 (3.65)	-.19 (3.03)	-.18 (4.36)		
Criminal Cognitions Scale								
Short-term	WL/TAU	9.56(2.41)	9.66 (2.89)	9.79 (2.46)	-.09 (2.17)	-.22 (1.92)	.42	.09
	AD S&T	10.23 (2.67)	10.70 (1.98)	10.70 (2.28)	-.47 (2.67)	.23 (2.25)		
Entitlement	WL/TAU	8.88 (2.43)	8.92 (3.05)	9.23 (3.39)	-.03 (2.47)	-.35 (2.22)	.21	.10
	AD S&T	9.13 (2.67)	9.46 (2.00)	8.93 (2.37)	-.34 (2.5)	.20 (2.17)		
Responsibility	WL/TAU	11.47 (2.95)	11.33 (3.00)	11.29 (3.39)	.14 (1.92)	.18 (1.94)	.10	.01
	AD S&T	11.75 (2.67)	11.62 (2.77)	11.33 (1.96)	.13 (2.93)	.42 (1.86)		
Authority	WL/TAU	10.61 (2.69)	10.93 (2.95)	11.29 (3.12)	-.32 (2.13)	-.69 (1.98)	.01	.15
	AD S&T	11.37 (2.92)	10.95 (1.71)	10.88 (2.18)	.42 (2.47)	.49 (2.43)		
Intensivity	WL/TAU	8.88 (2.91)	8.47 (3.46)	8.52 (3.02)	.42 (2.06)	.36 (1.30)	.11	.14
	AD S&T	8.96 (2.61)	8.8 (2.23)	8.92 (2.66)	.16 (2.47)	.04 (1.96)		
Reparation	WL/TAU	21.60 (5.29)	22.06 (5.52)	23.18 (4.68)	-.46 (4.39)	-1.59 (3.22)	.04	.23
	AD S&T	20.76 (4.31)	21.85 (3.83)	22.15 (4.13)	-1.09 (4.01)	-1.39 (3.67)		
CCS Total	WL/TAU	70.71 (12.18)	71.19 (13.92)	73.13 (13.27)	-.48 (9.02)	-2.42 (5.66)	.20	.08
	AD S&T	72.10 (12.06)	73.4 (7.41)	72.23 (8.36)	-1.29 (10.76)	-.13 (5.40)		

Note: WL/TAU= Wait Listed/ Treatment as Usual control arm; AD S&T= Adequate Dose Survive & Thrive experimental intervention arm. ITT data based on MI: WL/TAU control arm (n= 42); AD S&T: Survive & Thrive Experimental arm (n=33) except for BAC-R: (n= 28 and n=25 respectively). Between-group effect sizes (Cohen's *d*) based on post assessment and also at follow up assessment.

Table 20. ITT comparison of the WL/TAU and AD S&T treatment arms at T1 - T3

	T2 *(Post)* **x WL/TAU vs AD S&T**			**T3** *(Follow Up)* **x WL/TAU vs AD S&T**			**Time x WL/TAU vs AD S&T**		
	β	(95% CI)	p	β	(95% CI)	p	β	(95% CI)	p
PTSD Checklist Civilian Population									
Intrusion	-.84	(-2.92 to 1.24)	.422	.52	(-1.65 to 2.69)	.635	-.29	(-1.38 to .80)	.601
Avoidance	1.56	(-1.35 to 4.47)	.285	2.76	(-.33 to 5.85)	.079	-1.44	(-2.98 to .09)	.065
Arousal	-1.29	(-3.41 to .84)	.230	.28	(-1.99 to 2.55)	.804	-.23	(-1.37 to .91)	.686
PCL Total	-.711	(-6.19 to 4.77)	.795	3.42	(-2.38 to 9.21)	.242	.173	(-4.91 to .90)	.173
Behavioural Assessment Checklist- Revised									
Belligerence	1.59	(-2.86 to 6.04)	.475	-.36	(-4.65 to 3.92)	.864	.88	(-.94 to 2.69)	.339
Distress	-1.62	(-4.50 to 1.26)	.262	-3.51	(-6.55 to -.47)	.024	1.75	(.24 to 3.26)	.023
Withdrawal	-.856	(-3.33 to 1.62)	.490	-1.06	(-3.47 to 1.36)	.384	.48	(-.70 to 1.65)	.418
Impulsivity	2.71	(-2.87 to 8.29)	.334	-3.01	(-8.23 to 2.22)	.251	2.36	(-.11 to 4.84)	.061
Egocentricity	.09	(-4.01 to 4.18)	.966	-2.13	(-5.93 to 1.67)	.264	1.41	(-2.97 to 3.12)	.104
Problem solving	.47	(-3.89 to 4.83)	.830	2.89	(-1.38 to 7.16)	.179	-1.68	(-3.75 to .40)	.111
Total BAC-R	2.08	(-11.99 to 16.15)	.767	-7.65	(-21.15 to 5.844)	.258	5.15	(-1.15 to 11.42)	.107
Difficulties in Emotional Regulation Scale									
Non-Acceptance	.730	(-2.60 to 4.06)	.663	3.61	(.36 to 6.86)	.030	-1.92	(-3.52 to -.31)	.020
Goals	1.18	(-.96 to 3.32)	.271	.181	(-2.28 to 2.65)	.884	-.02	(-1.25 to 1.21)	.970
Impulsive	.488	(-1.59 to 2.57)	.638	2.43	(-.23 to 5.09)	.073	-1.27	(-2.60 to .07)	.063
Aware	2.25	(-1.59 to 6.09)	.248	-.91	(-3.51 to 1.68)	.481	.60	(-.68 to 1.88)	.351
Strategies	1.42	(-1.62 to 4.46)	.352	1.51	(-1.93 to 4.95)	.383	-.72	(-2.43 to .99)	.404
Clarity	1.82	(-.26 to 3.90)	.085	1.63	(-.70 to 33.96)	.168	-.64	(-1.78 to .51)	.439
DERS Total	8.41	(-2.82 to 19.63)	.139	9.45	(-2.03 to 20.93)	.105	-4.38	(-10.05 to 1.28)	.127
Dissociative Experiences Scale									
Amnesia	-.52	(-6.82 to 5.79)	.879	2.90	(-4.33 to 10.14)	.424	-1.56	(-5.17 to 2.05)	.390
Absorption	.44	(-5.51 to 6.39)	.883	5.35	(-.37 to 11.07)	.066	-2.49	(-5.33 to .35)	.084
Depersonalisation	-1.19	(-6.72 to 4.34)	.666	3.15	(-3.86 to 10.15)	.371	-1.48	(-5.01 to 2.04)	.402
DERS Total	-.65	(-6.53 to 5.23)	.824	3.87	(-2.49 to 10.23)	.227	-1.91	(-5.10 to 1.29)	.236

Table 20 continued

	T2 *(Post)* **x WL/TAU vs AD S&T**			**T3** *(Follow Up)* **x WL/TAU vs AD S&T**			**Time x WL/TAU vs AD S&T**		
	β	(95% CI)	p	β	(95% CI)	p	β	(95% CI)	p
Hospital Anxiety Depression Scale									
Anxiety	-.32	(-1.98 to 1.34)	.699	1.22	(-.53 to 2.97)	.169	-.76	(-1.62 to .10)	.083
Depression	-.43	(-2.21 to 1.35)	.628	-2.03	(-3.83 to -.23)	.028	1.14	(.27 to 2.00)	.011
HADS Total	-.70	(-2.93 to 1.53)	.534	-.78	(-2.78 to 1.21)	.435	.33	(-.59 to 1.25)	.483
Criminal Cognitions Scale									
Short-term	.32	(-.96 to 1.59)	.620	.318	(-.96 to 1.59)	.620	-.20	(-.84 to .44)	.536
Entitlement	.452	(.49 to .64)	.452	.098	(-1.21 to 1.41)	.881	-.05	(-.70 to .611)	.889
Responsibility	.11	(-1.12 to 1.33)	.861	-.088	(-1.29 to 1.11)	.883	.04	(-.56 to .64)	.893
Authority	.75	(-.63 to 2.12)	.281	1.25	(-3.1 to 2.81)	.113	-.61	(-1.38 to .15)	.114
Intensivity	.16	(-1.01 to 1.34)	.785	.01	(-1.02 to 1.04)	.984	-.01	(-.53 to .50)	.963
Reparation	1.23	(-.98 to 3.44)	.270	.10	(-2.14 to 2.33)	.931	-.05	(-1.16 to 1.07)	.936
CCS Total	3.54	(-.94 to 8.01)	.119	1.14	(-2.97 to 5.26)	.570	-.82	(-2.83 to 1.19)	.408

Note: WL/TAU= Wait Listed/ Treatment as Usual control arm; AD S&T= Adequate Dose Survive & Thrive experimental intervention arm. Linear mixed model estimates of the treatment effects: β (Estimate of Fixed Effects) at post and follow up (1 month after treatment). Full Information Maximum Likelihood used to account for missing data. The analysis used time since randomisation as a categorical variable, with time, participant and slope random effects, treatment, treatment by time interaction, and time as fixed effects and treatment group specified as a baseline covariate. All results presented utilising a linear time trajectory.

4.8. Clinically significant and reliable change analysis

To further investigate the potential effectiveness of S&T a *post hoc* clinical change analysis was undertaken. This sought to compliment the understanding that had been gained in previous analyses which accounted for differences in mean change. The strength of clinically significant and reliable change analyses is in considering change at an individual level and where possible comparing the study population with normative data established from other non-clinical or relevant populations. Such comparisons also help to identify cut points thereby identifying individuals who have achieved change from participating in the intervention as well as potentially those that have achieved change from care as usual. As such those participants whose scores improve and cross the cut point, as assessed between different time points (i.e. from T1 to T2), may be considered to have moved from clinical levels to adaptive or more stable levels of functioning.

Following criteria established by Jacobson & Truax (1991) two complementary outcomes were computed. These were *Clinically Significant Change* (CSC) and the *Reliable Change Index* (RCI) which sought to ascertain the extent of change that had occurred and whether these were considered to be within the range of adaptive functioning. The RCI specifies the amount of change an individual must show for that change to be considered reliable (i.e., larger than that reasonably expected due to measurement error alone). Only if change is reliable can the clinically significant degree of these changes be analysed. CSC was considered to have occurred if this move was also statistically significant. All outcomes were calculated using the *Leeds Reliable Change Indicator* (Morley & Dowzer, 2014). See Methods section for further details.

As noted the RCI and CSC calculations utilised individual participant T1 and T2 scores. This was undertaken for all measures. However, an analysis comparing T1 and T3 scores was also completed for the PCL-C measure to consider the potential sustainability of outcomes. To further investigate the impact of adequate dose all analyses were based on outcomes from individuals previously identified as AD S&T participants. Only WL/TAU participants who completed T2 measures were included in

the CSC and RC analyses to ensure comparable sized arms from which to make indirect comparisons.

Figures 14 - 15 illustrate those participants in the AD S&T arm and those in the WL/TAU arm who achieved reliable change at T2 using the PCL-C measure. This graphical output also includes the line of no change, the RCI (red, parallel lines) and the cut scores. Individual data points are colour coded and the average of all the data is also shown. These outcomes are also presented in Table 21. The PCL-C total scores, as illustrated in Figures 14 - 15, utilises an 'external' cut point set at 45, as recognised in the relevant clinical literature (Blanchard, Jones-Alexander, Buckley, & Forneris, 1996). Where possible cut points and normative data were used for the other measures (see Methods section and Appendix 2.1.). It should be noted that this was not possible for either the BAC-R or CCS given adaptions made and due to a lack of relevant psychometric data. For these measures psychometric properties as established in this study were used.

From these analyses it was noted that similar CSC and RCI values and rates of change were computed for both arms for the PCL-C measure (see Table 21). However, there were more AD S&T participants who achieved positive change at T2 with reference to the RCI values for the PCL-C total score (30.3% vs 17.6%, OR 2.03 [95%CI, .64 to 6.43]) and for the *Intrusion* (15.2% vs 2.9%, OR 6.07 [95%CI, .67 to 55.04]) and *Arousal* (15.2% vs 5.9%, OR 2.86 [95%CI, 5.13 to 15.90]) subscales. This was not the case for the *Avoidance* subscale where more WL/TAU participants made progress (14.7% vs 9.1%, OR.58 [95%CI, .13 to 2.65]). These differences between the arms in the PCL-C subscales were however not evident by T3. There was however consistently less negative change associated with the AD S&T arm across the scales in the PCL-C.

Table 21: PCL- C clinically significant and reliable change for S&T AD and WL/TAU arm at T2 and T3

	Arm[1]	T1-T2 / T1- T3 ES (*d*)	SEM	RCI Value	No. of Change None	-	+	No. meeting CSC	CSC Criteria Used
PCL-C Outcomes at T2 (T1-T2)									
Intrusion	AD S&T	0.35	2.24	6.22	28	0	5	4	C (12.54)
	WL/TAU	0.18	2.12	5.88	32	1	1	1	
Avoidance	AD S&T	0.33	2.56	7.11	30	0	3	3	C (16.56)
	WL/TAU	0.15	2.58	7.16	28	1	5	5	
Arousal	AD S&T	0.29	1.85	5.14	27	1	5	5	C (12.18)
	WL/TAU	0.02	2.07	5.73	30	2	2	2	
Total PCL-C	AD S&T	0.36	3.87	10.73	22	1	10	5 [5*]	C (41.48)*
	WL/TAU	0.13	3.97	11.00	25	3	6	4 [4*]	
PCL-C Outcomes at T3 (T1-T3)									
Intrusion	AD S&T	0.41	2.24	6.22	28	0	5	4	C (12.54)
	WL/TAU	0.43	2.12	5.88	29	1	4	4	
Avoidance	AD S&T	0.18	2.56	7.11	31	0	2	1	C (16.56)
	WL/TAU	-0.09	2.58	7.16	28	5	1	1	
Arousal	AD S&T	0.11	1.85	5.14	29	1	3	3	C (12.18)
	WL/TAU	0.02	2.07	5.73	31	2	1	1	
Total PCL-C	AD S&T	0.26	3.87	10.73	28	1	4	4 [3*]	C (41.48)*
	WL/TAU	0.12	3.97	11.0	28	2	4	3 [3*]	

Notes: WL/TAU= Wait Listed/ Treatment as Usual control arm; AD S&T= Adequate Dose Survive & Thrive experimental intervention arm. [1]AD S&T: N= 33; WL/TAU: N= 34. *External cut off set at 45. T1- T2 and T1- T3 Effect Size (ES). SEM: Standard Error of Measurement. RCI: Reliable Change Index. No. of [participants who] Change: none i.e. 'no change', - i.e. 'deteriorate', + i.e. 'improved'. CSC: Clinically Significant Change

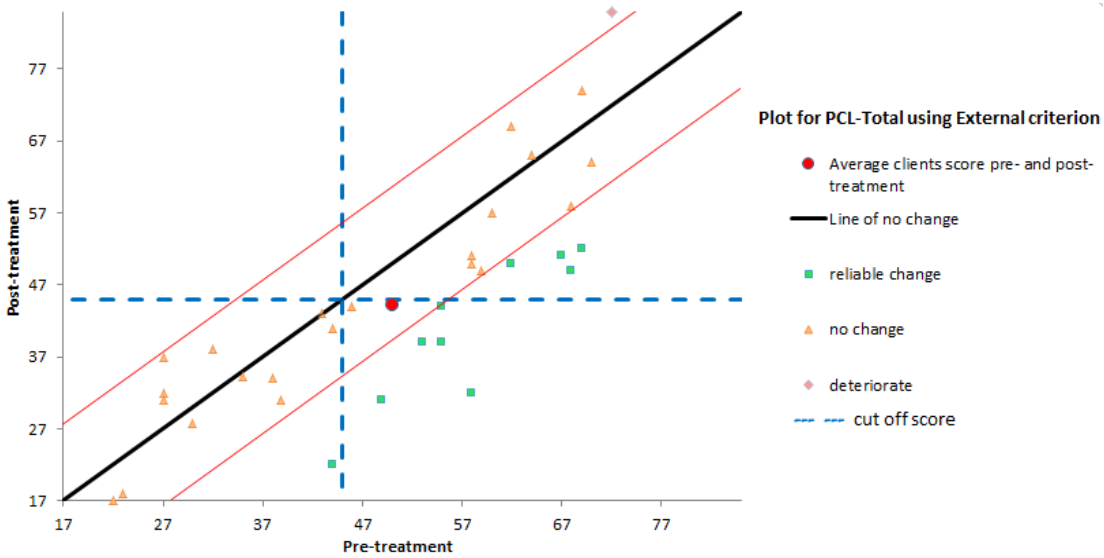


Figure 14: PCL-C total reliable change score for AD participants in the S&T arm at T2

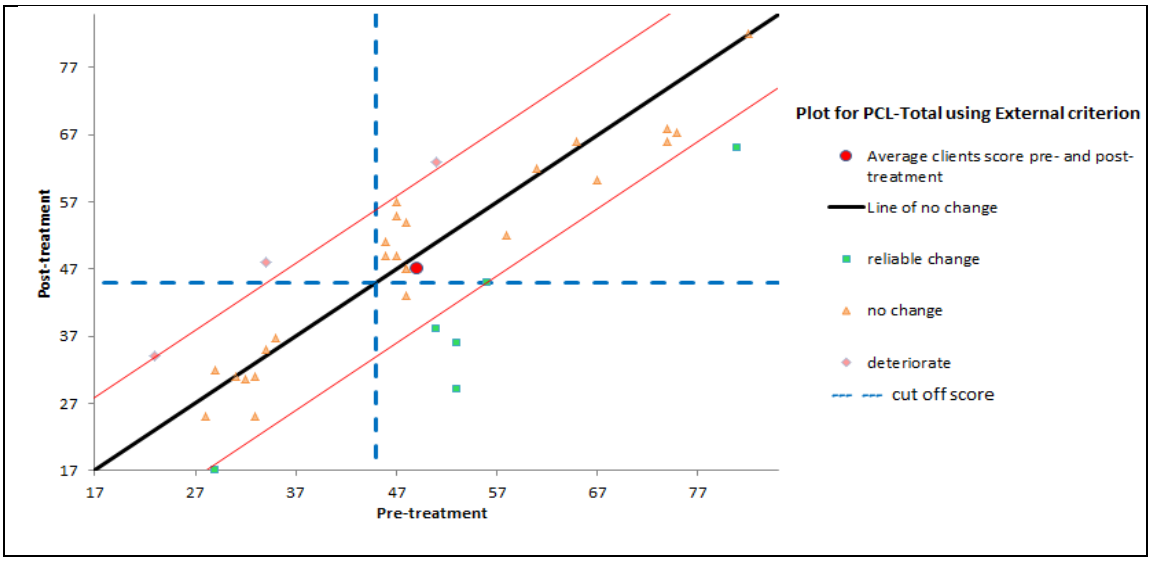


Figure 15: PCL-C total reliable change score for participants in the WL/TAU arm at T2

The RCI values computed for the BAC-R measure indicated that the WL/TAU arm was more effective at stabilising pathological behaviours (see Table 22). This was particularly evident in the *Distress* subscale at T2 with those in the WL/TAU arm achieving more positive change, i.e. decreasing behaviours as observed by staff (35.7% vs 12.0%, OR .25 [95%CI, .06 to 1.03]). This was similarly the case for the BAC-R Total score (25.0% vs 12.0%, OR .41 [95CI%, .09 to 1.80]) which is also illustrated in Figures 16 and 17. There were no participants in either arm who met the CSC criteria for the BAC-R measure.

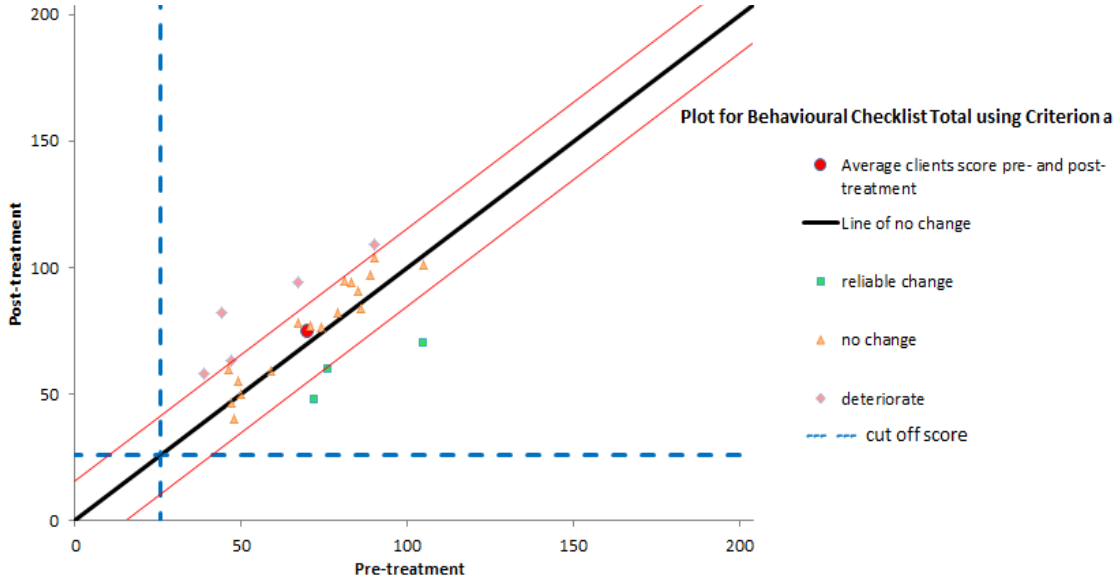


Figure 16: BAC-R at T2 changefor AD participants in the S&T arm

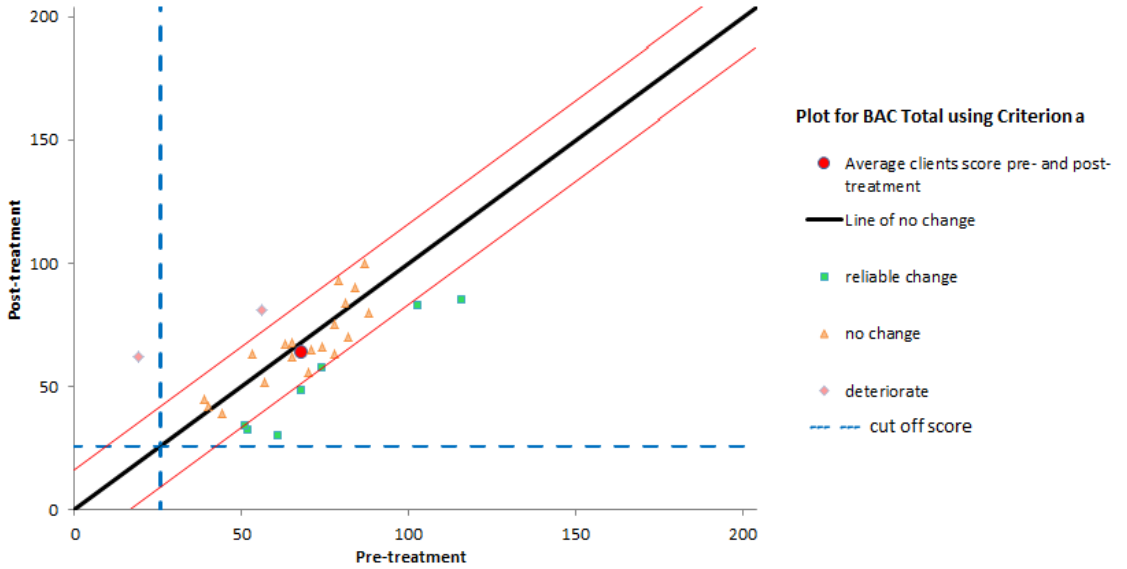


Figure 17: BAC-R at T2 changefor participants in the WL/TAU arm

RCI values and CSC outcomes for the AD S&T and WL/TAU arms are presented below in Table 22 for all the outcome measures except for the PCL-C which is presented in Table 21. The total outcomes of each measure, in respect of AD S&T participants, are also graphically presented in the Appendix 3.8. These outcomes were all undertaken comparing individual change from T1 to T2 scores.

No differences in the CSC and RCI outcomes between the two arms were noted. The exception to this was the HADS *Anxiety* subscale. In this respect it was noted that 36.4% of AD S&T participants achieved a positive improvement in their ability to manage anxiety based symptoms compared to 2.9% of participants in the WL/TAU arm (OR 18.86 [95%CI, 2.28 to 155.86]). Although statistical significance was not achieved in the previous LMM analyses for this subscale it gives support to the larger mean differences for the S&T arm compared to the WL/TAU arm in both the AD and ITT analyses. These results also indicate that no participants produced positive change with reference to the RCI on the HADS *Depression* subscale again further contextualising other statistically significant results involving this subscale (0.0% vs 11.8%)

Table 22. Clinically significant and reliable change for AD S&T and WL/TAU arms at post intervention

	Arm[1]	T1-T2 ES (*d*)	SEM	RCI Value	No. of Change None	-	+	No. meeting CSC	CSC Criteria Used
Behavioural Assessment Checklist-Revised									
Bellig'ence	AD S&T	-0.18	1.55	4.28	19	4	2	0	A
	WL/TAU	0.08	1.26	3.48	23	2	3	0	(-4.93)
Distress	AD S&T	-0.12	1.47	4.06	17	5	3	0	A
	WL/TAU	0.29	1.41	3.90	16	2	10	0	(-1.30)
Withdrawal	AD S&T	0.08	1.83	5.06	25	0	0	0	A
	WL/TAU	0.10	1.78	4.98	19	4	5	0	(0.86)
Impulsivity	AD S&T	-0.46	2.41	6.67	15	9	1	0	A
	WL/TAU	0.22	2.38	6.59	21	2	5	0	(-1.27)
Egocen'city	AD S&T	-0.41	1.63	4.53	16	8	1	0	A
	WL/TAU	0.01	1.61	4.47	24	3	1	0	(-1.95)
Problem	AD S&T	0.37	2.94	8.16	21	1	3	0	A
solving	WL/TAU	0.01	2.74	7.60	27	0	1	0	(11.11)
Total BAC-R	AD S&T	-0.25	5.73	15.87	17	5	3	0	A
	WL/TAU	0.18	5.90	16.34	19	2	7	0	(25.70)
								No external cut off utilised	
DERS									
Non-Accept	AD S&T	0.38	2.67	7.41	27	1	5	1	C
	WL/TAU	-0.04	2.41	6.68	25	5	4	0	(12.74)
Goals	AD S&T	0.18	1.64	4.54	27	2	4	2	C
	WL/TAU	0.35	1.82	5.05	27	1	6	3	(13.92)
Impulse	AD S&T	0.33	2.46	6.81	28	0	5	3	C
	WL/TAU	0.08	2.99	8.29	32	0	2	1	(11.56)
Aware	AD S&T	-0.41	2.59	7.18	28	5	0	0	C
	WL/TAU	0.17	2.73	67.56	32	0	2	1	(14.39)
Strategies	AD S&T	0.17	2.85	7.89	28	2	3	1	C
	WL/TAU	0.23	2.79	7.73	26	2	6	4	(16.73)
Clarity	AD S&T	0.10	1.93	5.35	30	2	1	1	C
	WL/TAU	0.15	1.69	4.67	30	2	2	0	(10.58)
DERS	AD S&T	0.18	7.78	21.56	23	3	7	3	C
Summary	WL/TAU	0.20	7.75	21.47	26	1	7	1	(79.79)
								No external cut off utilised	
DES II									
Amnesia	AD S&T	0.13	6.10	16.90	27	2	4	1	C
	WL/TAU	-0.04	6.19	15.40	30	3	1	1	(9.27)
Absorption	AD S&T	0.18	4.63	12.83	23	2	8	5	C
	WL/TAU	-0.06	4.66	12.93	23	6	5	4	(22.34)
Deperson'	AD S&T	0.19	6.94	19.25	29	1	3	0	C
	WL/TAU	-0.05	5.99	16.90	31	3	0	0	(10.90)
Total DES II	AD S&T	0.17	5.54	15.35	28	1	4	1 [2]*	C
	WL/TAU	-0.06	5.41	14.98	30	3	1	1	(17.89)
								**External set at 30*	
HADS									
Anxiety	AD S&T	0.61	1.02	2.83	19	2	12	2	*
	WL/TAU	0.01	1.54	4.26	33	0	1	0	

Table 22 continued

	Arm[1]	T1-T2 ES (*d*)	SEM	RCI Value	No. of Change			No. meeting CSC	CSC Criteria Used
Depression	AD S&T	-0.49	1.20	3.60	27	6	0	0	*
	WL/TAU	0.04	1.06	2.94	26	4	4	0	
HADS Total	AD S&T	-0.01	1.09	3.02	25	4	4	0	C
	WL/TAU	0.04	0.97	2.69	21	6	7	0	(9.82)
				**External cut off: Depression and Anxiety set at 7*					
Criminal Cognitions Scale									
Short-term	AD S&T	-0.20	1.93	5.36	32	1	0	0	A
	WL/TAU	0.11	1.80	4.99	33	1	0	0	(5.85)
Entitlement	AD S&T	-0.12	1.41	3.90	27	4	2	0	A
	WL/TAU	0.05	1.32	3.67	28	3	3	0	(4.56)
Respons'ity	AD S&T	0.03	1.55	4.29	31	1	1	1	A
	WL/TAU	0.10	1.76	4.89	33	0	1	0	(6.47)
Authority	AD S&T	0.12	1.52	4.23	31	2	0	0	A
	WL/TAU	-0.06	1.45	4.02	33	1	0	0	(6.70)
Insensiti'ty	AD S&T	0.11	1.45	4.03	31	1	1	0	A
	WL/TAU	0.24	1.67	4.63	32	0	2	0	(-4.12)
Reparation	AD S&T	-0.24	2.29	6.34	27	3	3	0	A
	WL/TAU	0.05	2.90	8.01	32	1	1	0	(-4.86)
Summary	AD S&T	-0.11	5.96	16.51	30	2	1	0	A
CCS	WL/TAU	0.11	6.21	17.21	32	1	1	0	(57.56)
								No external cut off utilised	

Notes: WL/TAU= Wait Listed/ Treatment as Usual control arm; AD S&T= Adequate Dose Survive & Thrive experimental intervention arm. SEM (Standard Error of Measurement). [1]AD S&T: N= 33; WL/TAU: N= 34. Participant numbers for the BAC-R were lower than in the other outcome measures: AD S&T N=25; WL/TAU N=28. T1- T2 and T1- T3 Effect Size (ES). SEM: Standard Error of Measurement. RCI: Reliable Change Index. No. of [participants who] Change: none i.e. 'no change', - i.e. 'deteriorate', + i.e. 'improved'. CSC: Clinically Significant Change. Abbreviated items: BAC-R- *Belligerence*, *Egocentricity*; DERS- *Non-Acceptance*; *Impulsivity*, *Awareness*, DES II – *Depersonalisation*; CCS – *Responsibility*, *Insensitivity*

4.9. Summary

This chapter sought to analyse the outcomes from female prisoners randomly assigned into either a group based psychoeducational intervention (Survive & Thrive) designed to stabilise interpersonal trauma symptomatology or to usual care. As noted in the subsequent analyses there were few statistically significant differences between the two arms at T2 (i.e. 'post' treatment) or T3 (i.e. 'follow up' one month after treatment). Statistical modelling techniques were used to analyse the study's arm by time interaction for both ITT and AD data. The following summary seeks to provide answers to the research questions established prior to the trial.

RQ 1. *Will S&T be an efficacious intervention for promoting behavioural and emotional stability as associated with survivors of interpersonal trauma compared to a wait list control group (i.e. usual care) in a prison setting?*

It was noted that there appeared to be no association in the outcomes achieved by the two main outcome measures: the PCL-C and the BAC-R. In this respect PCL-C outcomes appear to favour the S&T arm whilst the BAC-R outcomes favour the WL/TAU control arm. This was particularly apparent when comparing potential differences between AD and Non-AD participants in the S&T arm using these two measures. Whilst, this is discussed more in the next chapter it is important to consider both the internal and external validity of the BAC-R and the difficulty of assessing behavioural stabilisation through staff observational ratings.

RQ 2. *Will S&T be an efficacious intervention for stabilising symptoms associated with PTSD and compared to a wait list control group in a prison setting?*

Whilst mean differences in PCL-C outcomes at T2 and T3 indicated larger reductions for the S&T arm than the WL/TAU arm these differences were not statistically significant. This indicates that only relatively small differences in symptom amelioration or stability were achieved by the intervention arm relative to usual care.

However, from indirect comparisons using clinical and reliable change analyses differences between the two arms were highlighted. With respect to the PCL-C total score 30.3% of participants in the AD S&T arm achieved positive reliable deductions between T1 and T2. This suggests that twice the number of AD S&T participant achieved a reduction in PTSD symptom reduction than would otherwise have been expected from usual care alone (OR 2.03 [95%CI, 0.64 to 6.43]). However, this difference in symptom reduction, in term of reliable change, does not appear to have been maintained at T3 and post hoc analyses should be treated with caution.

RQ 3. *Will S&T be an efficacious treatment for stabilising general symptoms of psychopathology compared to a wait list control group in a prison setting?*

As highlighted there were few statistically significant differences between the arms across all outcome measures. Exceptions to this were noted for the DERS *Non-Acceptance* subscale and the HADS *Depression* subscale. These outcomes suggest that

participants in the S&T arm had reduced levels of non-acceptance and higher rates of depression than the WL/TAU arm. In addition, when coding time as a categorical variable it was also possible to demonstrate that the statistically significant differences between the two arms seemed to be most evident between T2 and T3.

RQ 4. Will S&T be a more efficacious treatment for those participants who receive an 'adequate dose' compared to a waitlist control group in a prison setting?

A particularly large difference between AD and Non-AD S&T participants was noted with respect to outcomes using the PCL-C measure. This indicated that the few participants (n=11) who received < 7 sessions seemed to have higher levels of PTSD symptomology than those who received an adequate dose (n= 33).

The reliable and clinical change analysis in this study was computed using AD S&T participants. As such it was noted that AD S&T participants achieved more positive change in the HADS *Anxiety* subscale and more WL/TAU participants achieved positive change on the HADS *Depression* subscale. However, the magnitude of these differences does not seem to correspond to those obtained from direct statistical comparisons.

Further consideration needs to be given to these outcomes and caution should be exercised with regard to the number of statistical tests undertaken. *A priori* statistical power calculations initially indicated that this trial required n=70 participants. However, an intra-class correlational coefficient (ICC) was used to take into account the anticipated variance inflation associated with group based interventions. This indicated that an n=118 or greater would be required to achieve sufficient statistical power. Whilst, the statistical modelling techniques used (LMM) will also have improved statistical power caution should be given to the results, particularly those from the BAC-R, which should be considered underpowered.

5. Discussion

It's like teaching someone to swim and then throwing them into a stormy sea.

Eldar Shafir
(Bregman, 2016)

5.1. Summary of the thesis

This thesis was undertaken during a time of increased awareness of the impact that adverse childhood experiences and interpersonal trauma can have on survivor's lives. The Scottish Government, like many other progressive Governments, has increasingly sought to develop an effective response to this endemic public and mental health concern (Scottish Government, 2017). These concerns relate to the long term effect that interpersonal trauma has on survivors physical and mental health as well as their overall life chances and circumstances (Dugal, Bigras, Godbout, Belanger, 2016; Kessler, Aguilar-Gaxiola, Alonso et al, 2017; Magruder, McLaughlin & Borbon, 2017). The extremely high prevalence of interpersonal trauma in female prison populations, up to 91% in a recent study, is an important example of this (Karatzias, Power, Woolston, Apurva, Begley et al 2018). In this respect, interpersonal trauma within forensic populations has been associated with violent offending, substance misuse and difficulties mediating levels of interpersonal distress (Moretti, Odgers Reppucci, Catherine, 2011; Sadeh & McNiel, 2015; Howard, Karatzias, Power, & Mahoney, 2017).

A recent increased awareness of this situation has been particularly evident within the female prison estate where I worked as a Psychology Manager and where I have sought to develop services that responded thoughtfully and appropriately to these endemic levels of interpersonal trauma (Mahoney, 2012; Mahoney & Karatzias, 2012; Mahoney, Chouliara, Karatzias, 2015; Karatzias et al, 2018; Howard, Karatzias, Power & Mahoney, 2017; Howard, Karatzias, Power & Mahoney, 2017a). This included introducing psychoeducational interventions to the female prison estate, such as *Survive & Thrive*, that were considered as state of the art and potentially to have good levels of responsivity.

Psychoeducational interventions have been recommended as part of a helpful phased based approach from which interpersonal safety and emotional stability can be promoted (UKPTSD, 2017). It has been argued this an essential pre-requisite for survivors recovery (Ford, Cortois, Steele, van der Hart & Mijenhuis, 2005; Harris & Fallot, 2001; Herman, 1992). However, it was also apparent that there were substantial gaps in the evidence base concerning the efficacy of interpersonal trauma interventions and particularly for group based interventions undertaken with offenders. An initial review of the prison based literature revealed only small low quality randomised control trials (RCTs) for group based psychoeducational interventions or those based on interventions that had a dual focus on PTSD and substance use disorder. As such there were no adequately powered and controlled RCTs for group based psychoeducation interventions for interpersonal trauma with female prisoners from which to make appropriate recommendations for practice.

The primary objective of this thesis was therefore to investigate the efficacy of psychoeducational group based interventions for the stabilisation of complex interpersonal trauma symptoms in incarcerated female offenders. This objective was considered within two substantive pieces of work. The first piece of work was an important first step in establishing the evidence base for group treatments and consisted of a systematic review and meta-analysis. The second piece of work was an RCT investigating the efficacy of Survive & Thrive, as a prototypical psychoeducational intervention, within a prison setting. Survive & Thrive, which had been piloted successfully in other treatment settings (Ball, Karatzias, Mahoney, Ferguson, & Pate, 2013; Karatzias, Ferguson, Chouliara, Gullone, Cosgrove, Douglas, 2014), was considered a theoretically sound and pure psychoeducational group based intervention that could enable the efficacy of such approaches to be established.

5.1.1. Meta-analysis findings

One of the important findings of the meta-analysis undertaken for this thesis was the small effect sizes that psychoeducational group based interventions achieved. This included those outcome domains (PTSD, Psychological Distress and Depression) which had the largest number of studies. This conclusion was reached after outliers (i.e. Garland et al, 2016) were accounted for. Presenting synthesised results both with and

without this outlier was central in assisting the reader's understand of how such studies impacted on the findings. Indeed, this situation is indicative of the heterogeneity found throughout the meta-analysis. This also enabled an indirect comparison between psychoeducational and trauma memory processing (TMP) interventions which establish that the later were potentially more efficacious (when compared to usual care). This finding has important implications, particularly when considering the outcomes of this thesis's RCT, and the decisions that service providers might make when deciding to deliver either a TMP or a psychoeducational intervention.

However, direct comparisons between psychoeducation and TMP interventions indicated that there was no evidence that one approach was more effective than the other in the trials that compared them. This represents a key finding from the meta-analysis particularly as small, statistically non-significant, effect sizes were computed in favour of psychoeducational interventions for the Depression and Psychological Distress domains. Similarly, equally efficacious results were computed for trauma group based interventions compared to non-trauma group based interventions. These results would seem to concur with previous findings in terms of the lack of superiority for TMP and trauma specific interventions for individuals with complex interpersonal trauma (Foa, McLean, Zang et al, 2018; Greger, Munder & Bath, 2014). These direct comparisons have more statistical credibility than the indirect findings and seem to challenge the theoretical basis of current group based psychological treatments on offer to survivors of interpersonal trauma.

This meta-analysis was also the first to make a distinction between different psychoeducational interventions that have been designed as phase 1 responses to interpersonal trauma. In this respect, the term *Psychoeducation Plus* was used to differentiate interventions of increased intensity and focus. When compared to usual care Psychoeducational Plus interventions were more efficacious and computed medium to large effect sizes suggesting that that the design and focus of psychoeducational interventions is highly important.

Another particularly notable finding of this meta-analysis was the numerous 'unclear' Risk of Bias ratings given to many of the included RCTs due to a lack of reported detail.

This included a lack of reporting on important methodological considerations such as randomisation as well other trial administration and statistical procedures. These difficulties were particularly apparent in the Substance Misuse and the Dissociation domains. It was also noted that there were a lack of studies assessing behavioural stability, a core treatment endeavour of phase 1 interventions, which the RCT for this thesis sought to address. In addition, the meta-analysis for this thesis also helped inform and put into context the efficacy that could be expected from the Survive & Thrive (S&T) trial.

5.1.2. RCT findings

Following the results of the systematic review and meta-analysis, an RCT was conducted to explore the effectiveness of group based psychoeducational interventions for behavioural and emotional stability in female prisoners. Results from the RCT illustrated that there were few statistically significant differences between the S&T and the WL/TAU control arms across the outcome measures. In this respect S&T did not demonstrate superiority to usual care and the results from this trial do not support the use of this intervention with female offenders. As such in the ITT analysis the only statistically significant difference between the arms were with respect to subscales which indicated a rise in depression and emotional acceptance for participants in the S&T arm.

A priori tests also explored the impact of adequate dose (AD) (set at ≥ 7 sessions) in the S&T arm. Whilst the number of non-AD participants was very small (n=11), results from the main outcome measures suggested that there were important statistical differences when compared with those that had completed the intervention. Subsequent differences were apparent in the AD versus WL/TAU analysis with larger mean differences computed for the intervention arm in the PCL-C *Intrusion* and *Avoidance* subscales. However, these differences in favour of AD S&T participants were represented by small non-statistically significant effect sizes with wide confidence intervals. This again suggests that even AD participants were unlikely to significantly benefit from S&T.

Further, *post hoc* Clinically Significant Change (CSC) and Reliable Change (RC) analyses, using clinically accepted cut-off points, identified that 30.3% of participants in the AD

S&T arm (versus 17.6% in the WL/TAU arm) achieved positive 'reliable' reductions in PTSD symptomatology at post treatment. This indicated that twice the number of AD participants benefited from an increase in PTSD symptom reduction than would otherwise have been achieved by usual care alone (30.3% vs 17.6%, OR 2.03 [95%CI, 0.64 to 6.43]). However, such *post-hoc* analyses should not be given greater weight than the neutral results from the ITT analyses and are suggestive only of future avenues of investigation.

Similarly, the ITT analysis for the S&T arm at one month follow up (i.e. T3) indicated that the mean PCL-C total score had reduced and was approaching the diagnostic cut-off point. This diagnostic cut–off point was set at 45, as suggested for 'speciality mental health clinics' (US Veterans Affairs, 2012). In this respect a 5 point reduction, as noted the PCL-C total score for the ITT analysis and also in the AD analysis, is thought to be suggestive of reliable change (Blanchard, Jones-Alexander, Buckley & Forneris, 1996; Monson, Gradus, Young-Xu, Schnurr, Price & Schumm, 2008, US Veterans Affairs, 2012). A conclusion could therefore be reached that this is reasonable performance for a psychoeducational intervention particularly given the other trauma informed benefits that delivering such an intervention could bring to an institution.

In contrast to the outcomes obtained from the PCL-C small statistically non-significant outcomes in favour of the WL/TAU control arm were computed using the BAC-R (behavioural stability measure). These outcomes were as a result of the apparently negative impact that the S&T intervention had on participant's behaviour. As noted in the AD analysis statistically significant differences between the arms were noted specifically for the BAC-R *Distress* subscale; but only one month after treatment. This suggests that the increase in distress was not immediately apparent upon treatment completion. Similarly, the contrast between the BAC-R and PCL-C outcomes also indicates that any amelioration or stabilisation in PTSD symptoms was internalised and that such processes maybe expressed in negative behavioural changes.

Similarly, and although not main outcomes, statistically significant differences were apparent between the S&T and WL/TAU control arms for the DERS *Non-Acceptance* subscale and the HADS *Depression* subscale in both the ITT and AD analyses. It was

also demonstrated that these statistically significant differences became most evident one month after treatment. This suggests that these responses weren't perhaps immediately apparent on treatment completion or at least intensified over time. However, again it suggests a potentially deleterious impact from participating in S&T. Indeed, improvements in terms of a greater acceptance of emotional functioning may be part of a temporal sequence involving increased behavioural *Distress* (BAC-R subscale) and *Depression* (HADS subscale) as observed within these results (Medrano & Trogolo, 2016).

Another finding was the data from the HADS *Anxiety* subscale suggested the direction of effect favoured the S&T intervention for reducing feelings of anxiety. Whilst these findings were not statistically significant in both the ITT and AD analyses the *post hoc* RC analysis highlighted that 36.4% of participant in the S&T arm (versus 2.9% in the WL/TAU arm; OR 18.86 [95%CI, 2.28 to 155.86]) achieved reliable positive change at post treatment. It could therefore be hypothesised from these results that the techniques used in S&T to down regulate emotional distress are useful although perhaps not at a level which is indicative of clinical change (as only 6.1% of the AD S&T arm achieved CSC for this subscale).

5.2. Comparisons with previous research

The results of the RCT conducted for this thesis correspond with those from the meta-analysis in respect of the small statistically non-significant effect sizes computed in favour of psychoeducational interventions when compared to usual care. This similarity was particularly the case with respect to PTSD outcomes, although not for *Depression* outcomes, in the ITT analysis. Similarly, the meta-analysis conducted for this thesis did not consider anxiety outcomes separately from symptoms associated with overall psychological distress. This RCT also provides further information with respect to a very limited evidence base (k=1 study) in the meta-analysis for the efficacy of group based psychoeducation on dissociative symptoms compared to usual care. As with the meta-analysis the outcomes from the RCT suggests that it is important for treatment providers to have realistic expectations about the extent of change that can be expected from psychoeducational interventions alone.

Arguably the outcomes from this thesis's meta-analysis and RCT also concur with other recent meta-analytic reviews that have found all psychotherapeutic responses generally promote recovery in PTSD symptoms (Erford et al 2016; Lenz, 2018). Obviously, consideration needs to be given as to how active the usual care (i.e. the TAU component) was in this RCT's WL/TAU arm. In this respect, previous research has suggested that non-specific interventions are equally efficacious particularly for individuals with complex clinical presentations (Greger et al, 2014). Foa et al (2018) recent large scale RCT, comparing the effectiveness of individually delivered TMP and 'present centred' interventions, reported no significant differences. There are of course important differences between Foa et al (2018) and this RCT; not least when comparing active duty military personal to multiply traumatised female offenders from socially disadvantaged backgrounds. As such, avoiding a simple 'one size fits all' approach to interpersonal trauma and its treatment is important (Cloitre, 2015).

The initial literature review undertaken for this thesis identified five previous RCTs with prison based female populations. Whilst there are considerable variations in the treatment protocols employed by these studies arguably the most similar to the current trial was Ford et al (2013). This was also the only previous psychoeducational intervention as the other studies were classified in the meta-analysis as either Psychoeducation Plus or TMP interventions. As with this RCT, Ford et al (2013) noted a slight increase in negative mood for their experimental arm.

However, the only statistically significant difference that Ford et al (2013) identified was an increase in the experimental arm's 'forgiveness' (described as the 'self-perceived ability to forgive self and others for transgressions') which the authors link to the emotional resolution of past interpersonal trauma. Whilst there might be some parallels to the significant increase in emotional acceptance as demonstrated by the S&T arm in this RCT these measures ultimately have a very different theoretical basis and the reliability of Ford et al (2013) outcomes can easily be questioned.

The RCTs that were classified as Psychoeducation Plus (Messina, Grella, Cartier & Torres, 2010; Zlotnick, Johnson & Najavtis, 2009) in the meta-analysis were based on interventions that sought to address co-occurring substance use disorder. Both these studies involved trials in well-established treatment settings for substance misuse

which formed the usual care WL/TAU control arms. In this respect for both these studies participants in the trauma informed arms and the WL/TAU control arms improved significantly on measures of PTSD, SUD and psychological wellbeing. These outcomes of course are similar to those obtained in this RCT particularly with respect to the lack of statistically significant differences in the main outcome measures.

As such the work undertaken for this thesis concurs with other RCT studies investigating the efficacy of psychoeducational interventions and the limited effectiveness they have demonstrated. However, this is also the first time that CSC and RC analyses have been undertaken and it may be possible that a more idiographic approach to analysing outcomes is required. Positive reliable change was demonstrated in this thesis's RCT for PTSD and anxiety based symptomology. Other relevant outcomes in terms of increased emotional acceptance and depression were also observed for the first time.

5.3. Summary of contributions to the literature

This thesis has provided the most comprehensive overview to date of group based interventions for interpersonal trauma (Sloan, Feinstein, Gallagher, Beck, & Keane, 2013; Barrera, Mott, Hofstein, & Teng, 2013). This includes investigating the different treatment outcomes that can be expected from different phase based treatments. The meta-analysis also highlighted the potential increased efficacy that Psychoeducation Plus interventions might achieve. Similarly, the purity of the psychoeducational model that constitutes Survive & Thrive also helped to establish the efficacy that can be expected from such brief group based interventions. As Pelekis & Dahl (2005) note, brief short-term group psychotherapy has been a major part of the treatment offered to survivors of interpersonal trauma. The comprehensive assessment and statistical analysis undertaken within the trial established that small non-significant effect sizes are associated with psychoeducational interventions. However, increases in depression and emotional acceptance as distinct outcomes were apparent which had not previously been evidenced. It is therefore hoped that a major contribution of this thesis is in how psychoeducational interventions are conceptualised and the impetus it provides to improving the psychological treatments offered to survivors including those in prison.

5.4. Clinical practice implications

Prevalence of interpersonal trauma versus symptom expression

The baseline measures used in the RCT draw attention to the high rates of abuse reported across the lifespan for women in Scottish prisons. These outcomes provide greater detail to those already reported by previous studies with this population and in the international literature (Bowen, Jarrett, Stahl, Forrester & Valmagia, 2018, Fazel, Ramesh, Hawton, 2017; Karatzias et al, 2018; MacDonald, 2013). Baseline characteristics also demonstrated a high prevalence of violent offending which further suggests an association between this type of offending and experiences of interpersonal trauma (Howard, Karatzias, Power & Mahoney, 2017).

However, this study also indicates that a potentially important issue when working with this highly complex population is a reduction or underreporting of current trauma symptoms. The SIDES-SR authors note that in their prior clinical experience an absence of a formal PTSD diagnosis is not uncommon for individuals with histories of chronic traumatisation (Luxenberg et al, 2001 pg. 384). This is thought to be a possible outcome of severe avoidance and suppression of trauma memories often through self-medication as seen in substance abuse. In addition, associated effects of numbing and over developed dissociative mechanisms may mask the expression of symptoms that would otherwise have met diagnostic criteria. Other forms of psychopathology may also have masked symptoms that would have otherwise been attributed to complex PTSD, for example, paranoid ideation, perceptual and ongoing relational disturbances. Prison management and psychologists therefore have a unique role to play in developing a greater understanding of possible symptom expression and provide environments which are both supportive and help to address these concerns.

Increases in distress and depression

It is of course important to consider statistically significant differences in the increase of *Distress* (BAC-R subscale) and *Depression* (HADS subscale) as seen in the S&T arm. Obviously, an important question that any clinician needs to consider is whether there is the potential for adverse effects and the harm from the interventions they deliver (Berk & Palmer, 2009; Crawford, Thana, Farquharson, Palmer, Hancock, Bassett,

Clarke & Parry, 2016). For example, there is the potential for any sort of cognitive or behavioural dissonance to occur that results in adverse behaviours (Linden, 2013). As such, elevated mood or distress from those undergoing challenging treatments and personal change processes need to be supported appropriately by all staff involved in participant's day to day interactions and in other treatment planning processes (Harris & Fallot, 2001).

Given that participants in the S&T arm reported an increase in the *Depression* HADS subscale, it is important to give some consideration to whether trauma informed treatments, particularly psychoeducational interventions lead to significant, or even temporary, rises in depression and/or negative affect more generally. Certainly, both arms continued to return mean scores above the cut-off range considered as clinically meaningful for both HADS subscales. These mean scores would indicate that overall at post treatment, participants in both arms reported scores indicating 'moderate' levels of depression (Stern, 2014).

It is therefore important that clinicians and facilitators are aware of the potential increase in negative affect from participating in trauma informed interventions. As Ratcliffe (2018) notes: 'changes in the structure of interpersonal experience are central to most of those predicaments labelled as depression'. Such changes could include how an individual understands their interpersonal (i.e. abusive) experiences. Whilst there is considerable recognition of the comorbidity of PTSD and depression symptoms there is little recognition of phase 1 trauma treatments resulting in increases in low mood (Heim, Newport, Mletzko, Miller & Nemeroff, 2008).

It can also be hypothesised that increases in emotional acceptance may be linked to increases in depression. As Gratz & Roemer (2004) describe, emotional acceptance as having important implications for emotional regulation and for it being essential in reducing maladaptive coping responses. In this respect, participants in the S&T arm may have developed a greater acceptance of their emotional experience and overall this may have important benefits for their recovery from interpersonal trauma. However, an increased awareness and acceptance of the impact that interpersonal trauma has had may also be a trigger for feelings of low mood and depression. This would hopefully be a temporary situation and improve as participants' progress

further in their recovery. It is also important to recognise that this has been regarded as an essential part of the remembrance and mourning process that Herman (1992) has described. Herman (1998) articulates this difficult situation as:

The descent into mourning is a necessary but dreaded part of the recovery process. Patients often fear that the task is insurmountable, that once they allow themselves to start grieving, they will never stop.

(Herman, 1998)

Depression and indeed Major Depressive Disorders (MDD) have also long been associated with interpersonal trauma and as a complicating factor within the recovery process (Bedard-Gilligan, Jakob, Doane, Jaeger, Eftekhari, Feeny & Zoellner, 2015). In this respect, an increase in depression within the S&T arm may also be an indication of an increased ability to identify, label and describe emotional experiences. However, facilitators need to be prepared that this situation may also represent the further activation of cognitive structures common to both PTSD and MDD such as rumination (Angelakis & Nixon, 2015). It has also argued that MDD might also present a phenotype or possibly a subtype of interpersonal trauma (Flory & Yehuda, 2015). If this is the case then consideration should be given as to whether this is more prevalent in forensic populations.

Dropout rates were similar to other prison based studies (Bradley & Follingstad, 2003; Wolff et al, 2015). In this respect, the challenges of conducting research within prisons have long been commented upon. These challenges include the management processes involved in release and inter-prison transfer, which in this study impacted on participant retention and subsequent statistical power (Tully et al, 2014). These 'real world' difficulties and the complexities involved in implementing PTSD interventions are further exacerbated in populations with complex presentations such as violent behaviours and comorbid depression. Such complexities are generally associated with high dropout rates and make establishing an effective evidence base challenging (Flory & Yehuda, 2015; Najavatis, 2015).

High dropout rates in trauma-focused treatments per se have been attributed to an initial symptom increase linked to accessing trauma memories which then activates avoidance mechanisms (Schottenbauer, Arnkoff, Tendick & Gray, 2008). As such there,

are indications that present centred therapies, such as psychoeducation, may have lower dropout rates (Imel, Laska, Jakcupcak & Simpson, 2013). In this respect, it could be hypothesised that higher dropout rates would have been expected in this study if a TMP intervention had been selected. However, the reason for participant dropout can be multifaceted and not necessarily to do with worsening of symptoms (Fernandez, Salem, Swift, Ramtahal, 2015; Szafranski, Smith, Gros, Resick, 2017). Similarly, interventions such as S&T have been designed to target avoidance, which is considered to be an essential part of the DSM-5 diagnosis for PTSD. These well-established difficulties along with elevated levels of depression for the S&T arm suggests that phase 1 interventions should specifically target the depression associated with interpersonal trauma and develop strategies that help manage the fear response associated with the potential activation of traumatic memories (see also Nixon & Nearmy, 2011).

An alternative hypothesis to the increase in depression experienced by those in the S&T arm is one linked to treatment termination and a lack of further service provision. As this significant interaction was only noted at one month following treatment completion, it might be equally conjectured that attachment and loss based issues are being reflected in the HADS Depression subscale (Dagnino, Pérez, Gómez, Gloger & Krause, 2017; Mangione & Forti, 2018). As such, treatment endings are an important consideration for service providers.

Improvements in anxiety management

As noted in the AD analysis as well as the RC analysis the HADS *Anxiety* subscale noted improvements for the S&T arm. Whilst not evident in the more robust ITT analysis, this nevertheless indicates that those who completed the majority of the intervention (defined as completing at least 7 sessions) were likely to derive some benefit in managing anxiety based symptoms. PTSD has of course in the past been classified as an anxiety disorder. Arguably many of the treatment approaches for PTSD and complex trauma are focused on therapeutic techniques that help survivors down regulate symptoms of distress (Pai, Suris & North, 2017; Hyland, Shevlin, Fyvie & Karatzias, 2018). S&T with its inclusion of relaxation and brief mindfulness based skills is no exception to this and it would appear that this is reflected in this thesis RCT

results. Whilst any improvements in the management of relevant symptomology are welcomed, it is also important for a wide range of clinically significant improvements to be realised. A greater focus in relational functioning may also lead to wider and more robust outcomes (Ford et al, 2013, Mahoney et al, 2012).

Rehabilitation of offending behaviour

Prior to the trial, there was some initial concern expressed by staff at the RCT test sites about the impact that a trauma interventions might have on addressing criminogenic cognitions and attitudes. This included concern at how trauma based narratives and awareness might influence offence focused rehabilitative processes (Heide & Solomon, 2006). Such early concerns included the impact Survive & Thrive might have on promoting feelings of guilt and shame (Mahoney et al, 2015). Increasingly, it has become recognised that addressing trauma histories is an essential component to helping individuals manage their risk of future offending (Jones, 2015, Sandhu, 2017; Welfare & Hollin, 2011). However, from the CCS outcome scores, which were very consistent over time, it is apparent that criminogenic cognitions were not influenced by Survive & Thrive. This would fit with Survive & Thrive's overall design and treatment targets.

It is also important to note that this trial did not investigate the interaction effect of participating in a psychoeducation intervention with subsequent treatment outcomes resulting from an offending behaviour intervention. This is something that psychologists and relevant staff should consider in the future.

Prison as a stabilising environment

As already noted one of the findings with respect to the baseline measures used to determine the severity of symptoms associated with complex interpersonal trauma (SIDES-SR) was the reduced number of current symptoms compared to the lifetime experience of symptoms reported by participants. This was most noticeable in the symptom domain of alterations in affect regulation, self-perception and meaning. Similarly, very few participants reported the full range of complex interpersonal trauma symptoms as described by the authors of the SIDES-SR (Luxenberg, Spinazzola, van der Kolk, 2001; Pelcovitz, van der Kolk, Roth, Mandel, Kaplan, & Resick, 1997).

Parallels can perhaps be drawn with the lack of substantive change reported across the measures for both arms in the RCT. These outcomes may be reflecting the prisons potential role as a stabilising environment and a place of safety for individuals from otherwise very chaotic community and family based environments.

The Scottish Prison Service (SPS), where the RCT for this thesis was conducted, has promoted a 'whole system approach' that seeks to ensure the welfare of those in its care (SPS, 2014). This has included promoting a trauma informed environment particularly in the female prison estate. This may have provided a relatively active and potentially sophisticated TAU which constituted the control arm of this study. This may have been particularly the case where multi-agency working practices have led to a number of community based partner agencies and various therapists operating within the study sites. This presents challenges in terms of ascertaining the effectiveness of any particular intervention and isolating possible active components. Whilst the RCT endeavoured not to include individuals undertaking other trauma informed treatments, it was difficult to enforce this after enrolment. Certainly, the environmental impact of prisons, stability or otherwise, is one that has received considerable attention and needs to be carefully considered in future research (Blauw, Roozen, Van Marle, 2007; Hassan, Birmingham, Harty et al, 2011, van Ginneken, 2016).

The future development of phase 1 interventions

It is important to note that the outcomes of this study should not be interpreted as brief trauma based psychoeducation interventions being an irrelevant component within a suite of trauma informed interventions. From the experience of staff involved in delivering S&T it was evident that many participants did not elect to engage in further TMP interventions or were liberated shortly after their involvement in the study. This makes brief interventions potentially the only viable option for a large number of prisoners and the overall therapeutic and institutional climate in prisons may indirectly benefit from the inclusion of trauma informed interventions.

However, on the basis of the statistical evidence, i.e. mean change scores and effect sizes, which have been produced from the work undertaken for this thesis it is not possible to recommend that S&T, or indeed any standard CBT based psychoeducation

intervention, as an effective stand-alone treatment. Delivering a TMP based package of care may led to increased PTSD symptom amelioration but as of yet this has not been tested within a prison population using robust RCT procedures. It is, therefore, recommended that improvements are made to S&T, or indeed a new set of protocols designed, that seek to replicate some of the larger effect sizes seen for Psychoeducation Plus interventions within this thesis's meta-analysis. In this respect, developing protocols that address depression and other emotional difficulties may be particularly relevant to the maladaptive behaviours apparent within forensic populations. It is also important that, as with any psychoeducational intervention, realistic expectations are set as to their effectiveness and how the treatment needs of specific participants map on to specific treatment outcomes.

5.5. Strengths and limitations

The greatest strength of this thesis, and particularly the RCT, was to evaluate the efficacy of what can be considered as a pure psychoeducational intervention, uncontaminated from any processing elements or indeed attempting to alleviate any other co-occurring disorders. This enabled the trial to focus on establishing the treatment efficacy that can be expected from psychoeducation interventions alone which has important implications when helping to develop future treatment protocols.

As noted from the systematic review this is also the first RCT to be undertaken in the UK for the treatment of interpersonal trauma using a group treatment modality. It is also the first substantive RCT of a pure psychoeducational intervention to be undertaken in a prison setting. As such this RCT has importance for custodial treatment providers in the UK particularly as interpersonal trauma has been overlooked in their efforts to reduce recidivism (Jones, 2015). Using a gold standard research approach also avoided conclusions being reached that might simply have been based on a regression towards the mean (Linden, 2013; Morton & Torgerson, 2003). Research undertaken with more complex populations, as is the case in forensic settings, might be particularly vulnerable to this. Indeed, previous non-randomised trials with S&T had shown greater efficacy for forensic populations (Ball et al 2013; Karatzias et al, 2012).

Another strength of this thesis's RCT was that it used material that had been deigned in the community and specifically adapt this for use within a prison context and also delivered within pragmatic parameters (Patsopoulo, 2011). For example, it was noted from the meta-analysis that many RCTs in this field are designed and delivered by either an intervention's author or by other highly experienced experts. The present RCT used prison based facilitators and psychologists, which is the existing treatment delivery model in the SPS. The delivery of S&T was also embedded within the overall treatment delivery schedule for the duration of the study. As such, S&T was condensed or 'massed' so that sessions were delivered over 5 weeks (twice per week) instead of being delivered over 10 weeks (once per week) as usual in community settings. This has pragmatic implications as condensed interventions are more likely to be available to a greater number of short term offenders, although it may impact on the time participants have to learn new skills (i.e. 'homework'). In this respect the implications in having insufficient time to complete action plans and skills practices, an essential aspect of any CBT intervention, needs to be considered as a potential limitation and as an important deviation from the original S&T model (Kazantis, Brownfield, Usatoff & Flighty, 2017). The use of the S&T participant booklet, an important prompt in helping participants implement new skills and changes, was not measured. The recent trial by Foa et al (2018) would however question this assertion as they found no difference between 'massed' and 'spaced' treatment deliveries. Similarly, whether self-directed skills practice is regularly undertaken in prison settings is under investigated and maybe questioned.

Another strength of this study is the range of outcomes that were tested for. This sought to provide a comprehensive overview of the impact that might be expected from a psychoeducational intervention. This was followed by a detailed analysis of the subscales as well as the overall scores from the outcome measures used. For example, the HADS subscales returned relatively different outcomes with the *Anxiety* subscale favouring the S&T treatment arm and *Depression* subscale favouring the control arm. This subscale analysis was essential as differences were evened out in the HADS total score. It is also therefore reasonable to suggest that a synthesis of outcomes from different studies using different measures for psychological distress, as undertaken in the meta-analysis, may have potentially obscured key differences in treatment

efficacy. Using a range of outcomes also enabled this study to contribute towards the relatively sparse evidence base for dissociative symptoms and behavioural outcomes.

However, it is important given the extensive analysis of subscales used in the RCT that type 1 and other measurement errors are kept in mind (Barnett, van der Pols & Dobson, 2004). As mentioned, one feature of this RCT is the lack of substantive change across the assessment time points in respect of both arms for the main outcome measures. Similarly, where change did occur this cannot be directly attributed to S&T as this appears to be most significant in the assessment period one month after treatment.

This trial sought to address concerns regarding previous psychoeducation prison based RCTs being under powered. The original calculations set the sample size at n=70 which when taking into account the variance inflation factor for a group intervention (with a minimal ICC of 0.01) would require a sample size of n=118. As such given the dropout rate experienced in this study a larger sample size than was possible to recruit for this study would be needed, particularly for a psychoeducational intervention, to demonstrate suitable public health gains (Ali, Rhodes, Moreea, McMillan, Gilbody, Leach, Lucock, Lutz & Delgadillio, 2017; Hazell, Hayward, Cavanagh & Strauss, 2016). The sample size available to this RCT will therefore have restricted it to detecting only large group differences and this should be considered particularly when reflecting on some of the small to moderate differences discussed (Button, Ioannidis, Mokrysz, Nosek, Robinson & Munafo, 2013; Hughes, 2018). In this respect, it was also noted that many of the studies included in the meta-analysis had considerably smaller sample sizes. Given this situation results will need to be replicated in a definitive trial with greater power.

With respect to ensuring adequate power within the RCT analysis it is realised that the sporadic return rates from staff of the BAC-R measure presents a serious weakness in the interpretation of this measure. Due to substantial amounts of missing data we used MI conservatively with the BAC-R analyses which resulted in an imbalance between the arms with substantially more participants in the S&T arm having a BAC-R measure. Results from the BAC-R should therefore, along with the other difficulties inherent with this measure, be interpreted with extreme caution.

The decision to use the BAC-R was taken in the context of limited available options suited to the task of measuring behavioural stabilisation. Institutional records concerning self-harm and misconduct were considered too blunt for this purpose. Similarly, it was acknowledged from the outset that using measures such as the BAC-R are fraught with difficulties and efforts were made to guard against these difficulties (McDougal, Perry, Clarbour, Bowles & Worthy, 2009). This study however appears to have been no exception and it is noticeable that the results from the BAC-R contrast with those from other measures used. For example, the subscales from the BAC-R and the DERS for *Impulsivity* suggest different outcomes. This study also did not test for inter-rater reliability using the BAC-R and therefore again outcomes from this measure should be treated with particular caution.

It is also recognised that the RCT should have more clearly defined expected outcomes in terms of 'stabilisation' versus symptom amelioration. Arguably the goal of phase 1 interventions with respect to achieving suitable levels of resourcing and resilience (i.e. 'increased capacity to self-regulate and tolerate distress') prior to accessing phase 2 interventions could also be more specifically defined (UKPTS, 2017 pg. 14). It may therefore be more appropriate to conceptualise emotional stability and indeed behavioural stability as long term trait based outcome (Ho, Cheung, You, Kam, Kliewer, 2012). These difficulties in defining stabilisation, particularly for an offending population, are an important limitation with respect to the inclusion criteria used in this study. Although participants had experienced substantive levels of interpersonal trauma many individuals (as noted in the SIDES-SR) reported substantial differences in current versus lifetime levels of traumatic stress at recruitment. In this respect, very little is known about subthreshold CPTSD presentations within various populations (Franklin, Raines, Chambliss, Walton & Maieritsch, 2018). This current stability and inclusion of non-symptomatic survivors may have made it difficult for the intervention to demonstrate sufficient efficacy and indeed may have also limited further increases in depression (Zlotnick, Franklin & Zimmerman, 2002).

The use of linear mixed models (LMM) in the outcome analysis was essential in accounting for missing data and for optimising power. It should also be noted that whilst LMM will process MI data we considered it more appropriate to use the FIML procedure whilst using this statistical approach. In comparing the outcomes for the

both MI and FIML, both of which have similar advantages, it was concluded that the later offered more consistency with the original data (Jakobsen, Gluud, Wetterslev & Winkel, 2017).

Another distinct advantages of using LMM as the basis for statistical analysis in the RCT was that is that it measured change at the idiosyncratic rather than nomothetic level. LMM assumes that each participant has a unique intercept, or 'baseline' thereby accommodating individual change trajectories (Koerner & Zhang, 2017). As such to consider symptom improvement, particularly with respect to complex interpersonal trauma, as a discrete series of monochronic experiences may be an anathema. Therefore, considerable variations in respect to symptom expression, chronicity, stage of recovery and access of professional and other idiosyncratic sources of support may make any one-size fits all intervention difficult to evaluate. Similarly, given that there may be a variety of different presentations within a disorder such as complex trauma it may be important to ensure qualitative approaches are also used to investigate outcomes.

The specific limitations of the meta-analysis have been discussed previously in that chapter. However, it is also important to acknowledge the limitations of the RCT administration and management. An important limitation of this RCT, and indeed one that may studies within prison settings have to manage, is the attrition caused by participants being released back to the community (see also Ford et al, 2013). Whilst only individuals thought to have had sufficient time in prison were included, it was very difficult to gain access to information on decisions likely to have been made by the Home Detention Curfew (electronic tagging) process. In this respect, it was noted on a number of occasions (n=5) decisions to grant community access (i.e. release from prison) occurred after randomisation but prior to the commencement of assessments at T1. This, attrition occurred to participants randomised to the WL/TAU control arm as decisions to grant Home Detention Curfew orders were made on the basis that the individuals concerned were not actively engaged in treatment. Other difficulties included, as previously mentioned, the numerous different interventions and therapists that constituted the study sites usual care provision.

Other RCT administration difficulties included blinding the Research Assessor (RA) and participants and a lack of independent randomisation. Whilst efforts were taken to ensure concealed randomisation and blinding of the Principle Investigator and RA, it was not possible to ensure that assessments were undertaken within blinded conditions. Similarly, whilst various random factors and attrition may explain the increased mean sentence length in the WL/TAU arm compared to the S&T arm other biases relating to the RCT administration weakness need to be considered. Given the lack of other demographic and sentencing differences between the arms as well as the overall results of this trial it can be reasonably assumed that these had a minimal impact on the outcomes. Nevertheless, further testing would explore whether sentence length was an important covariate.

5.6. Directions for future research

The results of the RCT as well as the meta-analysis highlight the need for strengthening psychoeducational interventions if they are to be efficacious and capable of reducing emotional and behavioural symptoms associated with interpersonal trauma. The overall intensivity of brief psychoeducational interventions, such as S&T, needs consideration if individuals with long standing issues associated with emotional dysregulation are to be appropriately assisted and results produced that are better than usual care (Bohus et al, 2013).

The results of this thesis suggest that designing and testing specific interventions for specific populations have merit. As noted in meta-analysis studies in the category *Psychoeducational Plus*, had larger effect sizes than psychoeducational studies compared to usual care. This would suggest the need to design specific interventions for specific symptoms. The RCT results also suggest that helping participants to manage depression and negative affect, often associated with increased emotional acceptance and recovery would also have merit. One possible means of improving the efficacy of brief interventions might be to increase the mindfulness and introspective awareness components of treatment protocols. As seen in the meta-analysis there is some indication, albeit in respect of one outlier that needs further investigation, that mindfulness based interventions might be considerably more efficacious than CBT based psychoeducational interventions. Mindfulness has been established in several

reviews to be an efficacious treatment for depression and it would make sense that this is a useful avenue for future development and testing (Khoury, Lecomte, Fortin, Masse, Therien, Bouchard, Chapleau, Paquin & Hofmann, 2013).

Given the extent of phase 1 interventions being delivered as standalone interventions it is important to consider whether this is appropriate (de Jongh, Resick, Zoellner, van Minnen, Lee, Monson, Foa, et al 2016). Although it has been argued that some survivors may choose to exit treatment after a phase 1 intervention, it is also important to ensure that the opportunity to participate in a phase 2 intervention is extended to them. In this respect considering the efficacy of phase 1 and phase 2 treatments together, as compared to a phase 1 treatment on its own, would help provide a valuable insight into the recovery process. This also has important implications for the development of brief interventions where the inclusion of even limited TMP protocols could increase an interventions efficacy and be a more effective use of time (Bradley & Follingstad, 2003). Such a development would also concur with the increased effect sizes as evident in this thesis's meta-analysis for TMP interventions when interventions where compared to usual care.

In addition, it is also important to compare and contrast group and individual treatment modalities when seeking to establish the relative efficacy, and indeed effectiveness, of both approaches. It was noted from the meta-analysis that only one study (Stalker & Fry, 1999) had specifically compared these treatment modalities. Nevertheless, conducting and synthesising high quality RCT studies, for both individual and group based treatment modalities, remains an important endeavour in psychological trauma research.

There are a number of important questions that this study did not have scope to answer. This includes establishing the clinical and treatment differences between offenders and non-offenders with experiences of interpersonal trauma. As such important questions include what adaptations, if any; should be made to standard psychoeducational and TMP treatment protocols and whether these differ between male and female offenders. Consideration should also be given to identifying difficulties that some, perhaps an important minority of survivors, experience when participating in any trauma intervention and the additional support needed based on

schema or personality type. Other parallel and equally important questions include: what leads some survivors to offend against others and whether developing different interventions for distinct trauma profiles would be more effective? There is some indication, for example, that cognitive styles such as rumination and the externalisation of distress could be important correlates that need additional consideration in forensic populations (Kaplan, Palitsky, Carey, Crane, Havens, Medrano, Reznik, Sbarra & O'Connor 2018; Krunger & Eaton, 2015). This might have important implications for developing trauma informed offending behaviour interventions for those reporting minimal symptoms and more intensive trauma strategies for those with increased symptoms.

Whilst the randomisation process ensured that both arms of the RCT had similar baseline and demographic characteristics a noticeable difference was in sentence length with the WL/TAU control arm having longer sentences. As noted in the Results chapter this situation can be at least partially explained by a slight increase in the number of participants with Murder/Culpable Homicide offences in the WL/TAU arm. Nevertheless, it would be potentially useful to investigate the role of sentence length either as a covariate in any subsequent analysis or in studies further exploring behavioural and emotional stabilisation in prisons. This could have important implications both for the treatment of interpersonal trauma as well as for sentencing and rehabilitation options (Mews, Hillier, McHugh & Coxon, 2015).

Given the limitations inherent in the BAC-R it can be argued that further work is needed to develop a reliable measure of assessing the stabilisation of participant's behaviours. It is particularly difficult for research within forensic settings to distinguish between behaviours that might be purely driven by criminogenic cognitions as opposed to those driven by psychological distress. The BAC-R may have an inherent bias towards measuring institutionally disruptive behaviours without assessors having regard for less obvious behaviours related to psychological distress. Indeed, other psychometrics such as those measuring changes in motivation, self-efficacy, increased compassion based responses and coping might have more relevance than the BAC-R (Burlingame et al, 2003; Di Clemente, Schlundt, & Gemmell, 2004; Herman, 1992; McCrone et al, 2005; Mendelsohn et al, 2011).

In addition, the continuous assessment of learning and retention of psychoeducational material should be measured in future evaluations. This may also provide a more robust measure of quality assurance than what was undertaken for the current RCT (Rees, Norsworthy & Rowlands, 2009). Whilst it was beyond the scope of the current study, a qualitative evaluation of how effective S&T was perceived to be by participants would be of considerable benefit in establishing areas of development for S&T (Abbott, DiGiacomo, Magin & Hu, 2018; Sayer, Friedemann-Sanchez, Spoont, Murdoch, Parker, Chiros, & Rosenheck, 2009).

5.7. Overall conclusions

The primary objective of this thesis was to answer the important question: are psychoeducational interventions effective for the stabilisation of trauma symptomatology in a prison setting for female offenders? Results from this thesis, both from the meta-analysis and the RCT, with respect to Survive & Thrive (S&T), would support the assertion that it is a misnomer to expect such interventions to substantively ameliorate trauma related symptoms or associated dysfunctional behaviours. As such it was not possible from the results of this trial to demonstrate that Survive & Thrive was a useful intervention when delivered to female offenders.

The following research questions sought to provide more focus to the primary objective:

RQ. 1: *Will S&T be an efficacious intervention for promoting behavioural and emotional stability as associated with survivors of interpersonal trauma compared to a wait list control group (i.e. usual care) in a prison setting?*

It would appear that S&T was not more efficacious than usual care. From the AD analysis there was an indication that completing S&T may be associated with increased levels of distress. However, despite adjustments made to the main behavioural outcome measure, missing data and potentially inter-rater reliability may have impacted on the integrity of the data collected for this study. Similarly, the small non-significant gains that were made on the PTSD symptom amelioration seemed to be internalised and did not translate into any substantive behavioural change.

RQ 2. *Will S&T be an efficacious intervention for stabilising symptoms associated with PTSD compared to a wait list control group in a prison setting?*

Small, non-significant, effect sizes in favour of S&T reducing PTSD symptoms compared to usual care were evident. Post hoc analyses comparing pre and post treatment scores suggest that approximately a third of S&T participants who received an adequate dose achieved reductions in PTSD symptomatology. This equated to twice as many participants than would otherwise have been achieved by usual care alone. These reductions did not equate to clinically significant symptom amelioration and post hoc results require further investigation.

RQ 3. *Will S&T be an efficacious treatment for stabilising general symptoms of psychopathology compared to a wait list control group in a prison setting?*

There was no unequivocal statistical evidence for this. Results from outcomes measuring changes in general psychopathology indicate statistically significant decreases in non-acceptance of emotional states and potentially a corresponding increase in depression for S&T participants. These changes in affect can be accounted for theoretically and anecdotally as indicative of participants starting to emotionally process distressing experiences. It was also noted from post-hoc analyses that over a third of AD S&T participants made some improvements in managing experiences of anxiety; this result requires further exploration.

RQ 4. Will S&T be a more efficacious treatment for those participants who receive an 'adequate dose' compared to a waitlist control group in a prison setting?

The greatest treatment efficacy was noted for those participants who had achieved an adequate dose (≥ 7 sessions). This was particularly noted when comparing AD with non-AD participants in the S&T arm. However, limited statistical power makes such conclusions only tentative.

RQ 5. *What is current efficacy of psychoeducational group based interventions in the stabilisation of trauma associated symptomatology in comparison to various control conditions (i.e. usual care as well as trauma and non-trauma interventions)?*

As noted from this thesis's meta-analysis and the RCT the evidence does not support the efficacy of psychoeducational group based interventions compared to usual care. Notwithstanding the low quality of many RCTs as seen in the meta-analysis, it is suggested that the evidence base for psychoeducational group interventions is moving away from a position of clinical equipoise. Similarly, results from the meta-analysis suggest that such difficulties are also apparent when trauma and non-trauma interventions are compared. As such new and creative approaches are needed to help survivors stabilise symptoms and behaviours associated with complex interpersonal trauma (Cook & Sheets, 2011; London, 2017).

6. References

Abbott, P., DiGiacomo, M., Magin, P., & Hu, W. (2018). A scoping review of qualitative research. International Journal of Qualitative Methods, 17, 1–15.

Methods Used With People in Prison

Aki, E. A., Shawwa, K., Kahale, L. A., Agoristsas, T. et al (2015). Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide. BMJ Open, 5, e008431.

Alexander, P. C., Neimeyer, R. A., Follette, V. M., Moore, M. K., & Harter, S. (1989). A comparison of group treatments of women sexually abused as children. Journal of Consulting and Clinical Psychology, 57, 479−483.

Ali, S., Rhodes, L., Moreea, O., McMillan, D., Gilbody, S., Leach, C., Lucock, M., Lutz, W., & Delgadillo, J. (2017). How durable is the effect of low intensity CBT for depression and anxiety? Remission and relapse in a longitudinal cohort study. Behaviour Research and Therapy, 94, 1-8.

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington, DC, Author.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC, Author.

Andrews, D. A., & Bonta, J. (2010). The psychology of criminal conduct (5th ed.) Cincinnati, OH: Anderson.

Angelakis, S., & Nixon, R.D.V. (2015). The comorbidity of PTSD and MDD: implications for clinical practice and future research. Behaviour Change, 32(1), 1–25.

Baldwin, S. A., Murray, D. M., & Shadish, W. R. (2005). Empirically supported treatments or type I errors? Problems with the analysis of data from group-administered treatments. Journal of Consulting and Clinical Psychology, 73, 924–935.

Ball, S., Karatzias, T., Mahoney, A., Ferguson, S., & Pate, K. (2013). Interpersonal trauma in female offenders: a new, brief, group intervention delivered in a community based setting. Journal of Forensic Psychiatry & Psychology, 24(6), 795-802.

Barrera, T. L., Mott, J. M., Hofstein, R.F., & Teng, E. J. (2013). A meta-analytic review of exposure in group cognitive behavioral therapy for posttraumatic stress disorder. Clinical Psychology Review, 33, 24–32.

Bass, J. K., Annan, J., Murray, S., Kaysen, D., Griffiths, S., Cetinoglu, T., Wachter, K., Murray, L. K., & Bolton, P. A. (2013). Controlled Trial of Psychotherapy for Congolese Survivors of Sexual Violence. New England Journal of Medicine, 368, 2182-91.

Bedard-Gilligan, M., Jakob, J. M.D, Doane, L.S., Jaeger, J., Eftekhari, A., Feeny, N., & Zoellner, L.A. (2015). An investigation of depression, trauma history, and symptom severity in individuals enrolled in a treatment trial for chronic PTSD. Journal of Clinical Psychology, 71(7), 725–740.

Beck, J. G., Coffey, S. F., Foy, D. W., Keane, T. M., & Blanchard, E. B. (2009). Group cognitive behavior therapy for chronic posttraumatic stress disorder: An initial randomized pilot study. Behavior Therapy, 40, 82–92.

Benish, S.G., Imel, Z. E. & Wampold, B. E. (2008). The relative efficacy of bona fide psychotherapies for treating post-traumatic stress disorder: A meta-analysis of direct comparisons. Clinical Psychology Review, 28, 746–758.

Berk, M. & Palmer, G. (2009). The elephant on the couch: side-effects of psychotherapy. Australian and New Zealand Journal of Psychiatry, 43, 787794.

Berke, D., S., Macdonald, A., Poole, G. M., Portnoy, G. A., McSheffrey, S., Creech, S., K. & Taft, C. T., (2017). Optimizing trauma-informed intervention for intimate partner violence in veterans: The role of alexithymia. Behaviour Research and Therapy, 97, 222-229.

Bernstein, E. M., Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. Journal of Nervous Mental Disorders, 174, 727–735.

Bilić, V., Nemčić-Moro, I., Karšić, V., Grgić, V., Stojanović-Špehar, S., & Marčinko, D. (2010). Clinical experiences in treating PTSD patients by combining individual and group psychotherapy. Psychiatria Danubina, 22(2), 173–178.

Blaauw, E., Roozen, H. G., Van Marle, H. G. C., (2007). Saved by structure? The course of psychosis within a prison population. International Journal of Prisoner Health, 3(40), 248-256.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C. (1996). Psychometric properties of the PTSD Checklist (PCL). Behaviour Research and Therapy, 34, 669-673.

Bisson, J., & Andrew, M. (2007). Psychological treatment of post-traumatic stress disorder(PTSD). Cochrane Database of Systematic Review(3) Art. No.: CD003388.

Bisson, J.I., Ehlers, A., Matthews, R., Pilling, S., Richards, D., & Turner, S. (2007). Psychological treatments for chronic post-traumatic stress disorder. Systematic review and meta-analysis. British Journal of Psychiatry, 190, 97–104.

Bisson, J.I., Roberts, N.P., Andrew, M., Cooper, R., & Lewis, C. (2013). Psychological therapies for chronic post-traumatic stress disorder (PTSD) in adults (Review). Cochrane Database of Systematic Reviews 2013, Issue 12. Art. No.: CD003388.

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C. (1996). Psychometric properties of the PTSD Checklist (PCL). Behaviour Research and Therapy, 34, 669-673.

Blake, D.D., Weathers, F.W., Nagy, L.M., Kaloupek D., Klauminzer, G., Charney, D.S., Keane, T.M., & Buckley, T.C. (2000). Clinician-Administered PTSD Scale (CAPS) Instruction Manual. Boston, National Centre for PTSD.

Bohus, M., Dyer, A.S., Priebe, K., Krüger, A., Kleindienst, N., Schmahl, C., et al. (2013). Dialectical behaviour therapy for post-traumatic stress disorder after childhood sexual abuse in patients with and without borderline personality disorder: A randomised controlled trial. Psychotherapy and Psychosomatics, 82, 221–233.

Borenstein, M., Hedges, L. V., Higgins, J.P.T., & Rothstein, H.R. (2009). Introduction to Meta-Analysis. Chichester, John Wiley & Sons.

Borenstein, M., Hedges, L. Higgins, J. & Rothstein, H. (2014). Comprehensive Meta Analysis Version 3. Retrieved from www.meta-analysis.com

Bowen, K., Jarrett, M., Stahl, D., Forrester, A., & Valmaggia, L. (2018). The relationship between exposure to adverse life events in childhood and adolescent years and subsequent adult psychopathology in 49,163 adult prisoners: A systematic review. Personality and Individual Differences, 131, 74-92.

Boutron, I., Altman, D.G., Moher, D., Schulz, K.F., Ravaud, P., et al. (2017). CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. Annals of Internal Medicine, 167(1), 40-47.

Bowen, K., Jarrett, M., Stahl, D., Forrester, A., & Valmaggia, L. (2018). The relationship between exposure to adverse life events in childhood and adolescent years and subsequent adult psychopathology in 49,163 adult prisoners: A systematic review. Personality and Individual Differences, 131, 74-92.

Bowman, N. A. (2012). Effect sizes and statistical methods for meta-analysis in higher education. Research in Higher Education, 53, 375-382.

Bradley, R. & Davino, K. (2007) Interpersonal violence, recovery and resilience in incarcerated women. Journal of Aggression, Maltreatment &Trauma, 14(1/2), 123-146.

Bradley, R.G., & Follingstad, D.R. (2003). Group therapy for incarcerated women who experienced interpersonal violence: A pilot study. Journal of Traumatic Stress, 16, 337–340.

Bradley, R., Greene, J., Russ, E., Dutra, L., & Westen, D. (2005). A multidimensional meta-analysis of psychotherapy for PTSD. American Journal of Psychiatry, 162, 214–227.

Bregman, R. (2016). Why do the poor make such poor decisions? Retrieved from: https://thecorrespondent.com/4664/why-do-the-poor-make-such-poor-decisions/179307480-39a74caf

British Psychological Society (2014). Code of Human Research Ethics. Leicester, British Psychological Society.

Brookes, N., Barrett, B., Netten, A., & Knapp,E. (2013). Unit Costs in Criminal Justice (UCCJ). Personal Social Services Research Unit; PSSRU Discussion Paper 2855. Retrieved from: http://www.pssru.ac.uk/archive/pdf/dp2855.pdf

Browne, A., Miller, B., & Maguin, E. (1999). Prevalence and severity of lifetime physical and sexual victimization among incarcerated women. International Journal of Law and Psychiatry, 22(3-4), 301-322.

Brunton-Smith, I., Carpenter, J., & Kenward, M. (2014). Surveying Prisoner Crime Reduction (SPCR) Adjusting for Missing Data Technical Report. Ministry of Justice. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/296931/spcr-missing-data-report.pdf

Burgoyne, K. & Tyson, G. (2013). An evaluation of the 'Think First' program. Australian & New Zealand Journal of Criminology, 46(1), 88–100.

Burlingame, G.M., Fuhriman, A., & Mosier, J. (2003). the differential effectiveness of group psychotherapy: a meta-analytic perspective. Group Dynamics: Theory, Research, and Practice, 7(1), 3–12.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J. & Munafo, M.R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews, 14, 365-376.

Callahan, K. L., Price, J. L., & Hilsenroth, M. J. (2004). A review of interpersonal psychodynamic group psychotherapy outcomes for adult survivors of childhood sexual abuse. International Journal of Group Psychotherapy, 54, 491−519.

Carlson, E. B., & Putnam, F. W. (1993). An update on the Dissociative Experiences Scale. Dissociation: Progress in the Dissociative Disorders, 6(1), 16-27.

Chard, K.M. (2005). An evaluation of cognitive processing therapy for the treatment of posttraumatic stress disorder related to childhood sexual abuse. Journal of Consulting and Clinical Psychology, 73, 965–971.

Cima, M., Merckelbach, H., Hollnack, S., Butt, C., Kremer, K., Schellbach-Matties, R., & Muris, P. (2003). The other side of malingering: Supernormality. The Clinical Neuropsychologist, 17, 235-243.

Cislo, A. M. & Trestman, R. (2013). Challenges and solutions for conducting research in correctional settings: The U.S. experience. International Journal of Law and Psychiatry, 36, 304–310.

Clapp, J.D. & Beck, J. G. (2012). Treatment of PTSD in older adults: Do cognitive-behavioural interventions remain viable? Cognitive and Behavioural Practice, 19(1), 126–135.

Classen, C., Koopman, C., Nevill-Manning, K., & Spiegel, D. (2001). A preliminary report comparing trauma-focused and present-focused group therapy against a waitlisted condition among childhood sexual abuse survivors with PTSD. Journal of Aggression, Maltreatment and Trauma, 4, 265−288.

Classen, C. C., Palesh, O. G., Cavanaugh, C. E., Koopman, C., Kaupp, J. W., Kraemer, H. C., et al. (2011). A comparison of trauma-focused and present focused group therapy for survivors of childhood sexual abuse: A randomized controlled trial. Psychological Trauma: Theory, Research, Practice, and Policy, 3(1), 84.

Cloitre, M. (2015). The ‘‘one size fits all’’ approach to trauma treatment: should we be satisfied? European Journal of Psychotraumatology, 6:1, 27344.

Cloitre, M., Courtois, C. A., Ford, J. D., Green, B. L., Alexander, P., Briere, J., Herman, J.L., Lanius, R., Stolbach, B.C., Spinazzola, J., van der Kolk, B.A., van der Hart, O. (2012). The ISTSS Expert Consensus Treatment Guidelines for Complex PTSD in Adults. Complex Trauma Task Force (CTTF). Retrieved from: https://www.istss.org/ISTSS_Main/media/Documents/ISTSS-Expert-Concesnsus-Guidelines-for-Complex-PTSD-Updated-060315.pdf

Cloitre, M., Garvert, D. W., Brewin, C. R., Bryant, R. A. & Maercker, A. (2013). Evidence for proposed ICD-11 PTSD and complex PTSD: a latent profile analysis. European Journal of Psychotraumatology, 13 (4), 1-12.

Cloitre, M., Koenen, K. C., Cohen, L. R., Han, H. (2002). Skills training in affective and interpersonal regulation followed by exposure: a phase-based treatment for PTSD related to childhood abuse. Journal of Consulting and Clinical Psychology, 70(5), 1067.

Cloitre, M., Stolbach, B. C., Herman, J. L., van der Kolk, B., Pynoos, R., Wang, J., & Petkova, E. (2009). A developmental approach to complex PTSD: childhood and adult cumulative trauma as predictors of symptom complexity. Journal of Traumatic Stress, 22(5), 399-408.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. A.,,Mannarino, A., P., Kliethermes, M., & Murray, L. A. (2012). Trauma-focused CBT for youth with complex trauma. Child Abuse and Neglect, 36(6), 528-41.

Cole, K. L., Sarlund-Heinrich, P., & Brown, L. S. (2007). Developing and assessing effectiveness of a time-limited therapy group for incarcerated women survivors of childhood sexual abuse. Journal of Trauma and Dissociation, 8, 97−121.

Constantino, R., Kim, Y., & Crane, P.A. (2005). Effects of a social support intervention on health outcomes in residents of a domestic violence shelter: a pilot study. Issues in Mental Health Nursing, 26, 575–590.

Cook, C., & Sheets, C. (2011). Clinical equipoise and personal equipoise: two necessary ingredients for reducing bias in manual therapy trials. Journal of Manual and Manipulative Therapy, 19(1), 55-58.

Cooke, D. J. (1998). The development of the Prison Behaviour Rating Scale. Criminal Justice and Behaviour, 25(4), 482-506.

Corrigan, F. M., Hull, A. M. (2015). Neglect of the complex: why psychotherapy for post-traumatic clinical presentations is often ineffective. British Journal of Psychiatry Bulletin, 39(2), 86-89.

Courtois, C. A., Ford, J. D. (2009). Treatment of Complex Traumatic Stress Disorders: Scientific Foundations and Therapeutic Models. New York: The Guilford Press.

Courtois, C. A., Ford, J. D. (2013). Treatment of Complex Trauma: A Sequenced, Relationship-Based Approach. New York: The Guilford Press.

Courtois, C. A., Ford, J. D. (2016). Treatment of Complex Traumatic: A Sequenced, Relationship-Based Approach. New York: The Guilford Press.

Covington, S. S. (2008). Women and addiction: A trauma-informed approach. Journal of Psychoactive Drugs, 5, 377–385.

Covington, S. S. & Bloom, B. E. (2004). Gendered Justice: Women in the Criminal Justice System. Retrieved from: https://www.stephaniecovington.com/assets/files/4.pdf

Covington, S. S., Burke, C., Keaton, S., & Norcott, C. (2008). Evaluation of a traumainformed and gender-responsive intervention for women in drug treatment. Journal of Psychoactive Drugs, 5, 387–398.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2013). Developing and evaluating complex interventions. Medical Research Council. Retrieved from: https://mrc.ukri.org/documents/pdf/complex-interventions-guidance/

Crawford, M. J., Thana, L., Farquharson, L., Palmer, L., Hancock, E., Bassett, P., Clarke, J. & Parryvan, G. D (2016). Patient experience of negative effects of psychological

treatment: results of a national survey. The British Journal of Psychiatry, 208, 260–265.

Crespo, M., & Arinero, M. (2010). Assessment of the efficacy of a psychological treatment for women victims of violence by their intimate male partner. The Spanish Journal of Psychology, 13(2), 849-863.

Cuijpers, P. (2016). Meta-Analyses in Mental Health Research. A Practical Guide. Vrije Universiteit Amsterdam, Colofon. Retrieved from: meta-analysis@vu.nl

Dagnino, P., Pérez, C., Gómez, A., Gloger, S., & Krause, M. (2017). Depression and attachment: how do personality styles and social support influence this relation? Research in Psychotherapy: Psychopathology, Process and Outcome, 20, 53-62.

Davidson, J.R.T., Tharwani, H.M., & Connor, K.M. (2002). Davidson Trauma Scale (DTS): normative scores in the general population and effect sizes in placebo-controlled SSRI trials. Depression and Anxiety, 15, 75–78.

de Jong, T. L., & Gorey, K. M. (1996). Short-term versus long-term group work with female survivors of childhood sexual abuse: A brief meta-analytic review. Social Work with Groups, 19, 19−27.

de Jongh, A., Resick, P. A., Zoellner, L. A., Minnen, A., Lee, C. W., Monson, C. M., Foa, E., Wheeler, K., Broeke E. T., Feeny, N., Rauch S. A, Chard, K. M., Mueser, K. T., Sloan, D. M., van der Gaag, M., Rothbaum, B. O., Neuner, F., de Roos, C., Hehenkamp, L. M., Rosner, R., & Bicanic, .I.A. (2016). Critical analysis of the current treatment guidelines for complex PTSD in adults. Depression and Anxiety, 55 (5), 359-369.38.

Denis, C.M. Cacciola, J. C. & Alterman, A.I. (2013). Addiction Severity Index (ASI) summary scores: comparison of the recent status scores of the ASI-6 and the composite scores of the ASI-5. Journal of Substance Abuse Treatment, 45(5), 44-50.

de Vogel, V., & Nicholls, T. L (2016). Gender Matters: An Introduction to the Special Issues on Women and Girls, International Journal of Forensic Mental Health, 15(1), 1-25.

Devilly, G., & McFarlane, A.C. (2009). When Wait Lists Are not Feasible, Nothing Is a Thing That Does not Need to Be Done. Journal of Consulting and Clinical Psychology, 77(6), 1159–1168.

Di Clemente, C. C., Schlundt, D., & Gemmell, L. (2004). Readiness and stages of change in addiction treatment. American Journal on Addictions, 13(2), 103-119.

Dorrepaal, E., Thomaes, K., Hoogendoorn, A. W., Veltman, D. J., Draijer, N., & van Balkom, Anton J. L. M. (2014). Evidence-based treatment for adult women with child abuse-related Complex PTSD: A quantitative review. European Journal of Psychotraumatology, 5, 1–18.

Dorrepaal, E., Thomaes, K., Smit, J. H., van Balkom, A.J.L.M., Veltman, D.J., Hoogendoorn, A.W., & Draijer, N. (2012). Stabilizing group treatment for complex posttraumatic stress disorder related to child abuse based on psychoeducation and cognitive behavioural therapy: a multisite randomized controlled trial. Psychotherapy and Psychosomatics, 81, 217–225.

Dorrepaal, E., Thomaes, K., Smit, J.H., van Balkom, A.J.L.M., van Dyck, R., Veltman, D.J., et al. (2010). Stabilizing group treatment for complex posttraumatic stress disorder related to childhood abuse based on psycho-education and cognitive behavioral therapy: A pilot study. Child Abuse & Neglect, 34, 284–288.

Dorrepaal, E., Thomaes, K., Smit, J.H., van Balkom, A.J.L.M., Veltman, D.J., Hoogendoorn, A.W., et al. (2012). Stabilizing group treatment for complex posttraumatic stress disorder related to child abuse based on psychoeducation and cognitive behavioural therapy: A multisite randomized controlled trial. Psychotherapy and Psychosomatics, 81, 217–225.

Draycott, S., Kirkpatrick, T. & Askari, R. (2012). An idiographic examination of patient progress in the treatment of dangerous and severe personality disorder: a reliable change index approach. The Journal of Forensic Psychiatry & Psychology, 23:1, 108-124.

Duggan, C., Parry, G., McMurran, M., Davidson, K., & Dennis, J. (2014). The recording of adverse events from psychological treatments in clinical trials: evidence from a review of NIHR-funded trials. Trials. 15:335.

Dunn, G. (2013). Pragmatic trials of complex psychosocial interventions: methodological challenges. Epidemiology and Psychiatric Sciences, 22(2), 105-109.

Ehring, T., Welboren, R., Morina, N., Wicherts, J. M., Freitag, J., Emmelkamp, P. M. (2014). Meta-analysis of psychological treatments for posttraumatic stress disorder in adult survivors of childhood abuse. Clinical Psychology Review, 34(8), 645-657.

Eldridge, S.M., Ashby, D., Feder, G.S., Rudnicka, A.R., & Ukoumunne, O. C. (2004). Lessons for cluster randomized trials in the twenty-first century: a systematic review of trials in primary care. Clinical Trials, 1, 80–90.

Elliott, K. C., Cheruvelil, K. S., Montgomery, G.M. & Soranno, P. A. (2016). Conceptions of good science in our data-rich world. Bioscience, 66 (10), 880-88.

Enlow, M.B., Egeland, B., Blood, E., Wright, R.O. & Wright, R. J. (2013). Interpersonal Trauma Exposure and Cognitive Development in Children to Age 8 Years: A Longitudinal Study, J Epidemiology and Community Health, 66 (11), 1005–1010.

Falsetti, S., Resnick, H., & Davis, J. (2008). Multiple Channel Exposure Therapy for women with PTSD and comorbid panic attacks. Cognitive Behaviour Therapy, 37, 117–130.

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39, 175-191.

Fazel, S., Ramesh, T., & Hawton, K. (2017) Suicide in prisons: an international study of prevalence and contributory factors. Lancet Psychiatry, 4, 946–52.

Fazel, S., Yoon, I. A. Hayes, A. J. (2017). Substance use disorders in prisoners: an updated systematic review and meta-regression analysis in recently incarcerated men and women. Addiction, 112, 1725–1739.

Felitti, V.J., Anda, R.F., Nordenberg, D., Williamson, D.F., Spitz, A.M., Edwards, V., Koss, M.P., Marks, J.S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. American Journal of Preventative Medicine, 14(4), 245-58.

Ferguson, S. (2008, 2013). Survive & Thrive: Treatment Manual. Edinburgh: NHS Lothian.

Fernandez, E., Salem, D., Swift, J.K., et al (2015). Meta-analysis of dropout from cognitive behavioral therapy: magnitude, timing, and moderators. Journal of Consulting and Clinical Psychology, 83(6), 1108–22.

Finn, H., Warner, E., Price, M., & Spinazzola, J. (2018). The boy who was hit in the face: somatic regulation and processing of preverbal complex trauma. Journal of Child and Adolescent Trauma, 11 (3), 277-288.

Flory, J.D., & Yehuda, R. (2015). Comorbidity between post-traumatic stress disorder and major depressive disorder: alternative explanations and treatment considerations. Dialogues in Clinical Neuroscience, 17(2), 141-150.

Foa, E, B., Carmen, P., Zang, Y., Rosenfield, D., Yadin, E., Yarvis, J. S., Mintz, J., Young-McCaughan, S., Borah, E. V., Dondanville, K. A., Fina, B. A., Hall-Clark, B. N., Lichner, T., Litz, B. T., Roache, J., Wright, E. C., Peterson, A. L. (2018). Effect of Prolonged Exposure Therapy Delivered Over 2 Weeks vs 8 Weeks vs Present-Centred Therapy on PTSD Symptom Severity in Military Personnel: A Randomized Clinical Trial. JAMA, 319(4), 354-364.

Ford, J. D., Chang, R., Levine, J., & Zhang, W. (2013). Randomized clinical trial comparing affect regulation and supportive group therapies for victimization-related PTSD with incarcerated women. Behavior Therapy, 44(2), 262–276.

Foa, E, B., Carmen, P., Zang, Y., Rosenfield, D., Yadin, E., Yarvis, J. S., Mintz, J., Young-McCaughan, S., Borah, E. V., Dondanville, K. A., Fina, B. A., Hall-Clark, B. N., Lichner, T., Litz, B. T., Roache, J., Wright, E. C., Peterson, A. L. (2018). Effect of Prolonged Exposure Therapy Delivered Over 2 Weeks vs 8 Weeks vs Present-Centred Therapy on PTSD Symptom Severity in Military Personnel: A Randomized Clinical Trial. JAMA, 319(4), 354-364.

Ford, J. D., Chang, R., Levine, J., & Zhang, W. (2013). Randomized clinical trial comparing affect regulation and supportive group therapies for victimization-related PTSD with incarcerated women. Behavior Therapy, 44(2), 262–276.

Ford, J. D., Cortois, C. A., Steele, K., van der Hart, O., & Nijenhuis, E.R.S. (2005). Treatment of Complex Posttraumatic Self-Dysregulation. Journal of Traumatic Stress, 18(5) 437–447.

Ford, J. D. & Kidd, P. (1998). Early childhood trauma and disorders of extreme stress as predictors of treatment outcome with chronic posttraumatic stress disorder. Journal of Traumatic Stress, 11(4), 743-61.

Ford, J. D., Stockton, P., Kaltman, S., & Green, B. L. (2006). Associated with type and severity of interpersonal trauma exposure in a sample of healthy young women. Journal of Interpersonal Violence, 21(11), 1399-1416.

Fowler, J. C., Carak, R., Elhai, J. D., Allen, J. G., Frueh, J. C., & Oldham, J. M. (2014). Construct validity and factor structure of the difficulties in Emotion Regulation Scale among adults with severe mental illness. Journal of Psychiatric Research, 58, 175-180.

Franklin, C. L., Raines, A. M. Chambliss, J.L., Walton, J. L., & Maieritsch, K.P. (2018). Examining various subthreshold definitions of PTSD using the Clinician Administered PTSD Scale for DSM-5. Journal of Affective Disorders, 234, 256–260.

Freedland, K. E., Mohr, D. C., Davidson, K. W., & Schwartz, J. E. (2011). Usual and unusual care: existing practice control groups in randomized controlled trials of behavioral interventions. Psychosomatic Medicine, 73(4), 323-35.

Friedman, M. J. (2013). Finalising PTSD in DSM-5: Getting here from there and where to go next. Journal of Traumatic Stress, 26, 548- 556.

Frisman, L., Ford, J., Hsiu-Ju, L. Mallon, S., & Chang, R. (2008). Outcomes of trauma treatment using the TARGET Model. Journal of Groups in Addiction & Recovery, 3(3-4), 285-303.

Fritch, A.M., & Lynch, S.M. (2008). Group Treatment for Adult Survivors of Interpersonal Trauma. Journal of Psychological Trauma, 7(3), 145- 169.

Frude, N. Honess, T. & Maguire, M. (2013). CRIME-PICS II Manual. Retrieved from: http://www.crime-pics.co.uk/publications/cpicsmanual.pdf

Gamble C, Krishan A, Stocken D, et al. (2017). Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA, 318(23), 2337–2343

Garieballa, S. S., Schauer, M., Neuner, F., Saleptsi, E., Kluttig, T., Elbert, T., Hoffmann, K., & Rockstroh, B. S. (2006). Traumatic events, PTSD, and psychiatric comorbidity in forensic patients--assessed by questionnaires and diagnostic interview. Clinical Practice Epidemiology in Mental Health, 2:7.

Garland, E.L, Robert-Lewis, A., Tronnier, C. D., Graves, R., & Kelly, K. (2016). Mindfulness-Oriented Recovery Enhancement versus CBT for co-occurring substance dependence, traumatic stress, and psychiatric disorders: Proximal outcomes from a pragmatic randomized trial. Behaviour Research and Therapy, 77, 7-16.

Gene-Cos, N., Fisher, J., Ogden, P. & Cantrel, A. (2016). Sensorimotor psychotherapy group therapy in the treatment of complex PTSD. Annals of Psychiatry and Mental Health 4(6): 1080.

Gerger H., Munder, T., Gemperli, A., Nüesch, E., Trelle, S., Jüni, P., Barth, J., (2014). Integrating fragmented evidence by network meta-analysis: relative effectiveness of psychological interventions for adults with post-traumatic stress disorder. Psychological Medicine, 44(15), 3151-64.

Ghee, A.C., Bolling, L. C., & Johnson, C. S. (2009). The efficacy of a condensed seeking safety intervention for women in residential chemical dependence treatment at 30 days posttreatment. Journal of Child Sexual Abuse, 18, 475–488.

Graham-Bermann, S. A., & Miller, L. E. (2013). Intervention to reduce traumatic stress following intimate partner violence: An efficacy trial of the Moms' Empowerment Program (MEP). Psychodynamic Psychiatry, 41, 329–349.

Grant, S., Mayo-Wilson, E., Montgomery, P., Macdonald, G., Michie, S., Hopewell, S., Moher, D. (2018). CONSORT-SPI 2018 Explanation and Elaboration: guidance for reporting social and psychological intervention trials. Trials, 19:406.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. Journal of Psychopathology and Behavioral Assessment, 26(1), 41-54.

Greene, C.A., Ford, J. D., Wakefield, D.B., & Barry, L. C. (2014). Posttraumatic stress mediates the relationship between childhood victimization and current mental health burden in newly incarcerated adults. Child Abuse & Neglect, 38, 1569–1580.

Greger, H., Munder, T. & Barth, J. (2014) Specific and nonspecific psychological interventions for PTSD symptoms: a meta-analysis with problem complexity as a moderator. Journal of Clinical Psychology, 70(7), 601-615.

Haller, J. (2018). The role of the lateral hypothalamus in violent intraspecific aggression – the glucocorticoid deficit hypothesis. Frontiers in Systems Neuroscience, 12(26): doi: 10.3389/fnsys.2018.00026

Hansen, L., Birmingham, L., Harty, M. A., Jarrett, M., Jones, P., King, C., Lathlean, J., Lowthian, C., Mills, A., Senior, J., Thornicroft, G., Webb, R. & Shaw, J. (2011). Prospective cohort study of mental health during imprisonment. The British Journal of Psychiatry, 198, 37–42.

Harris, M., & Fallot, R. D. (Eds.). (2001). New directions for mental health services. Using trauma theory to design service systems. San Francisco, CA, US: Jossey-Bass.

Hassan, L., Birmingham, L., Harty, M. A., Jarrett, M., Jones, P., King, C., Lathlean, J., Lowthian, C., Mills, A., Senior, J., Thornicroft, G., Webb, R., & Shaw, J. (2011). Prospective cohort study of mental health during imprisonment. The British Journal of Psychiatry, 198, 37–42.

Hazell, C. M., Hayward, M., Cavanagh, K., & Strauss, C. (2016). A systematic review and meta-analysis of low intensity CBT for psychosis. Clinical Psychology Review, 45, 183–192.

Hedges, L.V., Tipton, E., & Johnson, M.C. (2009). Robust variance estimation in meta-regression with dependent effect size estimates. Research Synthesis Methods, 1(1), 39-65.

Heck, R. H., Thomas, S. L., Tabata, L. N. (2014). Multilevel and Longitudinal Modelling with IBM SPSS (2nd ed). NY, Routledge.

Heide, K. M., & Solomon, E. P. (2006). Biology, childhood trauma, and murder: rethinking justice. International Journal of Law and Psychiatry, 29, 220–223.

Heide, K. M., & Solomon, E. P. (2009). Female juvenile murderers: Biological and psychological dynamics leading to homicide. International Journal of Law and Psychiatry, 32, 244–252.

Heim, C., Newport, D. J., Mletzko, T., Miller, A. H. & Nemeroff, C. B. (2008). The link between childhood trauma and depression: Insights from HPA axis studies in humans. Psychoneuroendocrinology, 33, 693–710.

Herman, J. (1992). Trauma and recovery: The aftermath of violence – from domestic abuse to political terror. USA: Basic Books.

Herman, J. (1998). Recovery from psychological trauma. Psychiatry and Clinical Neurosciences, 52(1), S98-S103.

Herman, J. (2007). Shattered Shame States and their Repair. The John Bowlby Memorial Lecture. Retrieved from: https://rvtsmidt.no/wp-content/uploads/2014/03/Herman-Shattered-Shame-2011-2-9.pdf

Herman, J. (2009). Forward. In C. A. Courtoise & J. D. Ford (Eds.) Treating Complex Traumatic Stress Disorder: Scientific Foundations and Treatment Models. New York, The Guilford Press.

Hey, S., Weijer, C., Taljaard, M., & Kesselheim, A. S. (2018). Research ethics for emerging trial designs: does equipoise need to adapt? BMJ, 360:k226.

Hien, D., Cohen, L., Miele, C., Litt, C., & Capstick, C. (2004). Promising Treatments for Women With Comorbid PTSD and Substance Use Disorders. American Journal of Psychiatry, 161, 1426-1432.

Hien, D.A., Levin, F. R., Ruglass, L.M., Lopez-Castro, T., Papini, S., Hu, M., Cohen, L.R. & Herron, A. (2015). Combining Seeking Safety with sertraline for PTSD and alcohol use disorders: a Randomized Controlled Trial. Journal of Consulting and Clinical Psychology, 83(2), 359–369.

Hien, D. A., Wells, E. A., Jiang, H., Suarez-Morales, L., Campbell, A. N. C., Cohen, L. R., et al. (2009). Multisite randomized trial of behavioral interventions for women with co-occurring PTSD and substance use disorders. Journal of Consulting and Clinical Psychology, 77, 607–619.

Higgins J. P. T., Green, S. (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org

Hinton, D. E., Hofmann, S. G., Rivera, E., Otto, M. W., & Pollack, M. H. (2011). Culturally adapted CBT (CA-CBT) for Latino women with treatment-resistant PTSD: A pilot study comparing CA-CBT to applied muscle relaxation. Behaviour Research and Therapy, 49, 275–280.

Ho, Y.M., Cheung, F. M., You, J., Kam, C., Zhang, K., & Kliewer, W. (2013). The moderating role of emotional stability in the relationship between exposure to violence and anxiety and depression. Personality and Individual Differences, 55, 634–639.

Hoge, C. W. & Chard, K. M. (2018). A window into the evolution of trauma-focused psychotherapies for posttraumatic stress disorder. JAMA, 319(4), 343-345.

Hollifield, M., Sinclair-Lian, N., Warner, T. D., & Hammerschlag, R. (2007). Acupuncture for posttraumatic stress disorder. The Journal of Nervous and Mental Disease, 195, 504–513.

Hollis, S. & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ, 319:670, 670-4.

Howard, R., Karatzias, T., Power, K. & Mahoney, A. (2017). Posttraumatic stress disorder (PTSD) symptoms mediate the relationship between substance misuse and violent offending among female prisoners. Social Psychiatry and Psychiatric Epidemiology, 52 (1), 21–25.

Howard, R., Karatzias, T., Power, K., & Mahoney, A., (2017a). From childhood trauma to self-harm: an investigation of theoretical pathways among female prisoners. Clinical Psychology and Psychotherapy, 24(4), 942-951.

Hughes, B. (2018). Does psychology face an exaggeration crisis? The Psychologist, September 2018 issue, 8-10.

Hyland, P., Shevlin, M., Fyvie, C. & Karatzias, T. (2018). Posttraumatic Stress Disorder and Complex Posttraumatic Stress Disorder in DSM-5 and ICD-11: Clinical and Behavioral Correlates. Journal of Traumatic Stress, 31, 174–180.

Imel, Z. E., Laska, K., Jakcupcak, M., & Simpson, T. L. (2013). Meta-analysis of dropout in treatments for post-traumatic stress disorder. Journal of Consulting and Clinical Psychology, 81(3), 394–404.

Jacobson, N. S., Roberts, L. J., Berns, S. B., McGlinchey, J.B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. Journal of Consulting and Clinical Psychology, 67(3), 300-307.

Jacobson, N. S. & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology, 59(1) 12-19.

Johnstone, L. & Boyle, M. (2018). The Power Threat Meaning Framework: Towards the identification of patterns in emotional distress, unusual experiences and troubled or troubling behaviour, as an alternative to functional psychiatric diagnosis. Leicester: British Psychological Society.

Jakobsen, J.C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. BMC Medical Research Methodology, 17:162.

Jones, L. (2010). History of offence paralleling behaviour construct and related concepts. In M. Daffern, L. Jones, J, Shine (Eds.) Offence Paralleling Behaviour: A Case Formulation Approach to Offender Assessment and Intervention. Chichester, John Wiley & Sons.

Jones, L. (2015). The Peaks unit: from a pilot for 'untreatable' psychopaths to trauma informed milieu therapy. The Prison Service Journal, 218, 17-23.

Jüni, P., Altman, D. G., & Egger, M. (2001). Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ, 7, 323(7303), 42-6.

Kaplan, D. M., Palitsky, R., Carey, A., Crane, T. E., Havens, C. M., Medrano, M. R., Reznik, S. J., Sbarra, D. A. & O'Connor, M. (2018). Maladaptive repetitive thought as a transdiagnostic phenomenon and treatment target: An integrative review. Journal of Clinical Psychology, 74(7), 1126-1136.

Kaptchuck, T. J. (1998). Intentional Ignorance: A history of blind assessment and placebo controls in medicine. Bulletin of the History of Medicine, 72.3, 389-433.

Karatzias, T. Ferguson, S., Chouliara, Z., Gullone, A., Cosgrove, K., Douglas, A. (2014). Effectiveness and acceptability of group psychoeducation for the management of mental health problems in survivors of child sexual abuse (CSA). International Journal of Group Psychotherapy, 64 (4), 493- 514.

Karatzias, T., Jowett, S., Begley, A. & Deas, S. (2016). Early maladaptive schemas in adult survivors of interpersonal trauma: foundations for a cognitive theory of psychopathology. European Journal of Psychotraumatology, 7(1), 3071.

Karatzias, T., Power, K., Woolston, C., Apurva, P., Begley, A., Mirza, K. Conway, L., Quinn, C., Jowett, S., Howard, R. & Purdie, A. (2018). Multiple traumatic experiences, post-traumatic stress disorder and offending behaviour in female prisoners. Criminal Behaviour and Mental Health, 28(1), 72-84.

Karatzias, T., Shevlin, M., Fyvie, C., Hyland, P., Efthymiadou, E., Wilson, D., Roberts, N., Bisson, J. I., Brewin, C. R., Cloitre. M. (2016). Evidence of Distinct Profiles of Posttraumatic Stress Disorder (PTSD) and Complex Posttraumatic Stress Disorder (CPTSD) based on the New ICD-11 Trauma Questionnaire (ICDTQ). Journal of Affective Disorders, 207, 181-187.

Kaslow, N.J., Leiner, A.S., Reviere, S., Jackson, E., Bethea, K., Bhaju, J., Rhodes, M., Gantt, M., Senter, M., & Thompson, M.P. (2010). Suicidal, abused African American women's response to a culturally informed intervention. Journal of Consulting and Clinical Psychology, 78 (4) 449–458.

Kazantis, N., Brownfield, N. R., Usatoff, A.S., & Flighty, A. J. (2017). Homework in cognitive behavioral therapy: a systematic review of adherence assessment in anxiety and depression (2011–2016). Psychiatric Clinics of North America, 40, 625–639.

Kelly, A. & Garland, E. L. (2016). Trauma-Informed Mindfulness-Based Stress Reduction for female survivors of interpersonal violence: results from a stage I RCT. Journal of Clinical Psychology, 72(4), 311–328.

Kessler, R. C., Aguilar-Gaxiola, S., Alonso, J., Benjet, C., Bromet, E. J., Cardoso, G., Degenhardt, L., de Girolamo, G., Dinolova, R. V., Ferry, F., Florescu, S., Gureje, O., Haro, J. M., Huang, Y., Karam, E. G., Kawakami, N., Lee, S., Lepine, J., Levinson, D., Navarro-Mateu, F., Pennell, B., Piazza, M., Posada-Villa, J., Scott, K.M. Stein, D. J., Have, M. T., Torres, Y., Viana, M. C., Petukhova, M. V. Sampson, N. A., Zaslavsky, A. M. & Koenen, K. C. (2017). Trauma and PTSD in the WHO world mental health surveys. European Journal of Psychotraumatology, 8(5), 1353383.

Kezelman, C., Hossack, N., Stravropoulos, P., & Burley, P. (2015). The cost of unresolved childhood trauma and abuse in adults in Australia. Blue Knot Foundation. Retrieved from: https://www.blueknot.org.au/Portals/2/Economic%20Report/The%20cost%20of%20unresolved%20trauma_budget%20report%20fnl.pdf

Khoury, B., Lecomte, T., Fortin, G., Masse, M., Therien, P., Bouchard, V., Chapleau, M., Paquin, K., & Hofmann, S. G. (2013). Mindfulness-based therapy: A comprehensive meta-analysis. Clinical Psychology Review, 33, 763–771.

Koerner, T. K., & Zhang, Y. (2017). Application of Linear Mixed-Effects Models in Human Neuroscience Research: A Comparison with Pearson Correlation in Two Auditory Electrophysiology Studies. Brain Sciences, 7 (26), 1-11.

Koltz, R. (2012). The Compassionate-Mind Guide to Managing Your Anger: Using Compassion-Focused Therapy to Calm Your Rage and Heal Your Relationships. Oakland, New Harbinger Publications.

Krakow, B., Hollifield, M., Johnston, L., Koss, M., Schrader, R., Warner, T., Tandberg, D., Lauriello, J., McBride, L., Cutchen, L., Cheng, D., Emmons, S., Germain, A.,

Krueger, R. F., & Eaton, N. R., (2015). Transdiagnostic factors of mental disorders. World Psychiatry, 14 (1), 27-29.

Krupnick, J.L., Green, B.L., Stockton, P., Miranda, J., Krause, E.D., & Mete, M. (2008). Group interpersonal psychotherapy for low-income women with posttraumatic stress disorder. Psychotherapy Research, 18, 497–507.

Kubany, E. S., Hill, E.E., Owens, J. A. (2003). Cognitive trauma therapy for battered women with ptsd: preliminary findings. Journal of Traumatic Stress, 16(1) 81–91.

Lau, M., & Kristensen, E. (2007). Outcome of systemic and analytic group psychotherapy for adult women with history of intrafamilial childhood sexual abuse: A randomized controlled study. Acta Psychiatrica Scandinavica, 116, 96−104.

Lazaratou, H. (2017). Interpersonal trauma: psychodynamic psychotherapy and neurobiology, European Journal of Psychotraumatology, 8, 1351202.

Lenz, A. S. Haktanir, A., & Callender, K. (2017). Meta-analysis of trauma-focused therapies for treating the symptoms of posttraumatic stress disorder. Journal of Counselling & Development, 95, 339-353.

Lewis, C. (2006). Treating incarcerated women: Gender matters. Psychiatric Clinics of North America, 29(3), 773–789.

Lewis, C. F. (2010). Childhood antecedents of adult violent offending in a group of female felons. Behavioral Sciences and the Law, 28, 224–234.

Lenz, A. S. Haktanir, A., & Callender, K. (2017). Meta-analysis of trauma-focused therapies for treating the symptoms of posttraumatic stress disorder. Journal of Counselling & Development, 95, 339-353.

Liebman, R. E., Burnette, M. L., Raimondi, C., Nichols-Hadeed, C., Merle, P. & Cerulli, C. (2014). Piloting a psycho-social intervention for incarcerated women with trauma histories: lessons learned and future recommendations. International Journal of Offender Therapy and Comparative Criminology, 58(8), 894-913.

Lichev, V., & Wolfradt, U., (2016). Alexithymia and depersonalization in child sex offenders: the Role of Emotion Regulation. International Journal of Forensic Mental Health, 15(3), 274-282.

Lilienfeld, S.O., McKay, D., Hollon, S. D. (2018). Why randomised controlled trials of psychological treatments are still essential. The Lancet Psychiatry, 5(7), 536-538.

Linden, A. (2013). Assessing regression to the mean effects in health care initiatives. BMC Medical Research Methodology, 13:119

Linden, M. (2013). How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. Clinical Psychology and Psychotherapy, 20, 286–296.

Linehan, M. M. (1993). Diagnosis and treatment of mental disorders. Cognitive-behavioral treatment of borderline personality disorder. New York, NY, US: Guilford Press.

London, A. J. (2017). Equipoise in research integrating ethics and science in human research. JAMA, 317(5), 525-526.

Loucks, N. (1997). HMPI Cornton Vale: Research into drugs and alcohol,violence and bullying suicides and self-injury, and backgrounds of abuse. Scottish Prison Service Occasional Papers Report No 1/98. Edinburgh, UK: Scottish Prison Service.

Loewenstein, R. & Brand, B. (2014). Treating Complex Trauma Survivors. Psychiatric Times; October 2014, 40-45.

Lubin, H., Loris, M., Burt, J., & Johnson, D. (1998). Efficacy of Psychoeducational Group Therapy in Reducing Symptoms of Posttraumatic Stress Disorder Among Multiply Traumatized Women. American Journal of Psychiatry, 155, 1172-1177.

Luxenberg, T., Spinazzola, J., van der Kolk, B., (2001) Complex Trauma and Disorders of Extreme Stress (DESNOS) Diagnosis, Part One: Assessment. Directions in Psychiatry, 21(25), 373- 393.

Lynch, S. M., Heath, N. M., & Mathews, K. C. & Crepeda, G. J. (2012). Seeking safety: an intervention for trauma-exposed incarcerated women? Journal of Trauma and Dissociation, 13(1), 88-101.

MacDonald, M. (2013). Women prisoners, mental health, violence and abuse. International Journal of Law and Psychiatry, 36, 293-303.

Maercker, A., Brewin, C. R., Bryant, R. A., Cloitre, M., van Ommeren, M., Jones, L. M., et al. (2013). Diagnosis and classification of disorders specifically associated with stress: Proposals for ICD-11. World Psychiatry, 12(3), 198-206.

Magruder, K. M., McLaughlin, K. A., & Borbon, D. L. E. (2017). Trauma is a public health issue. European Journal of Psychotraumatology, 8, 1375338.

Mahoney, A. L. (2012). Working with survivors of sexual abuse and complex trauma in prison. Forensic Update, 106, 47-51.

Mahoney, A., Chouliara, Z., & Karatzias, T. (2015). Substance Related Offending Behaviour Programme (SROBP): an exploration of gender responsivity and treatment acceptance issues for female prisoners. The Journal of Forensic Psychiatry & Psychology, 26(6), 798-823.

Mahoney, A. L., & Karatzias, T. (2012). Violent female offending: an exploration of repeat and one-time offending. International Journal of Forensic Mental Health, 11, 191–202.

Mangione, L. & Forti, R. (2018). Beyond Midlife and Before Retirement: A Short-Term Women's Group. International Journal of Group Psychotherapy 68(3), 314-336.

Maniglio, R. (2009). The impact of child sexual abuse on health: A systematic review of reviews. Clinical Psychology Review, 29, 647–657.

Manigilo, R. (2011). The role of childhood trauma, psychological problems, and coping in the development of deviant sexual fantasies in sexual offenders. Clinical Psychology Review, 31, 748–756.

McCambridge, J., Kypri, K., & Elbourne, D. (2014). In randomization we trust? There are overlooked problems in experimenting with people in behavioral intervention trials. Journal of Clinical Epidemiology, 67, 247e253.

Matuschek, H., Kligl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. Journal of Memory and Language, 94, 305–315.

Mauritz, M.W., Goossens, P.J.J., Draijer, N., & van Achterberg, T. (2013). Prevalence of interpersonal trauma exposure and trauma-related disorders in severe mental illness. European Journal of Psychotraumatology, 4(1), 19985.

Mazzotti, E., Farina, B., Imperatori, C., Mansutti, F., Prunetti, E., Speranza, A., & Barbaranelli, C. (2016). Is the Dissociative Experiences Scale able to identify detachment and compartmentalization symptoms? Factor structure of the Dissociative Experiences Scale in a large sample of psychiatric and nonpsychiatric subjects. Neuropsychiatric Disease and Treatment, 12, 1295–1302.

McCrone, P., Weeramanthri, T., Knapp, M., Rushton, A., Trowell, J., Miles, G., & Kolvin, I. (2005). Cost-Effectiveness of Individual versus Group Psychotherapy for Sexually Abused Girls. Child and Adolescent Mental Health, 10(1), 26–31.

McLellan, A. T., Kushner, H., Metzger, D., Smith, P.R., Grisson, G., Pettinati, H., & Argeriou, M. (1992). The Fifth Edition of the Addiction Severity Index. Journal of Substance Abuse Treatment, 9(3), 199-213.

McDonagh A, Friedman, M., McHugo, G., Ford, J., Sengupta, A., Mueser, K., Demment, C. C., Fournier, D., Schnurr, P. F., Descamps M. (2005). Randomized trial of cognitive–behavioral therapy for chronic posttraumatic stress disorder in adult female survivors of childhood sexual abuse. Journal of Consulting and Clinical Psychology, 73(3) 515–524.

McDougall, C., Clarbour, J. Perry, A. E. & Bowles, R. (2009). Evaluation of HM Prison Service Enhanced Thinking Skills Programme: Report on the implementation of a randomised controlled trial. Ministry of Justice Research Series 4/09. London, Ministry of Justice.

McDougall, C., Clark, D. A., & Woodward, R. (1995). Application of operational psychology to assessment of inmates. Psychology, Crime and Law, 2, 85–99.

McCrone, P., Weeramanthri, T., Knapp, M., Rushton, A., Trowell, J., Miles, G., & Kolvin, I. (2005). Cost-Effectiveness of Individual versus Group Psychotherapy for Sexually Abused Girls. Child and Adolescent Mental Health, 10(1), 26–31.

McLean, L. M., & Gallop, R., (2003). Implications of childhood sexual abuse for adult borderline personality disorder and complex posttraumatic stress disorder. American Journal of Psychiatry, 160(2), 369-71.

McWhirter, P. T. (2011). Differential Therapeutic Outcomes of Community-Based Group Interventions for Women and Children Exposed to Intimate Partner Violence. Journal of Interpersonal Violence, 26(12), 2457–2482.

Meade, C.S., Drabkin, A.S., Hansen, N.B., Wilson, P.A., Kochman A., Sikkema, K.J. (2010). Reductions in alcohol and cocaine use following a group coping intervention for HIV-positive adults with childhood sexual abuse histories. Addiction, 105(11), 1942–1951.

Medrano, L. A., & Trogolo, M. (2016). Construct validity of the Difficulties in Emotion Regulation Scale: further evidence using confirmatory factor analytic approach. Abnormal and Behavioural Psychology, 2:2.

Melberg, H. O. (2004). Three problems with the ASI composite scores. Journal of Substance Use, 9(3-4) 120-126.

Melendrez, D., Sandoval, D., & Prince, H. (2001). Imagery Rehearsal Therapy for chronic nightmares in sexual assault survivors with posttraumatic stress disorder. A randomized controlled trial. JAMA, 286, 537–545.

Mendelsohn, M., Herman, J. L., Schatzow, E., Kallivayalil, D., Levitan, J. & Coco, M. 2011). The Trauma Recovery Group: A Guide for Practitioners. New York, Guilford Press.

Mendelsohn, M., Zachary, R. S., Harney, P. A. (2007). Group Therapy as an Ecological Bridge to New Community for Trauma Survivors. Journal of Aggression, Maltreatment & Trauma, 14(1-2), 227-243.

Merckelbach, H., & Collaris, J. (2012). Mother Teresa doesn't help here: lack of moral priming effects on malingered symptom reports and what we can learn from it. Psychologica Belgica, 52 2-3) 271-285.

Messina, N., Calhoun, S., & Warda, U. (2012). Gender-responsive drug court treatment: A randomized controlled trial. Criminal Justice and Behavior, 39(12), 1536–1555.

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.

Mews, A., Hillier, J., McHugh, M. & Coxon, C (2015). The Impact of Short Custodial Sentences, Community Orders and Suspended Sentence orders on Re-offending. London, Ministry of Justice Analytical Series.

Miller, P. W. (2016). EMDR for Schizophrenia and Other Psychoses. New York, Springer Publishing Company.

Ministry of Justice (2010). Compendium of re-offending statistics and analysis. Ministry of Justice Statistics Bulletin. London, Ministry of Justice. Retrieved from www.justice.gov.uk/publications/docs/compendium-of-reoffending-statistics-andanalysis.pdf. See also: www.justice.gov.uk/.../spcr-sample1-wave1-questionnaire-tables.xls

Mitchell, K. S., Dick, A. M., DiMartino, D., M., Smith, B. N., Niles, B., Koenen, K.C., & Street, A. (2014). A pilot study of a randomized controlled trial of yoga as an intervention for PTSD symptoms in women. Journal of Traumatic Stress, 27, 121–128.

Mohr, D.C., Ho, J., Hart, T.L., Baron, K.G., Berendsen, M., Beckner, V., Cai, X., Cuijpers P., Spring, B., Kinsinger, S.W., Schroder, K. E., Duffecy, J. (2014). Control condition design and implementation features in controlled trials: a meta-analysis of trials evaluating psychotherapy for depression. Translational Behavioural Medicine, 4, 407–423.

Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., Ockene, J., & Kaplan, R. (2009). The Selection and Design of Control Conditions for Randomized Controlled Trials of Psychological Interventions. Psychotherapy and Psychosomatics, 78, 275–284.

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., Elbourne, D., Egger, M., Altman, D.G. (2010). CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010, 340:c869.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G., The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097.

Moloney, K.P., van den Bergh, B.J. & Moller, L.F. (2009). Women in prison: The central issues of gender characteristics and trauma history. Public Health, 123, 426–430.

Monson, C. M., Gradus, J. L., Young-Xu, Y., Schnurr, P. P., Price, J. L., & Schumm, J. A. (2008). Change in posttraumatic stress disorder symptoms: Do clinicians and patients agree? Psychological Assessment, 20(2), 131-138.

Montgomery, P., Grant, S., Mayo-Wilson, E., Macdonald, G., Michie, S., Hopewell, S., Moher, D. (2018). Reporting randomised trials of social and psychological interventions: the CONSORTSPI 2018 Extension. Trials, 19:407.

Moretti, M.M., Odgers, C.L., Reppucci, N.D., The Gender and Aggression Project Team, & Catherine, N.L.A. (2011). Serious conduct problems among girls at risk: Translating research into intervention. International Journal of Child, Youth and Family Studies, 1, 142 -161.

Morley, S., & Dowzer, C. (2014). The Leeds Reliable Change Indicator. University of Leeds. Retrieved from: https://medhealth.leeds.ac.uk/download/2287/manual_for_leeds_rci

Morton, V. & Torgerson, D. J. (2003). Effect of regression to the mean on decision making in health care. BMJ, 326, 1083 – 1084.

Mott, J.M., Sutherland, R. J, Williams, W., Lanier, S.H., Ready, D.J., & Teng, E.J. (2013). Patient perspectives on the effectiveness and tolerability of group-based exposure therapy for posttraumatic stress disorder: preliminary self-report findings from 20 veterans. Psychological Trauma: Theory, Research, Practice, and Policy, 5(5), 453-461.

Murray, D. M., Catellier, D. J., Hannan, P.J., Treut, M. S., Stevens, J., Schmitz, K. H., Rice, J. C., & Conway, T. L. (2004). School-Level Intraclass Correlation for Physical Activity in Adolescent Girls. Medicine & Science in Sports & Exercise, 876- 882.

Najavits, L. M. (2002). Seeking safety: A treatment manual for PTSD and substance abuse. New York, Guilford Press.

Najavatis, L. M. (2015). The problem of dropout from "gold standard" PTSD therapies. Faculty of 1000 Prime Reports. Retrieved from: http://f1000.com/prime/reports/m/7/43

Najavits, L. M., & Hien, D. (2013). Helping vulnerable populations: a comprehensive review of the treatment outcome literature on substance use disorder and PTSD. Journal of Clinical Psychology, 69(5), 433e479.

Nelson, L. D., Simmons, J., & Simonsohn, U. (2017). Psychology's renaissance. Annual Review of Psychology. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3064811

NES (2017). Transforming Psychological Trauma: A Knowledge and Skills Framework for the Scottish Workforce. NHS Education for Scotland. Retrieved from: https://www.nes.scot.nhs.uk/media/3971582/nationaltraumatrainingframework.pdf

NES (2018). The Scottish Psychological Trauma and Adversity Training Plan. NHS Education for Scotland. Retrieved from: https://www.nes.scot.nhs.uk/media/4236974/FINAL REVISED CONSULTATION NESD0715-NATIONAL-TRAUMA-TRAINING-STRATEGY-V7-DL.pdf

Ney, B., Van Voohris, P. & Lerner, G. (2011). Achieving successful outcomes with justice involved women. National Resource Centre for Justice Involved Women (NRCJIW). Presented on 12 July 2011, as part of the BJA NTTAC's Webinar Series. Retrieved from www.bjatraining.org.

NICE (2018). Post –traumatic stress disorder: management (update). Evidence Review D: Evidence reviews for psychological, psychosocial and other non-pharmacological

interventions for the treatment of PTSD in adults. National Institute for Health and Care Excellence. Retrieved from: https://www.nice.org.uk/guidance/gid-ng10013/documents/evidence-review-3.

Nixon R.D.V. & Nearmy, D.M. (2011). Treatment of comorbid posttraumatic stress disorder and major depressive disorder: a pilot study. Journal of Traumatic Stress, 4(4), 451-455.

Nollett, C. Lewis, C., Kitchiner, N., Roberts, V. Addison, K., Brookes-Howell, L., Cosgrove, S., Cullen, K., Ehlers, A., Heke, S., Kelson, M., Lovell, K., Madden, K., McEwan, K., McNamara, R., Phillips, C., Pickles, T., Simon, N.& Bisson, J. (2018). Pragmatic RAndomised controlled trial of a trauma-focused guided self-help Programme versus InDividual trauma focused cognitive Behavioural therapy for post-traumatic stress disorder (RAPID): trial protocol. BMC Psychiatry, 18:77.

Norton, S., Cosco, T., Doyle, F., Done, J., & Sacker, A. (2013). The Hospital Anxiety and Depression Scale: a meta confirmatory factor analysis. Journal of Psychosomatic Research, 74(1), 74-81.

Nugent, F., Geohagan, K., & Travers, R. (2006). Cognitive skills assessment test battery guide (version 16.2.1). London: The Offending Behaviour Programme Unit, HM Prison Service.

Nunnaly, J. C., & Bernstein, I. H. (1994). Psychometric Theory. Sydney: McGraw-Hill.

Olsen, S. A., Clapp, J. D., Parra, G. R., & Beck, J. G. (2013). Factor structure of the dissociative experiences scale: an examination across sexual assault status. Journal of Psychopathology and Behavioural Assessment, 35(3), 394-403.

Opendak, M. & Sullivan, R. M. (2016). Unique neurobiology during the sensitive period for attachment produces distinctive infant trauma processing. European Journal of Psychotraumatology, 7, 31276.

Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. Journal of Educational Statistics, 8, 157–159.

Osborne, J. W. (2013). Is data cleaning and the testing of assumptions relevant in the 21st century? Frontiers in Psychology. Retrieved from: https://doi.org/10.3389/fpsyg.2013.00370

Pai, A., Suris, A.M., & North, C. S. (2017). Posttraumatic stress disorder in the DSM-5: controversy, change, and conceptual considerations. Behavioural. Sciences. 7, 7.

Parnell, L. (1999). EMDR in the Treatment of Adults Abused as Children. London, W.W. Norton & Company.

Patsopoulos, N. A. (2011). A pragmatic view on pragmatic trials. Dialogues in Clinical Neuroscience. 13(2), 217-24.

Peleikis, D. E., & Dahl, A. A. (2005). A systematic review of empirical studies of psychotherapy with women who were sexually abused as children. Psychotherapy Research, 15, 304−315.

Pelcovitz, D., van der Kolk, B.A., Roth, S., Mandel, F., Kaplan, S., & Resick, P., (1997). Development of a criteria set and a structured interview for disorders of extreme stress (SIDES). Journal of Traumatic Stress, 10, 3–16.

Palic, S., Zerach, G., Shevlin, M., Zeligman, Z., Elklit, A., & Solomon, Z. (2016). Evidence of complex posttraumatic stress disorder (CPTSD) across populations with prolonged trauma of varying interpersonal intensity and ages of exposure. Psychiatry Research, 246, 692–699.

Pietrzak, R. H., Goldstein, R. B., Southwick, S. M., & Grant, B. F. (2012). Prevalence and axis I comorbidity of full and partial posttraumatic stress disorder in the United States: Results from Wave 2 of the National Epidemiologic Survey on alcohol and related conditions. Journal of Anxiety Disorders, 25, 456–465.

Polaschek, D.L. (2011). Many sizes fit all: A preliminary framework for conceptualising the development and provision of cognitive-behavioural rehabilitation programmes for offenders. Aggression and Violent Behaviour, 16, 20–35.

Puetz, V. B., Zweerings, J., Dahmen, B., Ruf, C., Scharke, W., Herpertz-Dahlmann, B., & Konrad, K. (2016). Multidimensional assessment of neuroendocrine and psychopathological profiles in maltreated youth. Journal of Neural Transmission, 123, 1095–1106.

Resick, P. A., O'Brien Uhlmansiek, M., Clum, G. A., Galovski, T. E., Scher, C. D., & Young-Xu, Y. (2008). A Randomized Clinical Trial to Dismantle Components of Cognitive Processing Therapy for Posttraumatic Stress Disorder in Female Victims of Interpersonal Violence. Journal of Consulting and Clinical Psychology, 76(2), 243–258.

Ratcliffe, M. (2018). The interpersonal structure of depression. Psychoanalytic Psychotherapy, 32(2), 122-139.

Rees, R. J., Norsworthy, L. A., & Rowlands, S.R. (2009). Does a continuous feedback system improve psychotherapy outcome? Psychotherapy Theory, Research, Practice, Training, 46(4), 418–431.

Resick, P. A., Nishith, P., & Griffin, M. G. (2003). How Well Does Cognitive-Behavioral Therapy Treat Symptoms of Complex PTSD? An Examination of Child Sexual Abuse Survivors Within A Clinical Trial. CNS Spectrums, 8(5), 340–355.

Rieckert, J., & Möller, A. T. (2000). Rational-emotive behavior therapy in the treatment of adult victims of childhood sexual abuse. Journal of Rational-Emotive and Cognitive-Behavior Therapy, 18, 87−101.

Roberts, N. P., Roberts, P. A., Jones, N., & Bisson, J. I. (2015). Psychological interventions for post-traumatic stress disorder and comorbid substance use disorder: a systematic review and meta-analysis. Clinical Psychology Review, 38, 25-38.

Rogers, R. (2018). An introduction to response styles. In R. Rogers & S. D. Bender (Eds.)Clinical Assessment of Malingering and Deception, Fourth Edition. NY, Guilford Press.

Roth, S., Newman, E., Pelcovitz, D., Van der Kolk, B., & Mandel, F. S. (1997). Complex PTSD in victims exposed to sexual and physical abuse: results from the DSM-IV Field Trial for Posttraumatic Stress Disorder. Journal of Traumatic Stress, 10(4), 539-55.

Rose, S., Brewin, C.R., Andrews, B., Kirk, M. (1999). A randomized controlled trial of individual psychological debriefing for victims of violent crime. Psychological Medicine, 29, 793-799.

Rutterford, C., Taljaard, M. Dixon, S., Copas, A., & Eldridge, S. (2015). Reporting and methodological quality of sample size calculations. Journal of Clinical Epidemiology, 68, 716e723.

in cluster randomized trials could be improved: a review

Ruiz, M. A., Poythress, N. G., Lilienfeld, S.O., & Douglas, K. S. (2008). F actor structure and correlates of the dissociative experiences scale in a large offender sample. Assessment, 15(4) 511-521.

Sadeh, N. & McNiel, D. E. (2015). Posttraumatic Stress Disorder Increases Risk of Criminal Recidivism Among Justice-Involved Persons With Mental Disorders. Criminal Justice and Behavior, 42, 6, 573-586.

Sandhu, J. (2017). Treatment within a prison-based Therapeutic Community (TC): Experiences of extended assessment services, the impact of TC treatment, and interventions at HMP Grendon. Forensic Update, 125, 57-59.

Santa Ana, E. J., Martino, S., Ball, S., A., Nich, C., Frankforter, T. L., & Carroll, K. M. (2008). What is usual about “treatment-as-usual”? Data from two multisite effectiveness trials. Journal of Substance Abuse Treatment 35(4) 369–379.

Sarchiapone, M., Carli, V., Cuomo, C., Marchetti, M. & Roy, A. (2009). Association between childhood trauma and aggression in male prisoners. Psychiatry Research, 165, 187–192.

Sayer, N. A., Friedemann-Sanchez, G., Spoont, M., Murdoch, M., Parker, L. E., Chiros, C., & Rosenheck, R. (2009). A qualitative study of determinants of PTSD treatment initiation in veterans. Psychiatry 72(3), 238-255.

Sayın, A., Candansayar, S., & Welkin, L. (2013). Group psychotherapy in women with a history of sexual abuse: what did they find helpful? Journal of Clinical Nursing, 22, 3249–3258.

Saxena, P., Grella, C., & Messina, N, P. (2015). Continuing Care and Trauma in Women Offenders' Substance Use, Psychiatric Status, and Self-Efficacy Outcomes. Women & Criminal Justice, 0, 1–23.

Schottenbauer, M.A., Glass, C.R., Arnkoff, D.B., Tendick, V., & Gray, S.H. (2008). Nonresponse and dropout rates in outcome studies on PTSD: Review and methodological considerations. Psychiatry, 71, 134–168.

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ, 340:c332. Retrieved from: https://www.bmj.com/content/340/bmj.c332.

Scottish Government (2017). A Nation with Ambition: The Government's Programme for Scotland 2017-18. Edinburgh, Scottish Government. Retrieved from: https://beta.gov.scot/publications/nation-ambition-governments-programme-scotland-2017-18/

Seltman, H. J. (2018). Experimental Design and Analysis. Retrieved from: www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

Shea, M. T., McDevitt-Murphy, M., Ready, D. J., & Schnurr, P. P. (2009). Group therapy. In E. B. Foa, T. M. Keane, M. J. Friedman, & J. A. Cohen (Eds.), Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies (2nd ed.). New York, NY: Guilford Press.

Shevlin, M., Murphy, M., & Read, J. (2015). Testing complex hypotheses using secondary data analysis: Is the association between sexual abuse and psychosis is moderated by gender in a large prison sample? Journal of Criminal Psychology, 5(2), 92-92.

Shine, J. (2003) General Principles for Video Monitoring. Guidance notes for Prison and Probation Service Treatment Managers (Draft IV). London, HMPS & NPD.

Sloan, D.M., Feinstein, B.A., Gallagher, M.W., Beck, J.G., & Keane, T.M. (2013). Efficacy of group treatment for posttraumatic stress disorder symptoms: A meta-analysis. Psychological Trauma: Theory, Research, Practice, and Policy, 5(2), 176–183.

Sikkema, K.J., Hansen, N.B., Kochman, A., Tarakeshwar, N., Neufeld, S., Meade, C.S., et al. (2007). Outcomes froma group intervention for coping with HIV/AIDS and childhood sexual abuse: Reductions in traumatic stress. AIDS and Behavior, 11, 49–60.

Sikkema, K. J., Ranby, K. W., Meade, C., S., Hansen, N. B., Wilson, P. A., & Kochman, A. (2013). Reductions in Traumatic Stress following a Coping Intervention were Mediated by Decreases in Avoidant Coping for People Living with HIV/AIDS and Childhood Sexual Abuse. Journal of Consulting Clinical Psychology, 81(2), 274–283.

SPS, 2014. Report of the Scottish Prison Service. Organisational Review – Unlocking Potential, Transforming Lives. Edinburgh, SPS. Retrieved from: http://www.sps.gov.uk/Corporate/Publications/Corporate9.aspx.

Snaith, R. P. (2003). The Hospital Anxiety and Depression Scale. Health and Quality of Life Outcomes. 1,1:29.

Stalker, C. A., & Fry, R. (1999). A comparison of short-term group and individual therapy for sexually abused women. Canadian Journal of Psychiatry, 44, 168−174.

Steins Bisschop,C.N., Courneya, K. S., Velthuis, M. J., Monninkhof, E. M., Jones, L. W., Friedenreich, C. van der Wall, E., Peeters, P. H., & May, A. M. (2015). Control group design, contamination and drop-out in exercise oncology trials: a systematic review. PLoS One, 27;10(3), e0120996.

Stern, A. F., (2014). Questionnaire review: the Hospital Anxiety and Depression Scale Occupational Medicine, 64, 393–394.

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., Carpenter, J.R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ, 29, 338:b2393.

Stewart, L. (2015). Effective interventions for women offenders: a rapid evidence assessment. National Offender Management Service. Retrieved from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/448859/effective-interventions-for-women-offenders.pdf

Stewart, L.A., Wilton, G., & Sapers, J. (2016). Offenders with cognitive deficits in a canadian prison population: prevalence, profile, and outcomes. International Journal of Law and Psychiatry, 44, 7–14.

Stice, E., Burton, E., Bearman, S. K., & Rohde, P. (2007). Randomized trial of a brief depression prevention program: an elusive search for a psychosocial placebo control condition. Behavioural Research and Therapy, 45(5), 863-76.

Szafranski, D. D., Smith, B. N, Gros, D. F., Resick, P. A. (2017). High rates of PTSD treatment dropout: A possible red herring?Journal of Anxiety Disorders, 47, 91–98.

Tangney, J.P., Meyer, P., Furukawa, E., Cosby, B. (2002). The Criminogenic Cognitions Scale. George Mason University; Fairfax, VA.

Tangney, J. P., Stuewig, J., Furukawa, E., Kopelovich, S., Meyer, P., & Cosby, B. (2012). Reliability, Validity, and Predictive Utility of the 25-Item Criminogenic Cognitions Scale (CCS). Criminal Justice and Behaviour, 39(10), 1340–1360.

Taylor, J. E., & Harvey, S. T. (2009). Effects of psychotherapy with people who have been sexually assaulted: A meta-analysis. Aggression and Violent Behaviour, 14, 273−285.

Taylor, J. E. & Harvey, S. T. (2010). A meta-analysis of the effects of psychotherapy with adults sexually abused in childhood. Clinical Psychology Review, 30, 749–767.

Thornton, J., Alderson, P., Tan, T., Turner, C., Latchem, S., Shaw, E., et al (2013). Introducing GRADE across the NICE clinical guideline program. Journal of Clinical Epidemiology, 66, 124-131.

Tucker, M. & Oei, T. P. S. (2007). Is group more cost effective than individual cognitive behaviour therapy? the evidence is not solid yet. Behavioural and Cognitive Psychotherapy, 35, 77–91.

Tully, R. J., & Bailey, T. (2017). Validation of the Paulhus Deception Scales (PDS) in the UK and examination of the links between PDS and personality. Journal of Criminological Research, Policy and Practice, 3(1), 38-50.

Tully, A., Forrester, A. & Exworth, T. (2014). The development of equivalence as a mechanism to improve prison healthcare. Journal of the Royal Society of Medicine, 107(5), 179–182.

Tirado-Munoz, J., Gilchrist, G., Lligona, E., Gilbert, L., & Torrens, M. (2015). A group intervention to reduce intimate partner violence among female drug users. Results from a randomized controlled pilot trial in a community substance abuse centre. Adicciones, 27(3), 168-178.

Torchalla, I., Nosen, L., Rostam, H., & Allen, P. (2012). Integrated treatment programs for individuals with concurrent substance use disorders and trauma experiences: A systematic review and meta-analysis. Journal of substance abuse treatment, 42(1), 65–77.

UKPTS (2017). Guideline for the treatment and planning of services for complex post-traumatic stress disorder in adults. Retrieved from: http://www.ukpts.co.uk/links_6_2920929231.pdf

US Veterans Affairs (2012) Scoring the PCL-C. Retrieved from: https://sph.umd.edu/sites/default/files/files/PTSDChecklistScoring.pdf

van den Berg, D.P.G, de Bont, P.A.J.M., van der Vleugel, B.M., de Roos, C., de Jongh, A., van Minnen, A., & van der Gaag, M. (2016). Trauma-Focused Treatment in PTSD Patients With Psychosis: Symptom Exacerbation, Adverse Events, and Revictimization. Schizophrenia Bulletin, 42(3), 693–702.

van der Hart, O., Brown, P., & van der Kolk, B. A. (1989). Pierre Janet's treatment of posttraumatic stress. Journal of Traumatic Stress, 2, 379–396.

van der Kolk, B. (2005). Developmental Trauma Disorder: Towards a rational diagnosis for children with complex trauma histories. Psychiatric Annals, 401-408.

van der Kolk, B. (2010) Traumatic Antecedents Questionnaire. Trauma Assessment Packet. Brookline, MA., The Trauma Centre.

van der Kolk, B. (2014). The body keeps the score; mind brain and body in the transformation of trauma. UK: Penguin.

van der Kolk B.A., Pelcovitz, D., Roth, S.H., et al. (1996). Dissociation, somatization, and affect dysregulation: The complexity of adaptation to trauma. American Journal of Psychiatry, 153, 83–93.

van der Kolk, B. A., Stone, L., West, J., Rhodes, A., Emerson, D., Suvak, M., Spinazzola, J. (2014). Yoga as an adjunctive treatment for posttraumatic stress disorder: A randomized controlled trial. The Journal of Clinical Psychiatry, 75(6), 559-65.

van der Kolk, B. & van der Hart, O. (1989). Pierre Janet and the Breakdown of Adaptation in Psychological Trauma. The American Journal of Psychiatry, 146 (12), 1530- 1540.

van Ginkel, J. R. & Kroonenberg, P. M. (2014). Analysis of variance of multiply imputed data. Multivariate Behavioural Research, 49(1), 78–91.

van Ginneken, E., F. J. C. (2016). Making sense of imprisonment: narratives of posttraumatic growth among female prisoners. International Journal of Offender Therapy and Comparative Criminology, 60(2), 208- 227.

Viera, A. J. & Bangdiwala, S. I. (2007). Eliminating bias in randomized controlled trials: importance of allocation concealment and masking. Family Medicine, 39(2), 132-7.

Walsh, K., DiLillo, D. K., & Scalora, M. J. (2011). the cumulative impact of sexual revictimization on emotion regulation difficulties: an examination of female inmates. Violence Against Women, 17(8), 1103–1118.

Walters, G. D., & Schlauch, C. (2008). The psychological inventory of criminal thinking styles and level of service inventory-revised: Screening version as predictors of official and self-reported disciplinary infractions. Law and Human Behavior, 32(5), 454-462.

Watts, B.V., Schnurr, P.P., Mayo, L.,Young-Xu, Y., Weeks, W. B., & Friedman, M, J. (2013). Meta-Analysis of the Efficacy of Treatments for Posttraumatic Stress Disorder. Journal of Clinical Psychiatry, 74(6), 541-550.

Weisz, J. R., Weiss, B., Han, S. S., Granger, D. A., & Morton, T. (1995). Effects of psychotherapy with children and adolescents revisited: A meta-analysis of treatment outcome studies. Psychological Bulletin, 117, 450−468.

Weierstall, R., Huth, S., Knecht, J., Nandi, C., Elbert, T. (2012). Appetitive Aggression as a Resilience Factor against Trauma Disorders: Appetitive Aggression and PTSD in German World War II Veterans. PLoS ONE, 7(12), e50891.

Weisburd, D., Lum, C. M., & Yang, S. (2003) When can we conclude that treatments or programs "don't work"? The Annals of the American Academy, 587, 31-48.

Welfare, H. & Hollin, C. R. (2012). Involvement in extreme violence and violence-related trauma: A review with relevance to young people in custody. Legal and Criminological Psychology, 17(1), 89-104.

Wessely, S., Bryant, R., Greenberg, N., Earnshaw, M., Sharpley, J., & Hughes, J. (2008). Does psychoeducation help prevent post traumatic psychological distress? Psychiatry Interpersonal & Biological Processes, 71(4), 287-302.

Wolff, N., Frueh, B.C., Shi, J., Gerardi, D., Fabrikant, N. & Schumann, B. E. (2011). Trauma exposure and mental health characteristics of incarcerated females self-referred to specialty PTSD treatment. Psychiatric Services, 62(8), 954-8.

Wolff, N., Huening, J., Shi, J., Frueh, B. C., Hoover, D. R., & McHugo, G. (2015). Implementation and effectiveness of integrated trauma and addiction treatment for incarcerated men. Journal of Anxiety Disorders, 30, 66–80.

WHO (2002). World report on violence and health. Retried from https://apps.who.int/iris/bitstream/handle/10665/42495/9241545615_eng.pdf;jsessionid=D1BF47D630F126A5B835947B3B1DD659?sequence=1

WHO (2018). International Classification of Diseases-11. World Health Organisation. Retrieved from: https://icd.who.int8

Yeomans, P.D., Forman, E.M., Herbert, J.D., Yuen, E. (2010). A randomized trial of a reconciliation workshop with and without PTSD psychoeducation in Burundian sample. Journal of Traumatic Stress, 23(3), 305–12.

Zaleski, K.L., Johnson, D. K. & Klein, J. T. (2016). Grounding Judith Herman's Trauma Theory within Interpersonal Neuroscience and Evidence-Based Practice Modalities for Trauma Treatment. Smith College Studies in Social Work, 86(4), 377-393.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavica, 67, 361-370.

Zlotnick, C., Franklin, L., & Zimmerman, M. (2002). does “subthreshold” posttraumatic stress disorder have any clinical relevance? Comprehensive Psychiatry, 43 (6), 413 – 419.

Zlotnick, C., & Pearlstein, T. (1997). Validation of the structured interview for disorders of extreme stress. Comprehensive Psychiatry, 38(4), 243-7.

Zlotnick, C., Johnson, J.E., & Najavits, L.M. (2009). Randomized controlled pilot study of cognitive–behavioral therapy in a sample of incarcerated women with substance use disorder and PTSD. Behavior Therapy, 40, 325–336.

Zlotnick, C., Najavits, L. M., Rohsenow, D. J., & Johnson, D. M. (2003). A cognitive-behavioral treatment for incarcerated women with substance abuse disorder and posttraumatic stress disorder: Findings from a pilot study. Journal of Substance Abuse Treatment, 25, 99–105.

Zlotnick, C., Shea, T.M., Rosen, K., Simpson, E., Mulrenin, K., Begin, A., et al. (1997). An affect-management group for women with posttraumatic stress disorder and histories of childhood sexual abuse. Journal of Traumatic Stress, 10, 425–436.

1. Appendix for systematic review and meta-analysis

Contents	*page*
1.1. Literature search strategy	209
1.2. Author's categorisation decisions for treatment arms	210
1.3. Funnel for group-based trauma interventions compared to usual care	226
1.4. GRADE and risk of bias outcomes	227
1.5. Summary of moderator findings across outcome domains	242

Appendix 1.1. Literature search strategy

Search strategy for Psychinfo (EBSCO)

Date undertaken: 10 December 2016

The inclusion criteria was based on the following: 1) adult participants who were at least 18+ years of age; 2) at least one treatment arm designed or implemented as a specific psychological treatment to ameliorate, stabilise or inform participants about symptoms and experiences associated with complex trauma (Courtois & Ford, 2016 pg.9); 3) the predominant (at least 90%) treatment modality was undertaken in a group setting; 4) outcome measures were assessed using a validated instrument, either self-report or administered during a clinical interview; 5) outcome measures included PTSD severity or the severity of other internal, external or global symptoms associated with complex trauma (depression, psychological distress, substance misuse and dissociation); 6) a clear randomized trial design comparing at least one active psychological treatment to at least one control condition or another active treatment condition; 7) the article is published in English; 8) published in a peer-reviewed journal.

The search used the key terms [“Posttraumatic Stress Disorder" OR "Complex PTSD" OR "DESNOS" or "developmental trauma" OR "Acute Stress Disorder" OR "Stress Reactions" OR “Post-Traumatic Stress”], as well as other terms relevant to interpersonal violence, AND terms related to specific type of psychotherapies including ["Group Psychotherapy" OR (“group*” OR “psychotherapy*” OR “group”] AND [“Psychoeducation" OR “psychoeducat*”] AND [“random*” OR “control”]

Appendix 1.2. Author's decisions distinguish treatment arms as 'Psychoeducation', 'Psychoeducation Plus' and 'Trauma Memory Processing' (TMP) within included studies.

Study *Summary of judgements for interventions as specified as treatment or control conditions.*	**Notes[1] supporting categorisation of interventions and which arms were included in the Treatment condition.** *[1] Where possible description of the interventions is taken directly from the text published by authors of the included studies*
Alexander et al (1989) **1. *TMP:*** *Designed to be a peer support group process; limited details available however discussing abuse experiences are the 'norm'.* **2. *Psychoeducation:*** *Although very interactional the IT group is considered to have a psychoeducational approach to treatment particularly as compared to PG* ***Combined = Psychoeducation Plus:*** *given that both a Psychoed*	**Intervention 1: Process Group (PG). A format which** is a more peer orientated process, less structured and form available information seems to be more orientated towards a **supportive group process**. It is noted that the facilitators took more of a minimal role. Difficult to specifically categorise this group given limited information in the available paper regarding the content of what was actually discussed. However, in Alexander, Neimeyer & Follette (1991) more details about this treatment model and it is apparent that group members are asked to describe their abuse experiences and it is established that it is the 'norm' that the abuse would be discussed although the extent that this was focused on varied amongst groups. Whilst there was some uncertainty as to whether to categories this treatment as a TMP group it was decided that this was appropriate given that it was disclosure based and described by the authors as something that might be advantageous, for example, after an IT approach. **Intervention 2: Interpersonal Transaction group format (IT).** There are two therapeutic group formats are utilised within this study and which are described briefly. It is our opinion from the available description that the **IT** has greater psychoeducational content and structure as the group is introduced to various topics by the therapists each week. These included relevant feelings and cognitions: such as negative perception of the self, helplessness and trust. It is noted that during the course of the therapy greater disclosure (presumed to be about coping and general life experiences (**and not specifically linked to TMP**) and levels of intimacy were encouraged.

and TMP are combined together when compared against a WL	
Bass et al (2013) **1. *Psychoeducation Plus:*** *As this is the cognitive only version of CPT this is therefore not categorised as a TMP intervention but as it processes meaning is more than just psychoeducation*	**Intervention 1: Cognitive Processing Therapy – Cognitive only.** There was a degree of uncertainty as to how much there was a trauma memory processing component to this work. However, the authors very clearly highlighted that the 2 sessions usually within CPT based on the written account was removed. As such it was considered that there was a considerable treatment component to this work but that as the specific memory had not been addressed that it would not classified as a specific TMP treatment. However, given that the remaining components are an essential part of CPT and the extensive work undertaken in adjusting it to the needs of the specific African population that it would be categorised as Psychoed Plus. CPT-C vs Support group. **WL/TAU: Individual Support.** When women were informed of their eligibility, psychosocial assistants invited them to receive individual support services as desired, including psychosocial support and economic, medical, and legal referrals. Psychosocial assistants were available throughout the treatment period for women who sought their services. Workers who were delivering the experimental condition had previously been trained only in case management and individual supportive counselling. Not regarded as a sophisticated trauma therapy and in the west might be part of the therapeutic milieu that participants receive as standard. Therefore assigned to the WL/TAU group rather than an Active treatment.
Bohus et al (2013) **1. *TMP:*** *this is a comprehensive intervention with exposure based material.*	**Intervention 1: PTSD –Dialectical Behavioural Therapy (DBT).** This is a multi-component modular treatment programme. It is conducted as a 12-week residential programme. Its goals are to help patients achieve the following: (1) reduce their fear of trauma-associated primary emotions such as fear, disgust and powerlessness, (2) question non-justified secondary emotions such as guilt, shame and self-contempt and (3) radically accept trauma-related biographic facts. Exposure-based techniques are applied to reduce fear of trauma-associated emotions.
Bradley et al (2003) **1. *TMP****: Given the structured writing and***	**Intervention 1: Narrative based trauma focused treatment.** Group sessions were 2.5 hr. Nine treatment sessions focused on education about interpersonal victimization and affect regulation (e.g., identifying and naming emotions and precipitating factors; using breathing exercises to decrease distress). The skills were based on Linehan's DBT model (Linehan, 1993). Nine sessions focused on structured writing assignments: 'We encouraged women to create

disclosure of abuse experiences	meaningful *narratives of their life experiences, including interpersonal victimization.'*
Chard et al (2005) **1. *TMP**: designed to process traumatic events*	**Intervention 1: Cognitive Processing Therapy for Sexual Abuse (CPT-SA).** The treatment consisted of 17 weeks of manual-based group and individual therapy, with participants attending a 90-min group each week and a 60 min individual therapy session for the first 9 weeks and the 17th week. The authors claim that this format has a couple of advantages including processing their traumatic events with the sole attention of the individual therapist. Therefore, processing has planned within the overall within the overall therapeutic approach.
Classen et al (2001) **1 and 2. *Combined [by the authors] = Psychoeducation Plus:** A difficult decision based on the published combined results of a **Psychoeducation and TMP intervention.** This categorisation seeks to recognise a potential overall increase therapeutic intensity within the combined results.*	**Intervention 1: Present Focused Group Therapy (PFGT)**. Is described as helping participants to identify and modify the maladaptive patterns of behaviour that had arisen as a result of their traumatic past. In the present-focused treatment the assumption is that by focusing on the here-and-now survivors can alter their current functioning and thereby address the impact of their abuse history. The aim was to help members become aware of their own internal affective and cognitive states. To recognize the triggers for their trauma symptoms, recognize how they were affected by others, and to learn ways of managing their trauma symptoms and expressing their needs, concerns, or fears as soon as they arose. **Intervention 2: Trauma Focused Group Therapy (TFGT) [TMP].** Described as having an increased level of trauma memory processing not otherwise expected in a phase 1 intervention. *Claassen et al (2001) somewhat problematically presents the results from both of these treatment approaches together and not separately. Therefore, the TFGT has not been included within this analysis as a TMP treatment or PFTG as a psychoeducational intervention.*
Classen et al (2011) **1. *TMP:** as described by the protocol.*	**Intervention 1: Trauma Focused Group Therapy (TFGT)**. Considered a TMP intervention; as described above. **Intervention 2: Present Focused Group Therapy (PFGT)**. As described above.

2. *Psychoeducation:* *Given the stabilisation and phase 1 orientation of the PFGT. This categorisation is used for the Active comparison when PFGT is compared to TFGT.* ***Combined = Psychoeducation Plus:*** *This categorisation is used when the results of the PFGT & TFGT are combined and compared against the WL control group.*	
Cole et al (2007) **1.*TMP****: A comprehensive intervention that involves psychoeducation but also processing of abuse histories***	**Intervention 1:** A four phase treatment approach which included: (1) self-soothing and safety, (2) psychoeducation, (3) processing and (4) termination. The first seven sessions focused on boundary setting, self-esteem and identity, and relaxation. Sessions eight through to eleven focused on understanding trauma and sexual abuse, trauma and addiction, identifying interpersonal patterns of abuse and assertiveness training. Session twelve focused on the processing and writing of personal stories of trauma. Participants engaged in a full hour of writing their personal story of childhood sexual abuse. Immediately following the writing exercise, participants were given the opportunity to openly discuss the writing experience.
Constantino et al (2005) **1. *Psychoeducation:*** *some concerns about the*	**Intervention 1: Social Support and Stabilisation**. A brief intervention for women within a domestic violence shelter. The course being over 8 weeks, once a week, and focusing primarily on stabilization issues and resourcing including coping with stress. The course also provides peer contact. It should also be noted that intervention is described as providing resources to the participants as well as time to access resources when available, and an 'environment to

amount of trauma informed information as opposed to general resourcing.	chat with a counsellor and friends'.
Crespo & Arinero (2010) **1. *TMP*:** *exposure based protocol* **2. *Psychoeducation:*** *general focus on trauma and stabilisation of affect and communication skills.*	**Intervention 1: Exposure**. Both arms essentially use the same group based format however this arm had exposure content (and therefore categorised as being TMP intervention). **Intervention 2: Communication Skills Group**. In considering the description of this intervention it was apparent that the group included trauma specific content on cognitions and affect management with an additional focus on feelings of anger and the ability to express and communicate emotions. Same format as the control group however instead of having TMP sessions these were substituted with anger management and communication and expression of emotions psychoeducation and skills material. Thus, Sessions 1-5 and 8 (this last focused on relapse prevention) were equivalent in both groups, since Sessions 6 and 7 were focused on the specific technique for each group (that is, exposure technique or communications skills training).
Dorrepaal et al (2012) **1. *Psychoeducation****:*** *fits criteria as specified in this review.*	**Intervention 1: 'Stabilising' Group Treatment.** The aim of this intervention is to decrease core symptoms of complex PTSD. A psychoeducation model is describes as aimed at attaining cognitive behavioural skills focused on identifying and modifying dysfunctional behaviour, thoughts and beliefs about the trauma, in particular the trauma's meaning for one's self. The focus of the treatment was towards the here-and-now, on positive reinforcement and empowerment.
Falsetti et al (2008) **1. *TMP:*** *this intervention contains some very distinct elements of processing and writing about relevant trauma.*	**Intervention 1: Multiple Channel Exposure Therapy (M-CET).** Developed to treat posttraumatic stress disorder (PTSD) with comorbid panic attacks. Contains psychoeducation and skills but also TMP material. Contents includes psychoeducation about PTSD and panic breathing and breathing retraining, CBT principles including catastrophic thinking and other cognitive distortions, an introduction to interoceptive exposure to panic related physical sensations, introduced cognitive exposure by writing about the traumatic events, identification of components of the traumatic event that were most painful to recall, and any cognitive distortions that may interfere with processing the memory accurately, discussion of the writing assignment, further identification of distortions and difficulties about

	the event and education about how trauma can affect one's sense of safety. Material also included in vivo exposure, and application of cognitive skills to decrease cognitive distortions regarding trust, the traumatic event, and its consequences as well as a review and ongoing plans for treatment.
Ford et al (2013) **Frisman et al (2008)** **1. *Psychoeducation****:*** *fits criteria as specified in this review.* **2. *Non-Trauma-Focused active control:*** *no trauma specific content*	**Intervention 1: Trauma Affect Regulation: Guide for Education and Therapy (TARGET**; Ford & Russo, 2006). Specific psychoeducation material linking PTSD symptoms to affect dysregulation, explaining that both are the result of biological adaptations to survival threats sequential skill-set for affect regulation including modelling and coaching by the therapist, and in vivo in individualized homework assignments. **Intervention 2: Supportive Group Therapy (SGT).** This is considered as an active treatment as it has a protocol designed to engage women in identifying current stressors and coping behaviours that work for them or others. SGT includes experiential self-expression activities and nondirective assistance in identifying stresses and effective coping strategies. The first five sessions provided education on group rules and setting personal goals, recognizing symptoms of traumatic stress, personal boundaries, and styles of attachment. The remaining seven sessions are devoted to an open-ended "discussion and to work through some of the issues, needs and concerns raised as participants began to make the connections between past trauma and present problems and function", However, SGT did not include the therapeutic mechanisms (i.e., detailed education about traumatic stress and the brain, and emotion/self-regulation skills training) hypothesized to be crucial in TARGET.
Garland et al (2016) **1. *Psychoeducation:*** A condensed version of Seeking Safety this was combined with MORE where relevant as both were active treatments with similar intentions. *However, MORE was contrasted with Trauma*	**Intervention 1: CBT trauma informed intervention**. Described prior to Corrigendum as a condensed version of Seeking Safety. This intervention provided training in cognitive, behavioural, and interpersonal coping skills and delivered content from the Seeking Safety treatment manual on the following session topics PTSD: taking back your power; detaching from emotional pain; when substances control you; compassion; honesty; recovery thinking; setting boundaries in relationships; healing from anger; termination. **Intervention 2: Mindfulness-Oriented Recovery Enhancement (MORE).** Described as uniting 'complementary aspects of mindfulness training, third-wave cognitive-behavioural therapy, and principles from positive psychology into an integrative intervention strategy'. MORE is therefore a mindfulness addiction and psychological distress (generally rather than C-PTSD specifically). The MORE treatment manual has been used in across a number of different addiction related settings (Garland, 2013, 2014), *sessions offered instruction in applying mindfulness and related skills to the following topics: awareness of automaticity in addiction; disrupting the link between negative*

Informed –MBSR (Kelly & Garland, 2016) which was concluded to be more of an example of a TFG intervention. **2. *Non-Trauma-Focused active control***. *MOREs was described as an addiction intervention and not as a specific trauma based intervention. However, its distress tolerance has made it one of the more sophisticated NTFG interventions that are included.*	*emotions and addictive behaviour through reappraisal; refocusing attention from stress and craving to savour pleasant experiences; regulating craving through mindful attention and awareness; decreasing craving through mindful stress reduction; promoting acceptance instead of suppression of experience; awareness of the impermanence of the body; mindful relationships; interdependence and meaning in life; and developing a mindful recovery plan.* There was some debate over whether MORE was a specific intervention designed to ameliorate trauma. The authors described it as a treatment designed to help 'substance dependent individuals with trauma histories'. However nowhere in the protocol is trauma specifically mentioned as outlined in the paper although there were a lot of very relevant affect and recovery management skills. The therapeutic mechanisms, were hypothesized by the authors to produce greater improvements in primary outcomes including craving, post-traumatic stress symptoms, and psychiatric distress than CBT or TAU. In this sense, MORE uses mindfulness training to facilitate cognitive reappraisal; this potentially synergistic "mindful reappraisal" approach may have led to the observed reduction in post-traumatic stress. In that regard, more frequent use of cognitive reappraisal predicts attenuated PTSD symptoms. Although this is not a TMP intervention it could nevertheless be considered, because of its addiction component, a *Psychoeducation Plus* intervention. It was therefore combined with the CBT intervention as unlike any of the other NTFG treatments as it was recognised to have many particularly trauma relevant processes.
Ghee et al (2009) **1. *Psychoeducation****: fits criteria as specified in our review. Also this is only 6 sessions of SS and therefore a much briefer version.*	**Intervention 1: Seeking Safety.** A condensed 6 session version. The content is described as including (a) Introduction to Safety, (b) PTSD: Taking Back Your Power, (c) Detaching from Emotional Pain (Grounding), (d) Setting Boundaries in Relationships, (e) Asking for Help, and (f) Commitment. [Seeking Safety utilizes coping skills to integrate trauma-specific treatment with substance abuse treatment]. Study design compared against standard comprehensive addictions treatment.
Graham-Berman (2013)	**Intervention 1: Mom's Empowerment Program.** As referenced from the paper the intervention is described as being:

1. *TMP*: *fits with processing and disclosure aspect expected of such an intervention.*	'based in part on interpersonal theory, the MEP emphasizes the whole person and explores strengths and abilities that can be used to compensate for biopsychosocial dysfunction...the group was designed to provide a venue for exploring relationship issues, including parent– child relationships, expectations derived from their family of origin, and social support. By telling their IPV story, connecting events to emotional reactions, identifying their fears and worries, and enhancing their self-esteem, the women may reduce their level of traumatic stress and recover from PTSD. A healing feature offered by therapy for traumatic exposure is the normative and empathic response of group therapists and group members providing the opportunity for the woman to make sense of (give meaning to) otherwise overwhelming and senseless event(s) in a safe environment.'
Hein et al (2009) ***1. Psychoeducation****: fits criteria as specified in our review. Also this is only 6 sessions of SS and therefore a briefer version.*	**Intervention 1: Seeking Safety.** A condensed 12 session version (mean attended 6.2 sessions). The sessions covered topics such as: Safety, PTSD: Taking Back your Power; When Substances Control You; Honesty; Setting Boundaries in Relationships; Compassion; Healing from Anger; Creating Meaning; Integrating the Split Self; Taking Good Care of Yourself; Red and Green Flags, and Detaching from Emotional Pain (Grounding). Control: '…The active comparison group, Women's Health Education (WHE), was intended to control for therapeutic time and attention, but may have also included other active therapeutic elements…. It is a psychoeducational, manualized health curriculum focused on topics such as understanding the female body, human sexual behavior, pregnancy and childbirth, sexually transmitted diseases, HIV, and AIDS. WHE was designed to provide equivalent therapeutic attention, expectancy of benefit, and an issue-oriented focus, but without theory-driven techniques (i.e., of SS, nor any explicit focus or psychoeducation specific to substance abuse or trauma). All WHE sessions followed a common format: 1) introduction of topic; 2) review of group rules and between session assignment; 3) topic presentation, 4) video, storytelling and/or text readings; and 5) topic exercises in a variety of formats to facilitate group discussion and application of session materials; and 6) setting between-session goals.'
Hinton et al (2011) **1. *TMP:*** *Clear inclusion of an exposure based protocol which has been culturally adapted. Little*	**Intervention 1: Culturally Adapted CBT (CA-CBT) for PTSD.** Described as including loving-kindness meditation; modifying catastrophic cognitions about PTSD and anxiety symptoms, including those related to cultural syndromes; educating about PTSD and trauma-cued recall; positive reframing of trauma cues; teaching emotion regulation techniques; presenting key lessons with culturally appropriate; analogies ("cultural bridging"); interoceptive exposure to anxiety-type sensations (through head rotation and hyperventilation), in conjunction with re-association to positive images, to treat fear of somatic and mental symptoms; applied relaxation techniques, including both applied

mention of the trauma based memory processing specifically however concluded relevant to TMP category. **2. *Psychoeducation:*** *Whilst the AMR was designed as a control it is clearly more psychoed that the label/name given to this treatment might suggest and belies its description of it being a trauma based psychoed affect management intervention.*	muscle relaxation and applied stretching, with positive self-statements that pair bodily flexibility to emotional flexibility. However, importantly also includes exposure and modification of fear networks. Authors explain that 'exposure is followed by practicing of a trauma-processing protocol; a set of emotion regulation techniques, which includes visualizations that are culturally adapted for the group in question (e.g., applied stretching of tense muscles followed by head rotation and a palm-tree visualization in the case of Latino patients)'. **Intervention 2: Applied Muscle Relaxation (AMR).** Consisted of instructions on applied muscle relaxation using a manual (Hinton & Safren, 2009) and was also provided by the first author and patients' therapist. The AMR treatment condition consisted of the same number of sessions as the control and each lesson was likewise 1 hour. The AMR treatment was manualized and importantly it is described as containing psychoeducation about PTSD symptoms, anxiety, and panic, and about how those disorders produce somatic symptoms; it also provides education about how AMR can reduce anxiety and hence improve those disorders.
Hollifield et al (2007) **1. *TMP:*** *includes exposure techniques and makes reference to trauma [memory] content.*	**Intervention 1: CBT.** Sessions 1 through 3 uses psychoeducation, behavioural activation, and activity planning, collectively termed "Trajectory and Resource Loss Stabilization." Participants identify valued resources that have either been lost or are at-risk, and then make plans to engage in activities that will help establish a resource gain cycle. In sessions 4 to 10, participants are taught classic cognitive restructuring and imagery rehearsal, utilizing material from daily life experiences, much of which is either symbolic or direct trauma content. They are encouraged to also continue behavioural activation and activity planning to regain lost resources. In sessions 10 to 12, participants use classic exposure and desensitization techniques while being encouraged to practice earlier session skills. Participants are asked to identify remaining feared or avoided situations, make plans to engage in those situations in the following 6 weeks, and use cognitive restructuring and/or imagery rehearsal as desensitizing reciprocal inhibitors.
Kaslow et al (2010)	**Intervention 1: Nia**. The authors describe Nia as 'a culturally informed, empowerment focused psychoeducational

1. *Psychoeducation: fits criteria as specified in this review.*	group intervention'. The focus is on addressing psychological symptomatology (suicidal ideation, depressive symptoms, posttraumatic stress symptoms, general psychological distress). Culturally informed, empowerment-focused psychoeducational group intervention. Nia consisted of 10 manualized, 90-min group meetings. Meetings included three to five women. Nia expands upon existing evidence-based psychosocial treatments for suicidal behaviours and IPV by being designed as a culturally informed group intervention that is specific to African American women with a history of both IPV and suicide attempts.
Kelly et al (2016) **1. *Psychoeducation:** fits criteria as specified in this review.*	**Intervention 1: Trauma-Informed Mindfulness-Based Stress Reduction (TI-MBSR).** Described by the authors of this intervention as a 'phase I trauma intervention for female survivors of interpersonal violence (IPV)'. *The skills taught in the program focus are designed to cultivate a sustained, focused awareness of mental states, physical sensations, perceptions, thoughts, and imagery. The TI-MBSR treatment manual detailed modifications of the original MBSR protocol designed to target clinically salient issues for female survivors of IPV. No original content from the original MBSR protocol was cut in the modification; instead, additional psychoeducational content was embedded within each session.*
Krakow et al (2001) **1. *Psychoeducation Plus**: fits criteria as specified in our review in that it has a specific focus on a PTSD symptom.*	**Intervention 1: Imagery Rehearsal Therapy.** A specific brief for chronic nightmares associated with PTSD (sexual assault). Treatment protocols are manualised and focus on nightmares within a framework of imagery and cognitive restructuring. The authors are careful to specifically mention: 'descriptions of traumatic experiences and traumatic content of nightmares are discouraged throughout the program in a carefully designed attempt to minimize direct exposure. To facilitate this approach, participants are instructed to work first with a nightmare of lesser intensity and, if possible, one that does not seem like a "replay" or a "re-enactment" of a trauma.' Therefore, it was concluded that this was not a TMP intervention.
Krupnick et al (2008) **1. *Psychoeducation Plus**. The specific focus of this intervention is participant's interpersonal relationships.*	**Intervention 1: Interpersonal Psychotherapy (IPT).** Treatment consisted of 16 two-hour sessions. The authors describe the treatment as follows: initial phase (Sessions 1-4) aimed at establishing a safe atmosphere and developing group cohesion. Therapists educated participants about PTSD and how PTSD increased interpersonal difficulties. The intermediate phase (Sessions 5- 12) addressed relationship disputes, social deficits, role transitions, and relationship losses. Members identified relationship behaviours that decreased social support or led to further exploitation or abuse and explored how these patterns led to PTSD symptoms. In the termination phase (Sessions 13-

This therapeutic focus being considered somewhat greater that just psychoeducation.	16), participants then focus on mourning the loss of the group as well as prior relationship losses and disappointments, anticipated future problems, and identified relationship triggers that could reactivate PTSD symptoms, exposing them to further interpersonal trauma or withdrawal from others.
Lau & Kristensen (2007) **1. *TMP:*** Limited information trauma memory processing assumed from available description. **2. *Psychoeducation Plus****: Whilst psychoeducation plays a significant role in the systematic group given length of treatment it is assumed that additional material and possibly TMP might have taken place from the reframing.*	**Intervention 1: Analytic group psychotherapy.** This treatment arm was over a more prolonged time and consisted of one session each week for 12 months. The intervention focused on intra-psychic and interpersonal dynamics and difficulties both in past and present relationships and within the group. The theoretical position being that interpersonal psychoanalysis is particularly suited to facilitate change in women who were sexually abused as children because both the abuse and the therapy are fundamentally relational. There is a limited amount of information about the contents (and group discussions) and level of trauma memory processing has been assumed from further information in the follow up study by Elkjaer, Kristensen, Mortensen, Poulsen & Lau (2013). **Intervention 2: Systemic group psychotherapy.** This included psychoeducation every fourth session, and the therapeutic focus was flashbacks, guilt, and validation of perceptions and feelings. The authors describe systemic group psychotherapy as 'a short-term and focused treatment: i.e. of two sessions each week for 5 months. This group was highly structured framework with initial goal setting and rounds during sessions. At each session, it was decided who had speaking time, when and for how long, and participants were supported by an active therapist. The main task was to reframe the patients' life histories and make them construct more rewarding perceptions of themselves and their situations, thus creating new behavioural possibilities. Every second week there was 1 hr of psycho-education with the group choosing the topics'.
McWhirter (2010) **1. *Psychoeducation****: of the 2 treatments the emotion focused therapy best fits criteria as specified in our review in that it has a specific therapeutic focus.*	**Intervention 1: Emotion Focused Therapy.** Of the 2 interventions in this study, delivered to women who have experienced IPV, the emotional focused therapy is characterised by psychoeducational trauma and emotions management content. The therapeutic orientation is described as 'behavioural and gestalt'. The structure of sessions specifically includes cognitive behavioural psychoeducational and is "here and now" focused. Positive group relationships as well as healthy and non-healthy relationships are considered. The five-session curriculum focused on the following: (a) exploring personal belief systems, especially concerning difficult experiences; (b) understanding the various forms of abuse; (c) understanding and expressing feelings; (d) recognizing healthy relationships; (e) and finding healthy ways to cope with stress. In this respect whilst the alternative goal orientated group may educated

2. *Non-Trauma-Focused active control.* *General coping strategies group.*	participants to distinguish between adaptive and non-adaptive coping strategies this was not trauma informed. **Intervention 2: Goal Orientated Group.** The authors describe this intervention as one that that seeks to help participants to distinguish between adaptive and non-adaptive coping strategies. Not a specifically trauma informed stabilisation intervention however participants are encouraged to identify a non-adaptive coping strategies which will have relevance to participants with complex trauma histories.
Messina (2010) and (2012) **1. *Psychoeducation Plus:*** *Included both trauma content and also other gender and addictions specific therapeutic content.*	**Intervention 1: Gender Responsive Treatment Model.** This model encompasses two manualized curricula, *Helping Women Recover* and *Beyond Trauma* and is designed to be relevant to the needs of drug-dependent women under criminal justice supervision. The specific trauma relevant intervention of interest is *Beyond Trauma* (Covington, 2003) which consists of 11 sessions focused on three areas: teaching women what trauma and abuse are, helping them to understand typical reactions to trauma and abuse, and developing coping skills. *Helping Women Recover* (Covington, 2008) is a 17-session program organized into four Modules: self-module, relationship module, sexuality module and spirituality module.
Rieckert Moller 2000 **1. *TMP:*** *There was obviously a large amount of psychoeducation material in this study's treatment protocol. The decision to categorise as TMP intervention is based on limited information however disclosure has been assumed as has processing when endeavouring to help*	**Intervention 1: Rational-Emotive Behavioural Therapy for CSA.** As taken from the main paper: 'initially, during the first session, participants discussed their experiences of sexual abuse and its negative emotional and behavioural effects. The basic principles of the A-B-C model of emotions were then presented. Participants were trained, through examples from their own experiences, to understand the association between activating events (A), perceptions and evaluations of those events (B), and the resultant emotional-behavioural consequences (C). The authors also describe sessions as being structured to ensure optimal participation by group members

participants work on specific individual core beliefs.	
Sikkema et al (2007, 2013) **1. *Psychoeducation Plus**:** *fits criteria as specified in our review in that it has a specific therapeutic focus for an HIV population.* **2. *Non-Trauma-Focused active control.*** *General peer support group.*	**Intervention 1: Living in the Face of Trauma (LIFT)**. A group intervention to address coping with HIV and childhood sexual abuse (CSA). Based on a coping skills framework this is a 15-session HIV and trauma coping group intervention. The cognitive theory of stress and coping and effective cognitive-behavioural treatment strategies for sexual trauma being the main focus. **Intervention 2: Support Group.** The control group is described as a 'support group' and emphasized principles of group treatment. The support groups were characterised by a focus on the present including the establishment of mutual support by group members, and the interactive processing of interpersonal issues. Described as resembling interpersonal process group (Sikkema, 2008).
Triado-Munzo et al (2015) **1. *Psychoeducation**:** *fits criteria as specified in our review.*	**Intervention 1: Intimate Partner Violence Therapy-Cognitive Behavioural Therapy (IPaViT-CBT)** Described by the authors as a '10 session intervention…manualized small-group, cognitive behavioural intervention, designed to reduce IPV and improve depressive symptoms in female drug users.' No trauma processing was apparent in the description of the sessions. Whilst targeted at a specific population there was some consideration as to whether this intervention should be categorised as a *Psychoed Plus* intervention, however, it was concluded from the available information that it was more general in focus rather that symptom specific.
van der Kolk et al (2014) **1. *Psychoeducation:*** *fits criteria as specified in our review: protocoled trauma informed intervention*	**Intervention 1: Trauma- Informed Yoga.** A protocoled programme over 10 weeks of an hour-long trauma-informed yoga class, incorporating the central elements of hatha yoga: breathing, postures, and meditation. The program promotes curiosity about bodily sensations, in which self-inquiry is prominent. The authors argue that this seeks to encourage body awareness as a necessary aspect of effective emotion regulation. For example learning to notice, tolerate, and manage somatic experience may substantially promote emotion regulation. **Intervention 2: Women's Health Education.** A matched 10 week hour-long intervention focusing on women's health

alternative to mainstream CBT based approaches. **2. *Non-Trauma-Focused active control.*** *General health intervention.*	education. This intervention did not discuss issues related to personal trauma or disclosure of abuse or trauma.
Yeomans et al (2010) **1. *Psychoeducation****:*** *fits criteria as specified in our review.* **2. *TMP:*** whether designed or not to be TMP or an active treatment condition it seems (from descriptions available) to have become this. ***Combined = Psychoeducation Plus:*** *When the results of both arms are combined this was considered to be an appropriate classification. This is based on the control arm inadvertently having assuming a **TMP** process.*	**Intervention 1: Workshop with psychoeducation**. The 3-day workshop used discussions, experiential exercises aimed at fostering interpersonal exchange, and games to explore themes of trauma, loss, anger, trust, and the roots of violence. The authors describe the psychoeducational content on the first day of the workshop as including a 90-minute presentation and discussion of the 17 specific symptoms of PTSD. **Intervention 2: Workshop 'without psychoeducation'.** The authors described this arm as encouraging increased interpersonal dialogue and as of equal length with the treatment arm. However, the participants in the control condition were encouraged to recall and discuss how they have been affected by events/experiences. Although facilitators did not augment these discussions with any PTSD psychoeducational content participants discussed in pairs thoughts on a series of topics related to issues of trust, security, and interethnic relations in the community. The additional time devoted to this content in this arm was not apparent in the experimental (psychoeducational) arm. The assigned topics facilitated communication around perspectives on trust, safety, sense of security, and interethnic relations in the community (e.g. "someone I trust and why," "a time I overcame fear"). Importantly, participants were encouraged to discuss how they have been affected by events. In this respect parallels can made to the trauma processing content of interventions such as Cognitive Processing Therapy (CPT).
Zlotnick et al (1997)	**Intervention 1: Affect-Management (AM) group.** Described as 'adjunct' to individual therapy and pharmacological

1. *Psychoeducation: fits criteria as specified in our review.*	treatment. The authors describe treatment as comprising of a 'review of the previous week's homework, a psychoeducational presentation, followed by skill building, application of skills, and assignment of homework. Therapists used a standardized manual for the affect-management group that outlines, in detail, each session (i.e., the objectives, the "mini-lecture" on the selected topic, how to conduct the practice of new skills in session, and hand-outs of homework assignments and of relevant issues associated with the material presented in the session). Material covered in sessions included education regarding PTSD, dissociation, flashbacks, "safe" sleep, identification of emotions, crisis planning, anger management, and techniques for distraction, self-soothing, distress tolerance, and relaxation'.
Zlotnick et al (2009) **1. *Psychoeducation Plus**: This is not a condensed version of SS and also had booster sessions. In this respect and the overall specific responsivity to substance misuse it was decided that this should be recognised as more than only psychoeducation.*	**Intervention 2: Seeking Safety**. This intervention consisted of 25 group sessions (mean 15.6 sessions completed). In addition individual 'booster' treatment sessions were made available after prison release. The authors noted that the 'primary goals of SS are psychoeducation and the development of coping skills to help clients attain safety from both PTSD and SUD; it is present-focused, abstinence-oriented, and emphasizes an empowering, compassionate approach'.

Appendix 1.3. Funnel plots of standard error by Hedges's g for group-based trauma interventions compared to usual care

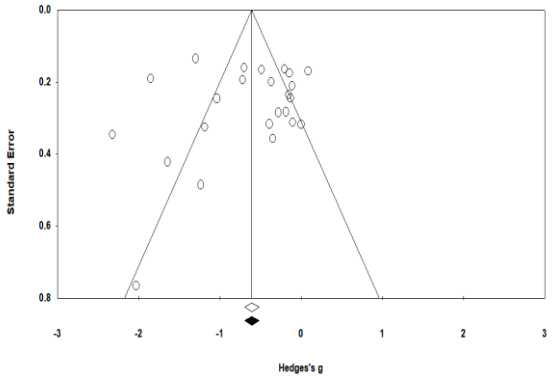


Figure 1: Funnel plot for PTSD Domain

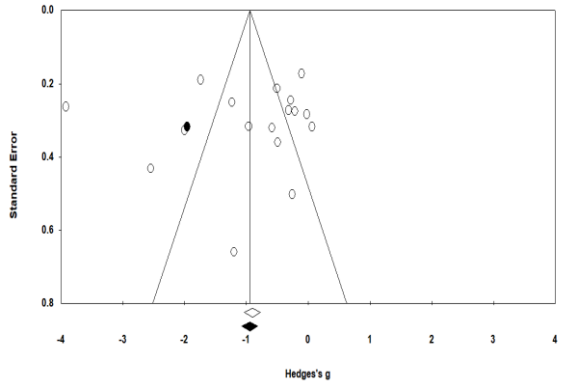


Figure 2: Funnel Plot for Depression Domain

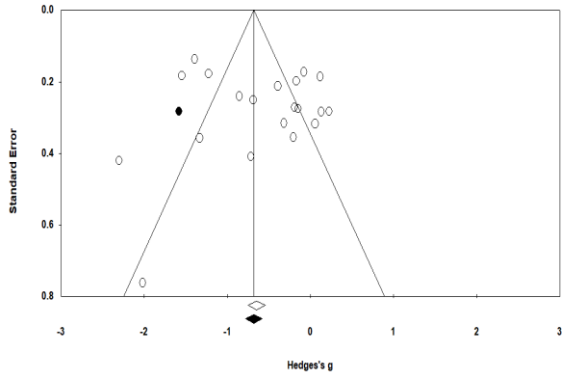


Figure 3: Funnel plot for Psychological Distress Domain

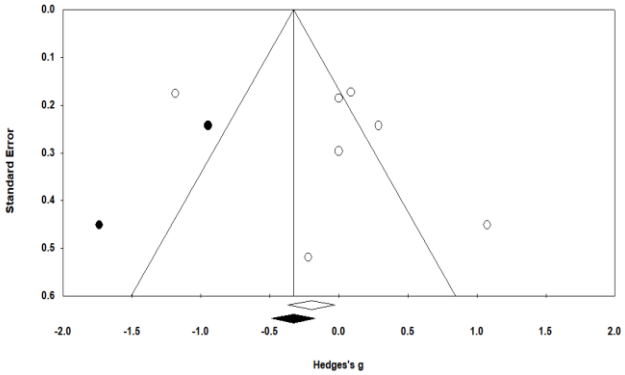


Figure 4: Funnel plot for Substance Misuse Domain

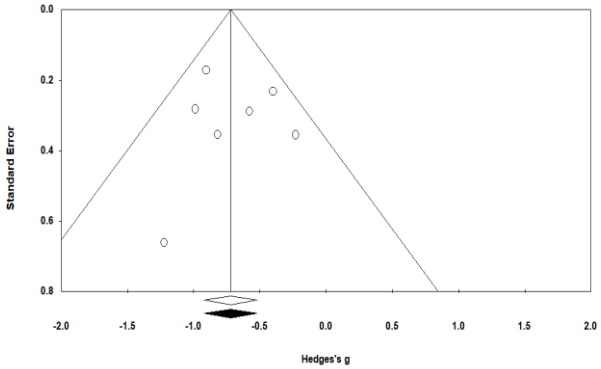


Figure 5: Funnel plot for Dissociation Domain

Note: All plots presented with observed and imputed studies

Appendix 1.4. GRADE and risk of bias assessment of outcomes

Method

All assessments were discussed within supervision and any disagreements were resolved through discussion, and an overall rating decided on.

Quality: For assessment of outcome **quality**, we downgraded by 1 point if ≥50% studies contributing to an outcome had at least one 'high risk' rating according to the Cochrane Risk of Bias assessment we conducted (excluding the ratings of 'Other Bias'), and 2 points if ≥50% relevant studies had at least two ratings of 'high'. However, we did not downgrade if the risk of bias did not affect that particular outcome. For example, if a study had significant missing data, or was at high risk of selective reporting bias, we only downgraded if the missing data and selective reporting directly affected the outcome in question.

Indirectness: ratings were informed by considerations of study population, treatment duration, and nature of control condition.

Inconsistency: we downgraded by 1 point for if the I^2 statistic was ≥40% in the context of an unclear direction of effect or ≥75% in the context of a clear direction of effect. We downgraded by 2 points if the I^2 statistic was ≥75% in the context of an unclear direction of effect.

Imprecision: we downgraded an outcome if "a recommendation or clinical course of action would differ if the upper versus the lower boundary of the CI represented the truth" (Guyatt, Oxman, Kunz, Brozek, Alonso-Coello & Rind, et al. 2011) and / or the number of events and sample size meant the optimal information size was not reached. For binary outcomes we based our judgements on absolute rather than relative estimates of effect. For the primary outcomes we considered statistical and clinical significance separately.

Publication bias:[1] we downgraded when, for outcomes with more than 10 studies (Ioannidis & Trikalinos, 2007), funnel-plots suggested asymmetry and this was not better explained by selective reporting bias or some other factor. See also Cochrane Handbook examples of hypothetical funnel plots (Figure 10.4.a).

Other considerations that might undermine the accurate measurement of the outcomes relevant to a particular domain were also recorded. Where ≥50% of relevant studies used combined data within a domain or multiple different measures were used there is the increased possibility that different constructs are being measured.

NB. 'Post' results based on variable reporting time periods based on first reporting of results. Mean calculation in weeks provided for each analysis.

- Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, et al. (2011). GRADE guidelines 6. Rating the quality of evidence--imprecision. J Clin Epidemiol. 64(12):1283-93.
- Ioannidis JP, Trikalinos TA. (2007) The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey. Canadian Medical Association Journal, 10;176(8):1091-6.

[1]Only assessed if *k* studies ≥10;

PTSD / Trauma Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations	Trauma group	Control	Hedges's g	(95% CI)		
Question and outcomes: Do group based trauma interventions overall have a significant effect on PTSD / trauma outcomes at post treatment when compared to TAU/WL comparators? [Mean 7.5 weeks; assessed with: various PTSD / trauma symptom tools]. Included studies: Bass et al (2013); Bohus et al (2013); Bradley et al (2003); Chard et al (2005); Classen et al (2001); Classen et al (2011); Cole et al (2007); Dorrepaal et al (2012); Falsetti et al (2008); Frisman et al (2008); Garland et al (2016); Ghee et al (2009); Graham-Bermann et al (2013); Hollifield et al (2007); Kaslow et al (2010); Kelly et al (2010); Krakow et al (2001); Krupnick et al (2008); Messina et al (2012); Mitchell et al (2014); Sikkema et al (2007/2013); Yeomans et al (2010); Zlotnick et al (1997); Zlotnick et al (2009).												
24	0 Under 50% of overall studies had high risk RoB (many 'unsure')	-1 Lack of symmetry	-1	0 Combination of different group treatment protocols however generalised research question	0	0	1253	976	-0.66***	-0.94; -0.37	⊕⊕◯◯ LOW	CRITICAL
Question and outcomes: Do group overall based trauma interventions have significant effect on PTSD / trauma outcomes at post treatment compared to other Non-Trauma informed active comparators? [Mean 0.8 weeks; assessed with: various PTSD / trauma symptom tools]. Included studies: Ford et al (2013); Garland et al (2016); Hien et al (2009); Sikkema et al (2013); van der Kolk et al (2013); Yeomans et al (2010a).												
5	0 High RoB in no studies	na	-2	0 All Psychoed treatment arms	-2 Wide CI	0	433	431	0.36	-1.06; 1.16	⊕◯◯◯ VERY LOW	CRITICAL

PTSD / Trauma Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations	Trauma group	Control	Hedges's g	(95% CI)		
Question and outcomes: Do phase 2 TMP group based trauma interventions have a significant effect on PTSD / trauma outcomes at post treatment compared to phase 1 Psychoeducation comparators? [Mean 7 weeks; assessed with: various PTSD / trauma symptom tools]. Included studies: Classen et al (2001); Crespo & Arinero (2010); Hinton et al (2011); Yeomans et al (2010).												
4	0 High RoB in no studies	na	-2	0 All TMP treatment arms	-2 Wide CI	0	132	131	-0.36	-1.01; 0.33	⊕◯◯◯ VERY LOW	CRITICAL
Question and outcomes: Do Psychoeducational group based trauma interventions have a significant effect on PTSD / trauma outcomes at post treatment compared to WL/TAU comparators? [Mean 3.5 weeks; assessed with: various PTSD / trauma symptom tools]. Included studies: Dorrepaal et al (2012); Frisman (2008); Garland et al (2016); Ghee et al (2009); Kaslow et al (2010); Kelly et al (2016); Mitchell et al (2014); Zlotnick et al (1997).												
8	0	na	-1	0	0	0	379	252	-0.49*	-1.00; -0.20	⊕⊕⊕◯ MODERATE	CRITICAL
Question and outcomes: Do Psychoeducation Plus group based trauma interventions have a significant effect on PTSD / trauma outcomes at post treatment compared to other WL/TAU comparators? [Mean 14.1 weeks; assessed with: various PTSD / trauma symptom tools]. Included studies: Bass et al (2013); Classen et al (2001); Classen et al (2011); Falsetti et al (2008); Krakow et al (2001); Krupnick et al (2008); Messina et al (2012); Sikkema et al (2013); Yeomans et al (2010); Zlotnick et al (2009).												
10	0 High RoB in 40% of studies.	na	-1	0	0	0	707	472	-0.60**	-1.00; -0.20	⊕⊕⊕◯ MODERATE	CRITICAL
Question and outcomes: Do TMP group based interventions have a significant effect on PTSD / trauma outcomes at post treatment compared to other WL/TAU comparators? [Mean 2 weeks; assessed with: various PTSD / trauma symptom tools]. Included studies: Bohus et al (2013); Bradley et al (2003); Chard et al (2005); Cole et al (2007); Graham-Bermann et al (2013); Hollifield et al (2007).												

PTSD / Trauma Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations	Trauma group	Control	Hedges's g	(95% CI)		
6	0 High RoB in 42.8% of studies.	na	-1	0	0	0	167	256	-0.98***	-1.53; -0.43	⊕⊕⊕◯ MODERATE	CRITICAL
Depression Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations	Trauma group	Control	Hedges's g	(95% CI)		
Question and outcomes: Do group based trauma interventions overall have a significant effect on Depression outcomes at post treatment when compared to TAU/WL comparators? [Mean 4.9 weeks; assessed with: various Depression measures]. Included studies: Alexander et al (1989); Bohus et al (2013); Bradley et al (2003); Chard et al (2005); Classen et al (2001); Classen et al (2011); Cole et al (2007) Falsetti et al (2008); Frisman et al (2008); Garland et al (2016); Hollifield et al (2007); Kaslow et al (2010); Kelly et al (2010); Krupnick et al (2008); Mitchell et al (2014); Rieckert & Moller (2000); Triado-Munzo et al (2015).												
17	0 Less than 50% high RoB rating	-1 Lack of symmetry	-1	0 Combination of different group treatment protocols	0	0	667	498	-0.60***	-0.89; -0.32	⊕⊕◯◯ LOW	IMPORTANT
Question and outcomes: Do group based trauma interventions overall, have significant effect on Depression Symptoms at post treatment when compared to other various Non-Trauma informed active comparators? [Mean 0 weeks; assessed with: various Depression measures]. Included studies: Garland et al (2016); McWhrter et al (2016); van der Kolk et al (2013)												

PTSD / Trauma Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations	Trauma group	Control	Hedges's g	(95% CI)		
3	0 Less than 50% high RoB rating	na	-2	0	-2 Wide CI	0	118	120	0.05	-1.06; 1.16	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do phase 2 TMP group based trauma interventions have a significant effect on Depression outcomes at post treatment compared to phase 1 Psychoeducation comparators? [Mean 8 weeks; assessed with: various Depression measures]. Included studies: Alexander et al (1989); Classen et al (2011); Crespo & Arinero (2010).												
3	0 Less than 50% high RoB rating	na	-2	0	0	0	103	97	0.29	-0.83; 1.4	⊕⊕◯◯ LOW	CRITICAL
Question and outcomes: Do Psychoeducation group based interventions have a significant effect on Depression outcomes at post treatment when compared to TAU/WL comparators? [Mean 4 weeks; assessed with: various Depression measures]. Included studies: Frisman et al (2008); Garland et al (2016); Kaslow et al (2010); Kelly et al (2010); Mitchell et al (2014); Triado-Munzo et al (2015).												
6	0 Less than 50% high RoB rating	na	-1	0	-1 Wide CI	0	315	192	-0.90	-1.85; 0.05	⊕⊕◯◯ LOW	IMPORTANT
Question and outcomes: Do Psychoeducation Plus group based trauma interventions have a significant effect on Depression symptoms at post treatment when compared to TAU/WL control comparators? [Mean 6 weeks; assessed with: various Depression measures] Included studies: Alexander et al (1989); Classen et al (2001); Classen et al (2011); Krupnick et al (2008).												

PTSD / Trauma Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations	Trauma group	Control	Hedges's g	(95% CI)		
4	0 Less than 50% high RoB rating	na	-2	0	-2 Wide CI	0	198	125	-0.77	-1.92; 0.39	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do TMP group based trauma interventions have a significant effect on Depression symptoms at post treatment when compared to TAU/WL control comparators? [Mean 1.7 weeks; assessed with: various Depression measures]. Included studies: Bohus et al (2013); Bradley et al (2003); Chard et al (2005); Cole et al (2007); Falsetti et al (2008); Hollifield et al (2007); Rieckert & Moller (2000).												
7	0 Less than 50% high RoB rating	na	-1	0	-1 Wide CI	0	154	181	-1.12**	-2.01; -0.23	⊕⊕◯◯ LOW	IMPORTANT

Psychological Distress Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations[2]	Trauma group	Control	Hedges's g	(95% CI)		
Question and outcomes: Do group based trauma interventions overall have a significant effect on Psychological Distress symptoms at post treatment when compared to TAU/WL comparators? [Mean 5.7 weeks; assessed with: various measures]. Included studies: Alexander et al (1989); Bass et al (2013); Bohus et al (2013); Bradley et al (2003); Chard et al (2005); Classen et al (2001); Classen et al (2011); Cole et al (2007); Constantino et al (2005); Falsetti et al (2008); Frisman et al (2008); Garland et al (2016); Hollifield et al (2007); Kaslow et al (2010); Kelly et al (2010);; Messina et al (2010); Mitchell et al (2014); Rieckert & Moller (2000); Yeomans et al (2010); Zlotnick et al (2009).												
20	0 Less than 50% high RoB rating	-1 Lack of symmetry	-1	-1 different group treatment protocol	0	-1 60% of studies outcome data used from combining different measures	959	715	-0.60***	-0.89; -0.32	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do group based trauma interventions overall have a significant effect on Psychological Distress symptoms at post treatment compared to various Non-Trauma informed active comparators? [Mean 0 weeks; assessed with: various measures]. Included studies: Ford et al (2013); Garland et al (2016); van der Kolk et al (2014).												
3	0 Less than 50% high RoB rating	na	0	0	-2 Wide CI	-1 66.7% of studies outcome data used from combining different measures	126	127	0.06	-0.66; 0.78	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do phase 2 TMP group based interventions have a significant effect on Psychological Distress symptoms at post treatment compared to phase 1 Psychoeducational comparators? [Mean 4.3 weeks; assessed with: various measures]. Included studies: Alexander et al (1989); Classen et al (2011); Crespo & Arinero (2010 Hinton et al (2011); Lau (2007); Yeomans et al (2010).												

Psychological Distress Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations[2]	Trauma group	Control	Hedges's g	(95% CI)		
6	0 Less than 50% high RoB rating	na	-1	0	-2 Wide CI	-1 50% of studies outcome data used from combining different measures	204	201	0.19	-0.34; 0.71	⊕◯◯◯ VERY LOW	MPORTANT
Question and outcomes: Do Psychoeducation group based trauma interventions have a significant effect on Psychological Distress outcomes at post treatment when compared to TAU/WL comparators? [Mean 4 weeks; assessed with: various measures]. Included studies: Constantino et al (2005); Frisman et al (2008); Garland et al (2016); Kaslow et al (2010); Kelly et al (2010); Mitchell et al (2014).												
6	0 Less than 50% high RoB rating	na	-1	-1 Combination of the WL/TAU and Active controls as well as different group treatment protocol	0	-1 66.7% of studies outcome data used from combining different measures	321	196	-0.51	-1.09; 0.08	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do Psychoeducation Plus group based trauma interventions have a significant effect on Psychological Distress outcomes at post treatment when compared to TAU/WL comparators? [Mean 11.3 weeks; assessed with: various measures]. Included studies: Alexander et al (1989); Bass et al (2013); Classen et al (2001); Classen et al (2011); Messina et al (2010); Yeomans et al (2010); Zlotnick et al (2009).												

Psychological Distress Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations[2]	Trauma group	Control	Hedges's g	(95% CI)		
7	0 Less than 50% high RoB rating	na	-2	-1 Combination of the WL/TAU and Active controls as well as different group treatment protocol	-2 Wide CI	0 42% outcome data used from combining different measures	484	338	-0.38	-0.91; 0.15	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do TMP group based interventions have a significant effect on Psychological Distress outcomes at post treatment when compared to TAU/WL comparators? [Mean 1.7 weeks; assessed with: various measures]. Included studies: Bohus et al (2013); Bradley et al (2003); Chard et al (2005); Cole et al (2007); Falsetti et al (2008); Hollifield et al (2007); Rieckert & Moller (2000).												
7	-1 50% high RoB rating	na	0	0 WL/TAU control or group vs individual	-2 Wide CI	-1 71.4% of studies outcome data used from combining different measures	154	181	-0.98***	-1.66; -0.40	⊕◯◯◯ VERY LOW	IMPORTANT

Substance Misuse Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations[2]	Trauma group	Control	Hedges's g	(95% CI)		
Question and outcomes: Do group based trauma interventions overall have a significant effect on Substance Misuse outcomes at post treatment compared to TAU/WL comparators? [Mean 12.1 weeks; assessed with: various measures]. Included studies: Classen et al (2011); Frisman et al (2008); Garland et al (2016); Ghee et al (2009); Messina et al (2010); Triado-Munzo et al (2015), Zlotnick et al (2009).												
7	0 Less than 50% high RoB rating	na	-1	0	-2 Wide CI	-1 Substantive variety of different measures	413	260	-0.03	-0.56; 0.50	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do group based trauma interventions overall, are prioritised as the treatment condition, have significant effect on Substance Misuse outcomes at post treatment compared to other Non-Trauma informed active comparators? [Mean 13.3 weeks; assessed with: various measures]. Included studies: Garland et al (2016); Hien et al (2009); McWhiter et al (2011); Meade et al (2010).												
4	0 Less than 50% high RoB rating	na	-2	0	-2 Wide CI	-1 Substantive variety of different measures	386	388	0.45	-0.21; 1.12	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do phase 2 TMP group based trauma interventions have a significant effect on Substance Misuse outcomes at post treatment compared to phase 1 Psychoeducational trauma group comparators? [Mean 24 weeks; assessed with: various measures]. Included studies: Classen et al (2011).												
1	0 Less than 50% high RoB rating	na	-1 No clear direction of effect	0	-2 Wide CI	-1 Only one study	30	33	1.10	-2.28; 2.48	⊕◯◯◯ VERY LOW	IMPORTANT

Substance Misuse Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations[2]	Trauma group	Control	Hedges's g	(95% CI)		
Question and outcomes: Do group based trauma Psychoeducation interventions have a significant effect on Substance Misuse outcomes at post treatment compared to TAU/WL control comparators? [Mean 8 weeks; assessed with: various measures]. Included studies: Frisman et al (2008); Garland et al (2016); Ghee et al (2009); Triado-Munzo et al (2015).												
4	-2 50% multiple high RoB rating	Na	-2 No clear direction of effect	0	-2 Wide CI	-1 Substantive variety of different measures	267	161	0.37	-0.86; 0.59	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do group based trauma Psychoeducation Plus interventions have a significant effect on Substance Misuse outcomes at post treatment compared to TAU/WL control comparators? [Mean 17.7 weeks; assessed with: various measures]. Included studies: Classen et al (2011); Messina et al (2010); Zlotnick et al (2009).												
3	0 Less than 50% high RoB rating	na	-1 No clear direction of effect	0	-2 Wide CI	-1 Substantive variety of different measures	146	99	0.41	-0.70; 0.89	⊕◯◯◯ VERY LOW	IMPORTANT

Dissociation Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations	Trauma group	Control	Hedges's g	(95% CI)		
Question and outcomes: Do group based trauma interventions overall have a significant effect on Dissociation outcomes at post treatment when compared to TAU/WL comparators? [Mean 8.6 weeks; assessed with: various symptom tools]. Included studies: Bradley et al (2003); Bohus et al (2013); Chard et al (2005); Classen et al (2001); Classen et al (2011); Cole et al (2007); Zlotnick et al (1997).												
7	-1 50% had at least one high RoB rating	na	0	0	0	0	227	193	-0.70***	-1.05; -0.35	⊕⊕⊕◯ MODERATE	IMPORTANT
Question and outcomes: Do group based trauma interventions overall have a significant effect on Dissociation outcomes at post treatment compared to Non Trauma informed active control comparators? [Mean 0 weeks; assessed with: various symptom tools]. Included studies; Ford et al (2013); van der Kolk et al (2013).												
2	0 No high RoB rating	na	-2	0	-2 Wide CI	0	61	62	0.18	-0.43; 0.80	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do phase 2 TMP group based trauma interventions for a significant effect on Dissociation outcomes at post treatment compared to phase 1 Psychoeducation group comparators? [Mean 24weeks; assessed with: various symptom tools]. Included studies: Classen et al (2011).												
1	0 No high RoB rating	na	-1	0	-2 Wide CI	-2 Only one study	55	56	-0.12	-0.92; 0.67	⊕◯◯◯ VERY LOW	CRITICAL
Question and outcomes: Do group based Osychoeducational trauma interventions have a significant effect on Dissociation outcomes at post treatment when compared to TAU/WL control comparators? [Mean 0 weeks; assessed with: various symptom tools]. Included studies: Zlotnick et al (1997).												

Dissociation Symptoms												
Quality assessment							№ of participants		Effect		GRADE Quality Rating	Importance
№ of studies	Quality	Publication Bias[1]	Inconsistency	Indirectness	Imprecision	Other considerations	Trauma group	Control	Hedges's g	(95% CI)		
1	-1 50% had at least one high RoB rating	na	0	0	-1 Wide CI	-2 Only one study	16	17	-0.82*	-1.60; -0.04	⊕◯◯◯ VERY LOW	IMPORTANT
Question and outcomes: Do group based Psychoeducation Plus trauma interventions have a significant effect on Dissociation outcomes at post treatment when compared to TAU/WL comparators? [Mean 24 weeks; assessed with: various symptom tools]. Included studies: Classen et al (2001); Classen et al (2011).												
2	-2 50% had at least 2 high RoB ratings	na	0	0	0	0	130	88	-0.79***	-1.19; -0.39	⊕⊕⊕◯ MODERATE	IMPORTANT
Question and outcomes: Do group based TMP trauma interventions have a significant effect on Dissociation outcomes at post treatment when compared to TAU/WL comparators? [Mean 3weeks; assessed with: various symptom tools]. Included studies: Bradley et al (2003); Bohus et al (2013); Chard et al (2005); Cole et al (2007).												
4	-1 60% had at least one high RoB rating	na	0	0	0	0	81	88	-0.61***	--0.97; -0.24	⊕⊕⊕◯ MODERATE	IMPORTANT

Risk of bias summary

Review authors' judgements about each risk of bias item for each included study. Procedures as described in the Cochrane Handbook for Systematic Review of Interventions (Higgins & Green 2011; version 5.1.)

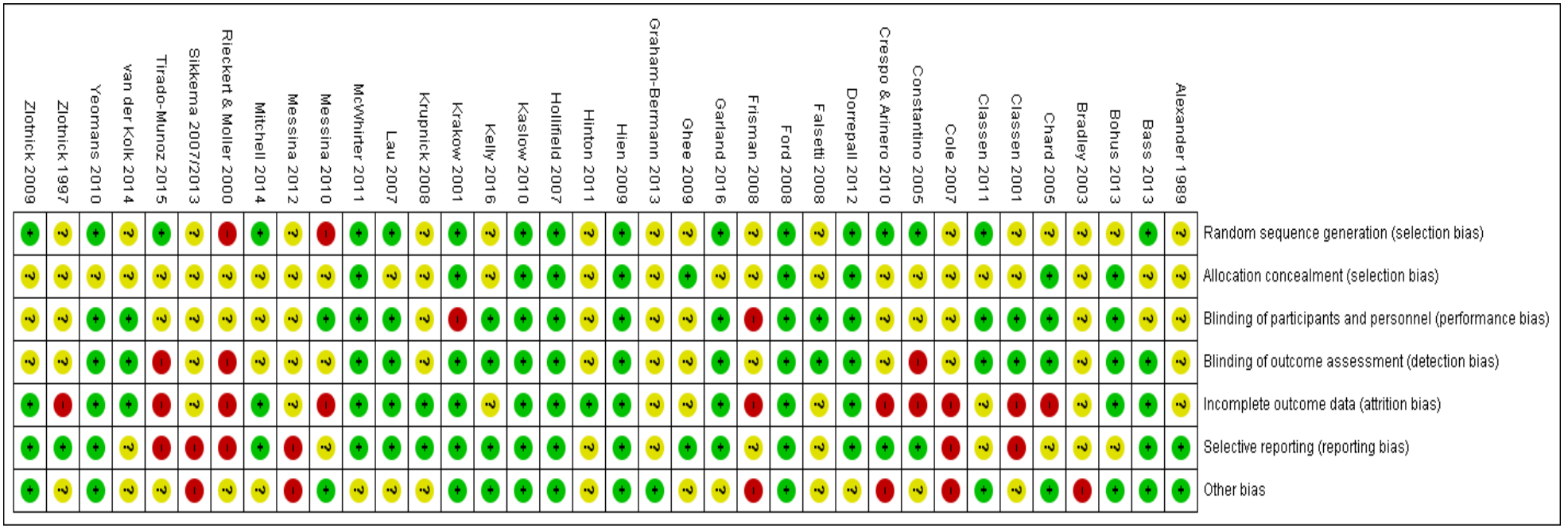


Figure 1: Risk of bias summary: review authors' judgements about each risk of bias item for each included study

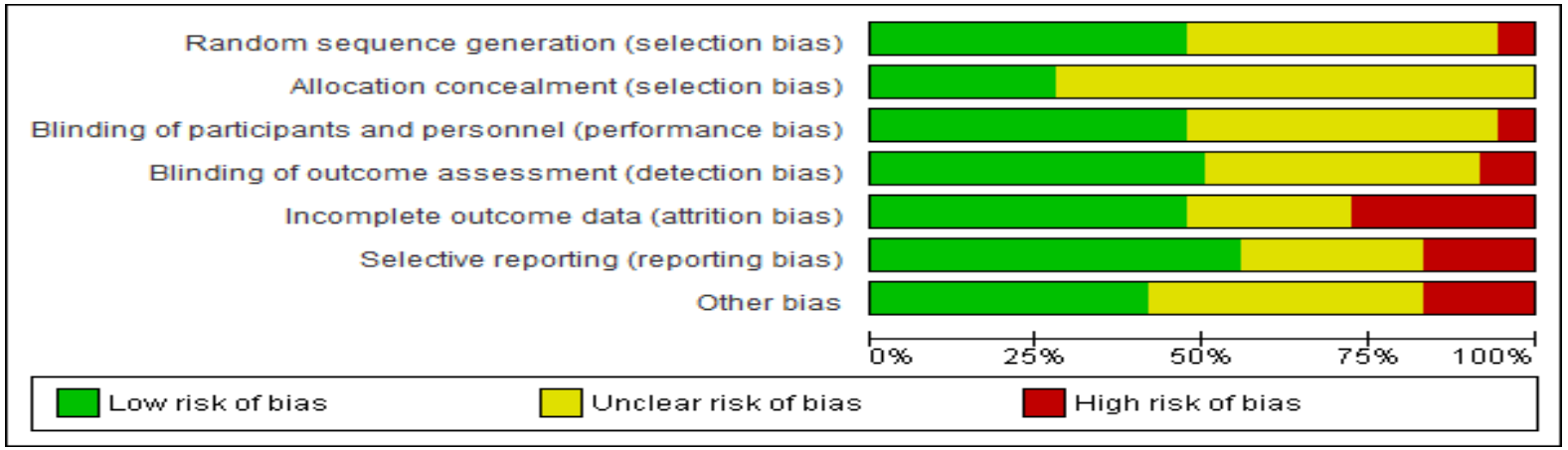


Figure 2: Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

Appendix 1.5. Summary of moderator findings across outcome domains for group treatments compared to WL/TAU

	PTSD Symptoms n= 24	Depression n=17	Psychological Distress n=20	Substance Misuse n=7	Dissociation n=7
	Q_M (df)	Q_M (df)	Q_M (df)	Q_M (df)	Q_M (df)
General Study Characteristics					
Year of publication	0.19 (1)	0.51 (1)	0.36 (1)	9.47 (1)*	0.01 (1)
Country (population/study conducted in) USA, Europe, Other (Africa)	0.25(2)	2.05 (2)	1.10(2)	0.06 (1)	2.36(1)

Participant Characteristic					
Age	6.27(1)**	2.62 (1)	2.70 (1)	0.71(1)	2.11(1)
Gender mix (% female)	0.21 (1)	5.28 (1)*	0.43 (1)	11.36 (1)**	-
SES Background Low SES; Mixed or Middle	0.35 (1)	0.19 (1)	0.01 (1)	0.24 (1)	2.11 (1)
Treatment setting Community population; specific population in institutional setting; prison / forensic population	0.38(2)	7.39(2)*	3.29(2)	0.33(2)	3.76 (2)
Therapist context					
Who facilitated? Researcher / special therapists; standard service provider	0.01(1)	0.44 (1)	0.23(1)	16.34 (1)***	0.01(1)
Level of expertise Experienced professional provider; other	2.37(1)	0.01(1)	0.64(1)	-	0.02(1)
Amount and Quality of Treatment					
Planned number of treatment sessions	0.07(1)	0.04(1)	0.01(1)	0.01(1)	2.26)
Mean number of treatment sessions attended	0.95 (1)	0.03 (1)	0.03 (1)	0.01 (1)	2.44 (1)
Number of sessions per week Once a week; twice a week; daily	0.37 (2)	0.11 (2)	0.79 (2)	1.14 (2)	2.78 (2)
Duration of treatment (in weeks)	0.24(1)	0.02 (1)	1.27(1)	0.01(1)	0.40(1)
Risk of bias? (i.e. RCT implementation problems): Yes; possible; no	1.33(2)	2.42(2)	2.74 (2)	1.37(2)*	1.03(2)
Treatment Content					
CBT treatment model	0.75(1)	0.24 (1)	0.02(1)	30.35 (1)***	-
Mindfulness / yoga model	0.09(1)	1.13 (1)	0.09(1)	2.93(1)	2.36(1)
Additional/booster or individual sessions	1.56(1)	0.82 (1)	0.61(1)	0.01(1)	0.10(1)

Emotions management component	0.06(1)	0.03(1)	0.08(1)	0.01(1)	0.96(1)
Behavioural social skills component	1.85(1)	1.87(1)	0.02(1)	2.19(1)	-
Coping Problem solving component	0.01(1)	1.16 (1)	0.15(1)	1.31(1)	0.64(1)
Interpersonal component	0.03 (1)	0.60(1)	0.18 (1)	30.34(1)***	0.09 (1)
Substance misuse component	1.45 (1)	2.98(1)	0.12 (1)	-	1.79 (1)
Note. Q_M = Q test for moderation; df = degrees of freedom . *p<.05. significant coefficient [regression of Hedge's g]					

2. Appendix for methods: Randomised control trial methods

Contents *page*

2.1. Psychometric properties of outcome measures 245

2.2. Missing data 252

2.3 Statistical analysis and interpretation 255

2.3.1. Effect size 255

2.3.2. Linear mixed models 257

2.3.3. Interpretation of LMM output 265

2.1. Psychometric properties of outcome measures

The information in the tables below was used in the Reliable Change Index/Clinically Significant Change (RCI/CSC) calculations. The relevant psychometric properties are presented with respect to the summary score for each measure as well as for each subscale. The reference in respect of Jacobson's (1999) criteria (i.e. a, b or c) is also presented as are, where possible, the clinical norms (M, SD) and comparison norms (M, SD) for a non-clinical population. References are provided to indicate the source of the normative information used to determine these scores.

Behavioural Assessment Checklist- Revised			
Total		**Subscales**	
Range of permissible scores (lowest-highest)	0-204	Range of permissible scores (lowest-highest)	Belligerence 0-24, Distress 0-24, Withdrawal 0-20, Impulsivity 0-52, Egocentricity 0-32, Problem solving 0-40
Direction of clinical gain	Decreasing	Direction of clinical gain	Decreasing

Reliability of Measure *As calculated from this study – see Results chapter*	Total .915	Reliability of Measure Cronbach's alpha coefficients *As calculated from this study – see Results chapter*	Belligerence .945, Distress .894, Withdrawal .628, Impulsivity .908, Egocentricity .906, Problem solving .799
Reference data for Clinically Significant Change (CSC)			
Mean (SD) of clinical norms *As calculated from overall norms (WL/TAU and S&T) from this study – see Results chapter*	68.98 (22.08)	Mean of clinical norms *As calculated from overall norms (WL/TAU and S&T) from this study – see Results chapter*	Belligerence 7.63 (6.41), Distress 7.62 (4.55), Withdrawal 6.78 (3.02), Impulsivity 14.74 (8.17), Egocentricity 8.63 (5.40) Problem solving 23.81 (6.48)
Mean (SD) of comparison norms	NA	Mean of comparison norms	NA
CSC criteria utilised		CSC criteria utilised	A
References: • No references regarded as applicable.			

PTSD Checklist –Civilian Version			
Total		**Subscales**	
Range of permissible scores (lowest-highest)	17-85	Range of permissible scores (lowest-highest)	Re-experiencing (5-25), Avoidance (7-35) Arousal (5-25)
Direction of clinical gain	Decreasing	Direction of clinical gain	Decreasing
Reliability of Measure	0.94	Reliability of Measure	Re-experiencing .85, Avoidance .85, Hyper-arousal .87

Cronbach's alpha coefficients		Cronbach's alpha coefficients	
Reference data for Clinically Significant Change (CSC)			
Mean (SD) of clinical norms	57.4 (17.0) (Ball et al, 2013)	Mean of clinical norms	Re-experiencing 17.0 (5.6), Avoidance 23.6 (8.8), Arousal 16.8 (5.0) (Ball et al (2013)
Mean (SD) of comparison norms	29.4 (12.9) (Ruggiero, 2003)	Mean of comparison norms	Re-experiencing 9.2 (4.2), Avoidance 12.0 (5.7), Hyperarousal 8.2(4.3) (Ruggiero, 2003)
CSC criteria utilised	External = 45 (Blanchard et al 1996)	CSC criteria utilised	C
References: • Ball, S., Karatzias, T., Mahoney, A., Ferguson, S., & Pate, K. (2013). Interpersonal trauma in female offenders: a new, brief, group intervention delivered in a community based setting. Journal of Forensic Psychiatry & Psychology, 24(6), 795-802. • Blanchard, E. B., Jones-Alexander, J., Buckley, T. C. (1996). Psychometric properties of the PTSD Checklist (PCL). Behaviour Research and Therapy, 34, 669-673. • Ruggiero, K. J., Del Ben, K., Scotti, J. R. & Rabalais, A. E. (2003). Psychometric Properties of the PTSD Checklist—Civilian Version. Journal of Traumatic Stress, 16(5) 495–502.			

Difficulties in Emotional Regulation Scale			
Total		**Subscales**	
Range of permissible scores (lowest-highest)	36-180	Range of permissible scores (lowest-highest)	Non-Acceptance 6-30; Goals 5-25; Impulsivity 6-30; Awareness 6-30; Strategies 8-40. Clarity 5-25

Direction of clinical gain	Decreasing	Direction of clinical gain	Decreasing
Reliability of Measure	0.93 (Gratz & Roemer, 2004)	Reliability of Measure Cronbach's alpha coefficients	Non-Acceptance .85; Goals .89; Impulsivity .86; Awareness .80; Strategies .88. Clarity .84 (Gratz & Roemer, 2004)
Reference data for Clinically Significant Change (CSC)			
Mean (SD) of clinical norms	100.23 (SD=32.57) Howard et al (2016) (Alternative noted: 82.61 (32.21), Walsh et al (2011)	Mean (SD) of clinical norms	Non-Acceptance 14.34 (6.95); Goals 13.22 (6.95); Impulsivity 12.61 (6.22); Awareness 14.46 (6.09); Strategies 17.44 (7.64). Clarity 10.54 (5.09) Walsh et al (2011)
Mean (SD) of comparison norms	77.99 (20.72) (Gratz et al 2004)	Mean of comparison norms	Non-Acceptance 11.65 (4.72); Goals 14.41 (4.95); Impulsivity 10.82 (4.41); Awareness 14.34 (4.60); Strategies16.16 (6.19). Clarity 10.61 (3.80) (Gratz, et al 2004)
CSC criteria utilised	C No external cut off	CSC criteria utilised	C No external cut off

References:

- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. Journal of Psychopathology and Behavioral Assessment, 26 (1), 41- 54.
- Walsh, K., DiLillo, D., & Scalora, M. L. (2011). The cumulative impact on Emotion Regulation Difficulties: An examination of female inmates. Violence against Women, 17(8), 1,103-1,118.

Dissociative Experiences Scale -II			
Total		**Subscales**	
Range of permissible scores (lowest-highest)	0-100	Range of permissible scores (lowest-highest)	0-100 – mean of each subscale
Direction of clinical gain	Decrease	Direction of clinical gain	Decrease
Reliability of Measure	.93 (Ruiz, et al 2008)	Reliability of Measure	Absorption .95; Depersonalisation .89; Amnesia .90 (Ruiz, et al 2008)
Reference data for Clinically Significant Change (CSC)			
Mean (SD) of clinical norms	20.8 (14.6) Ruiz 2008 [women offenders]	Mean (SD) of clinical norms	Abortion 24.35 (37.4), Depersonalisation 17.48 (19.78) , Amnesia 11.02 (14.26) (Mazzotti, 2016: Psychiatric patients)
Mean (SD) of comparison norms	15.55 (11.78) (Stockdale et al, 2002)	Mean (SD) of comparison norms	Absorption 21.56 (14.52); Amnesia 8.0 (10.30); Depersonalisation 7.07 (11.50) (Stockdale et al, 2002)
CSC criteria utilised Independent criteria cut off	30, a suggested cut-off point for dissociative pathology (Carlson & Putnam, 1993)	CSC criteria utilised	C

- Carlson, E. B., & Putnam, F. W. (1993). An update on the Dissociative Experiences Scale. Dissociation: Progress in the Dissociative Disorders, 6(1), 16-27.
- Ruiz, M. A., Poythress, N. G., Lilienfeld, S.O., & Douglas, K. S. (2008). Factor structure and correlates of the dissociative experiences scale in a large offender sample. Assessment, 15(4) 511-521.
- Mazzotti, E., Farina, B., Imperatori, C., Mansutti, F., Prunetti, E., Speranza, A., & Barbaranelli, C. (2016). Is the Dissociative Experiences Scale able to identify detachment and compartmentalization symptoms? Factor structure of the Dissociative

Experiences Scale in a large sample of psychiatric and non-psychiatric subjects. Neuropsychiatric Disease and Treatment, 12, 1295–1302.
- Stockdale, F. D., Gridley, B. E., Balogh, D, W., Holtgraves, T. (2002). Confirmatory factor analysis of single- and multiple-factor competing models of the dissociative experiences scale in a nonclinical sample. Assessment, 9(1), 94-106.

Hospital Anxiety Depression Scale			
Total		**Subscales**	
Range of permissible scores (lowest-highest)	0-42	Range of permissible scores (lowest-highest)	Anxiety and Depression 0-21
Direction of clinical gain	Decrease	Direction of clinical gain	Decrease
Reliability of Measure	0.84 (Stern, 2014)	Reliability of Measure	Anxiety 0.83, Depression 0.82 (Stern, 2014)
Reference data for Clinically Significant Change (CSC)			
Mean (SD) of clinical norms	12 (3) (Stern, 2014)	Mean of clinical norms	-
Mean (SD) of comparison norms	8 (2.5) (Stern, 2014)	Mean of comparison norms	-
CSC criteria utilised	Criterion C	CSC criteria utilised	External = 7 (Stern, 2014)
• Stern, A. F. (2014). The Hospital Anxiety and Depression Scale: Questionnaire Review. Occupational Medicine, 64, 393–394			

Criminal Cognitions Scale			
Total		**Subscales**	
Range of permissible scores (lowest-highest)	33-132	Range of permissible scores (lowest-highest)	Short Term, Entitlement, Responsibility, Authority, Insensitivity (all scales) 5-20; Reparation 8-32

Direction of clinical gain	Decreasing	Direction of clinical gain	Decreasing
Reliability of Measure Cronbach's alpha coefficients *As calculated from this study – see Results chapter*	.740	Reliability of Measure Cronbach's alpha coefficients *As calculated from this study- see Results chapter*	Short Term .442, Entitlement .703, Responsibility .641 Authority .710, Insensitivity .670, Reparation .701
Reference data for Clinically Significant Change (CSC)			
Mean (SD) of clinical norms *As calculated from overall norms (*WL/TAU *and S&T) from this study – see Results chapter*	74.57 (8.68)	Mean of clinical norms *As calculated from overall norms (*WL/TAU *and S&T) from this study- see Results chapter*	Short Term 10.24 (2.24), Entitlement 9.34 (2.44) Responsibility 11.78 (2.71) , Authority 11.57 (2.49) Insensitivity 9.15 (6.77), Reparation 22.50 (13.96)
Mean (SD) of comparison norms	NA	Mean of comparison norms	NA
CSC criteria utilised	A	CSC criteria utilised	A
References: • No references regarded as applicable.			

2.2. Missing data

Missing data – How much missing data is too much?

Dziura, Post, Zhao, Fu & Peduzzi (2013) highlight that the proportion of missing data in itself doesn't equate to study validity but that minimal missing data is likely to ensure more valid conclusions. For example, Schulz and Grimes (2002) suggest that losses to follow- up less than 5% usually have little impact whereas losses greater than 20% raise serious concerns about study validity. However, the often quoted 5-20% convention has no statistical justification and as Dziura et al (2013) note this can oversimplify the problem as the bias resulting from missing data also depends on the missing data mechanism and the analytic method.

See also Panel on Handling Missing Data in Clinical Trials commissioned recently by the National Research Council (2010) Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK209904/

CONSORT statement includes a set of checklists on ITT and missing data

- National Research Council: The prevention and treatment of missing data in clinical trials. In Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington DC: National Academies Press. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK209904/
- Dziura, Post, Zhao, Fu & Peduzzi (2013)
- Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gøtzsche, P.C., Devereaux, P.J., Elbourne, D., Egger, M., Altman, D.G. (2010). CONSORT 2010

explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010, 340:c869

- Schulz, K.F., Grimes, D.A. (2002). Sample size slippages in randomised trials: exclusions and the lost and wayward. Lancet,359(9308):781-5.

Multiple Imputation and Mixed Models – Sensitivity Analysis

CONSORT stipulates: *Participants with missing outcomes can be included in the analysis only if their outcomes are imputed (that is, their outcomes are estimated from other information that was collected). Imputation of the missing data allows the analysis to conform to intention-to-treat analysis but requires strong assumptions, which may be hard to justify. Simple imputation methods are appealing, but their use may be inadvisable. In particular, a widely used method is "last observation carried forward" in which missing final values of the outcome variable are replaced by the last known value before the participant was lost to follow up. This is appealing through its simplicity, but the method may introduce bias, and no allowance is made for the uncertainty of imputation. Many authors have severely criticised last observation carried forward.*

(CONSORT Statement 2010: 12a Statistical methods – missing outcomes)

Both multiple imputation and model-based approaches as described in the CONSORT guidance for ITT analysis were used for repeatedly measured outcomes. This ensured that outcomes were based on all observed data and correct for potentially biased methods for MAR data.

Sterne et al (2009) suggest reporting results from **Multiple Imputation (MI)** and a **Complete Case Analysis** within the online supplement to enable data to be contrasted. These authors also provide a useful checklist for imputation modelling including which variables were used in this process. In all MI generated data variables for age, randomised arm and all subscales of a measure were used in the model. Separate MI data sets were

computed for each measure.

Where possible, provide results from analyses restricted to complete cases, for comparison with results based on multiple imputation. If there are important differences between the results, suggest explanations, bearing in mind that analyses of complete cases may suffer more chance variation, and that under the missing at random assumption multiple imputation should correct biases that may arise in complete cases analyses.

van Ginkel et al (2014) note that there are no explicit rules for pooling F-tests of (repeated-measures) ANOVA and that the **Full information maximum likelihood (FIML)** available in SPSS Mixed Models procedures incorporates all available data without excluding incomplete cases. In this respect uses FIML same amount of information for handling the missing data as Multiple imputation does. Consequently, MI has no advantages over FIML. Therefore FIML is used within the LMM option as the gold standard. van Ginkel et al (2014) recommend that auxiliary variables are included in the imputation model. In this respect auxiliary variables have been used that help make sense of missingness including age, length of sentence and experimental arm.

It is assumed that MI will produce more conservative results. As such only results that are significant with MI or FIML procedures will be considered reliable.

Little's statistical test was used to help demonstrate that missing data were not MCAR (Little, 1988). However, Dziura, Post, Zhao, Fu & Peduzzi, (2013) note reliance on one statistical technique alone is inappropriate and exploring patterns of missing data are also crucial in establishing

whether missingness is MCAR/MAR or missing not at random (MNAR).

- Dziura, J. D., Post, L. A., Zhao, Q., Fu, Z., & Peduzzi, P. (2013). Strategies for dealing with missing data in clinical trials: from design to analysis. Yale Journal of Biology and Medicine, 86, 343-358.
- Little R. J. A. (1988). A Test of Missing Completely at Random for Multivariate Data with Missing Values. Journal of American Statistical Association, 83, 1198- 202.
- Sterne, J., White, I., Carlin, J., Spratt, M., Royston, P., Kenward, M., Wood, A., Carpenter, J. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ, 338:b2393.
- van Ginkel, J. R. & Kroonenberg, P. M. (2014). Analysis of variance of multiply imputed data. Multivariate Behavioural Research, 49(1), 78–91.

2.3. Statistical analysis and interpretation

2.3.1. Effect size

Sullivan & Feinn (2012) articulate the need for studies to report effect sizes and not just statistical significance (p values). As such the effect size can be considered as a main finding where the difference between the mean, outcomes in two different experimental arms are being considered. Wolff (1986) also notes that the interpretation of an effect size can depend upon context. As such a small ES can be impressive if the variable is difficult to change; for example a personality construct or an increase in life expectancy. However, if the focus of the investigation generated an incidentally large ES in a non-target variable, for example religious orientation, variable it should not necessarily be interpreted as being of practical value.

Cohen's *d*: measures the standardized difference between two means and conventional descriptions can be used to ascertain the effect size (small, 0.2; medium, 0.5, large, 0.8, very large 1.3). See Sullivan & Feinn (2012). Calculated using the Campbell Collaboration ES Calculator: F test, 2-group; unequal sample size (retrieved from https://campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD4.php) and from mean, standard deviation from University of Colorado Springs Cohens d calculator (retrieved from: https://www.uccs.edu/lbecker/).

In order to illustrate the magnitude of benefit compared to one another, the between-group effect sizes for the Treatment and Control arms at posttreatment was computed, using Cohen's d (Sherman, 1998), at T2 and T3. Effect sizes, are presented in the relevant tables with the initial Repeat Measures ANOVA for each measure.

- Sullivan, G. M. & Feinn, R. (2012). Using effect size—or why the p value is not enough. Journal of Graduate Medical Education, 4(3), 279-282.
- Sherman, J. J. (1998). Effects of psychotherapeutic treatments for PTSD: A meta-analysis of controlled clinical trials. Journal of Traumatic Stress, 11(3), 413-435.
- Wolff, F. M. (1986). Meta-Analysis: Qualitative Methods for Research Synthesis. Beverly Hills, CA: Sage.

2.3.2. Linear Mixed Models

Why use linear mixed model (LMM) procedures?

Summary of Reasons:

1. Perhaps the main advantage of using LMM is in being able to account for missing data in a way that would not be possible with GLM approaches which allow only a complete case analysis thereby not using all the data that is available (Heck et al, 2014). As such LMM offers more precise estimates of the differences between the two randomised control treatment arms. General Linear Models (RM-ANOVAs or ANCOVA) in respect repeated measures operate on a complete case analysis (CCA) to missing data and therefore discard subjects who may have a single missing measurement. LMM allow all of a subject's data to be used.
2. LMM also enables uneven spacing from repeated measurements to be handled within the model and can be extended to non-normal outcomes so that fixed and random effects can be considered with a hierarchy of levels (Seltman, 2018). This provides the option for simultaneous analyses of experimental effects and associated individual differences. As Matuschek, et al (2017) highlight a single LMM can replace what might otherwise be separate ANOVAs conducted on subjects or those conducted on items. However it is also noted that the disadvantages can also come at costs that come with LMM specification.
3. LMM allows an analysis of the relative contribution of multiple dependant and independent variables. This can help establish the efficacy of latent variables.
4. Blood et al (2010) conclude that LMM has good power coverage and it therefore more suitable to computing limited data sets compared to other statistical modelling, for example Structural Equation Modelling (SEM), particularly as the model with not be adding predictors where the univariate statics are not significant. Therefore, there is no rationale to attempt more statistical modelling.

5. A potential advantage of LMM is being able to compute the optimal time coding (Heck et al, 2014 pg. 227) within an analysis. Reference being made to the AIC and variance. As such it is possible to approximate the optimal estimate for the time coding. GLM procedures such as RM ANOVA assume spacing between measures is equidistant whereas LMM allows for spacing between time points to be varied and measured at an individual level.
6. GLM procedures consider growth trajectories in the aggregate and working on the assumption that there is a similar trending for all participants. LMM take into account that individuals may exhibit different growth curves.

- Blood, E. A., Cabral, H., Heeren, T., & Cheng, D. M. (2010). Performance of mixed effects models in the analysis of mediated longitudinal data. BMC Medical Research Methodology, 10:16.
- Heck, R. H., Thomas, S. L., Tabata, L. N. (2014). Multilevel and Longitudinal Modelling with IBM SPSS (2nd ed). NY, Routledge.
- Matuschek, H., Kligl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. Journal of Memory and Language, 94, 305–315.
- Seltman, H. J. (2018). Experimental Design and Analysis. Retrieved from : www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

LMM model design and selection

As Matuschek et al (2017) discuss there are several considerations that need to be accounted for when setting up an LMM procedure. Foremost includes ensuring a parsimonious mixed model that balances between Type 1 error and power considerations. To this end several different models

were trialled, particularly in respect to exploring how various variations in the time coding impacted on model fit, and these are documented below.

Model selection was primarily based on information criteria scores such as Akaike's Information Criterion (AIC) and the repeated covariate type selected accordingly. The PCL-C Total score was used to investigate model fit and this then guided model selection. A range of repeated covariance posibilities were also investigated at Level 1 using the AIC to ensure optimal model fit. A Diagonal covariance matrix was selected on this basis. Consideration was also given as to how well these covariance parameters reflected the research putpose. The final model was constructed with reference to Heck et al (2014 pg 233) 'Growth curve modelling with experimental groups in SPSS'. See also https://www.youtube.com/watch?v=Z2rxrQO1DAw&index=41&list=UU8r94_jZaoXv9qsgFwAdPQQ

Final LMM model

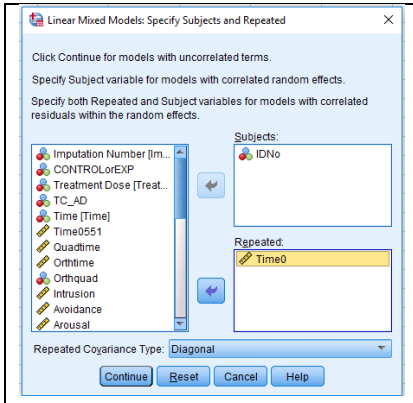
Model A: Time constructed as a categorical variable and as the reference group; intercept as the end status. This enables separate outcomes to be computed for T2 (in comparison to T0) and for T1 (in comparison to T0). See Heck et al (2014, pg 236) and Figure 2. AIC score 1463.775 based on 12 parameters.

Model B: Time constructed as a linear variable. Repeated covariate type: Diagonal. **Level 1: repeated measurement** of outcomes (significant difference in residuals at different time points i.e. T1, T2, T3 represented as T0, T1,T2); **Level 2:** based on subject level (participant ID) intercepts and slopes (growth trajectory and grand mean of the intercepts, initial status and expected variation in outcome scores as well as variation in their slopes). Fixed effects: Time, Treatment Arm, Interaction Treatment Arm * Time. Random effects: Intercept and individual subjects (see Figure 1).

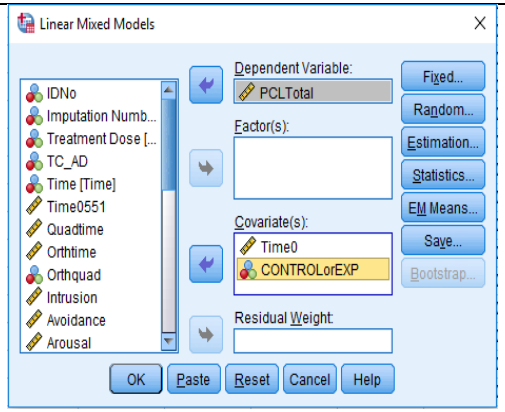
AIC score 1461.838 based on 8 parameters.

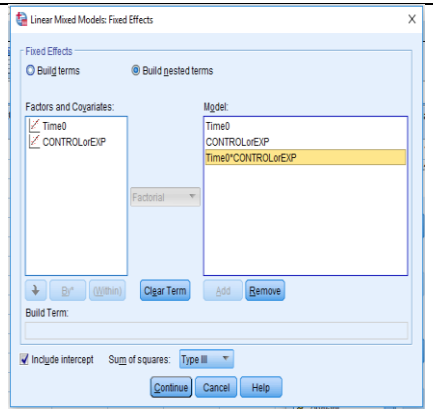
The SPSS dialogue boxes for LMM are also presented below for reference, ease of replication and supervision/peer scrutiny.

- Matuschek, H., Kligl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. Journal of Memory and Language, 94, 305–315.
- Heck, R. H., Thomas, S. L., Tabata, L. N. (2014). Multilevel and Longitudinal Modelling with IBM SPSS (2nd ed). NY, Routledge.

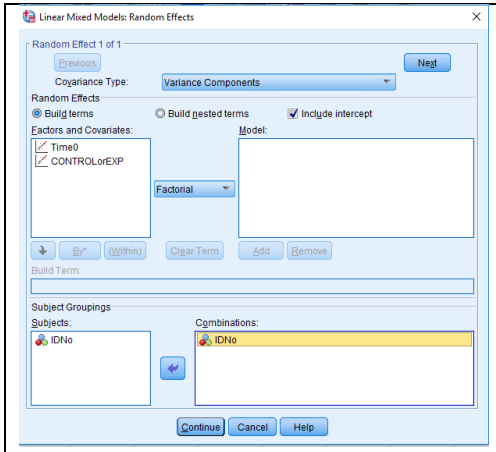


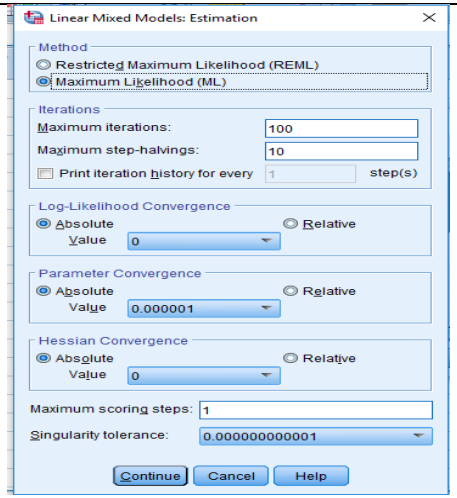
Repeated covariance type: *Diagonal*. This assumes unequal variances in the residuals across the time points and assumes no correlated residuals between time points. Choosing the repeated covariance structure also specified that the covariance was between time points rather than between the random intercept and random slope.

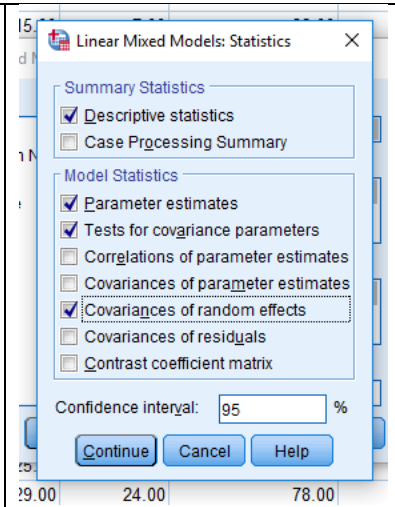




Fixed effects: Includes interaction (Treatment Arm * Time) as a nested term. Treatment is a fixed effect modelling the randomly varied slopes for time. Treatment as a predictor for the variation in the slope.







Random effects: Estimating the variance in participants intercepts from their growth trajectories as well as the variance in their slopes and intercepts. This helped to account for clustering in the data.
Select Variance Components = estimating both variance in slopes and slopes as well as the covariance between these two.
Time allowed to be a randomly varying component accounts for variation in terms of the slopes over time – participants can demonstrate variation in their linear trajectories

Figure 1: SPSS LMM procedures used to compute statistical significance of time (as a continuous variable) and treatment arm interaction as used for all measures and subscales (refered to as Modle A).

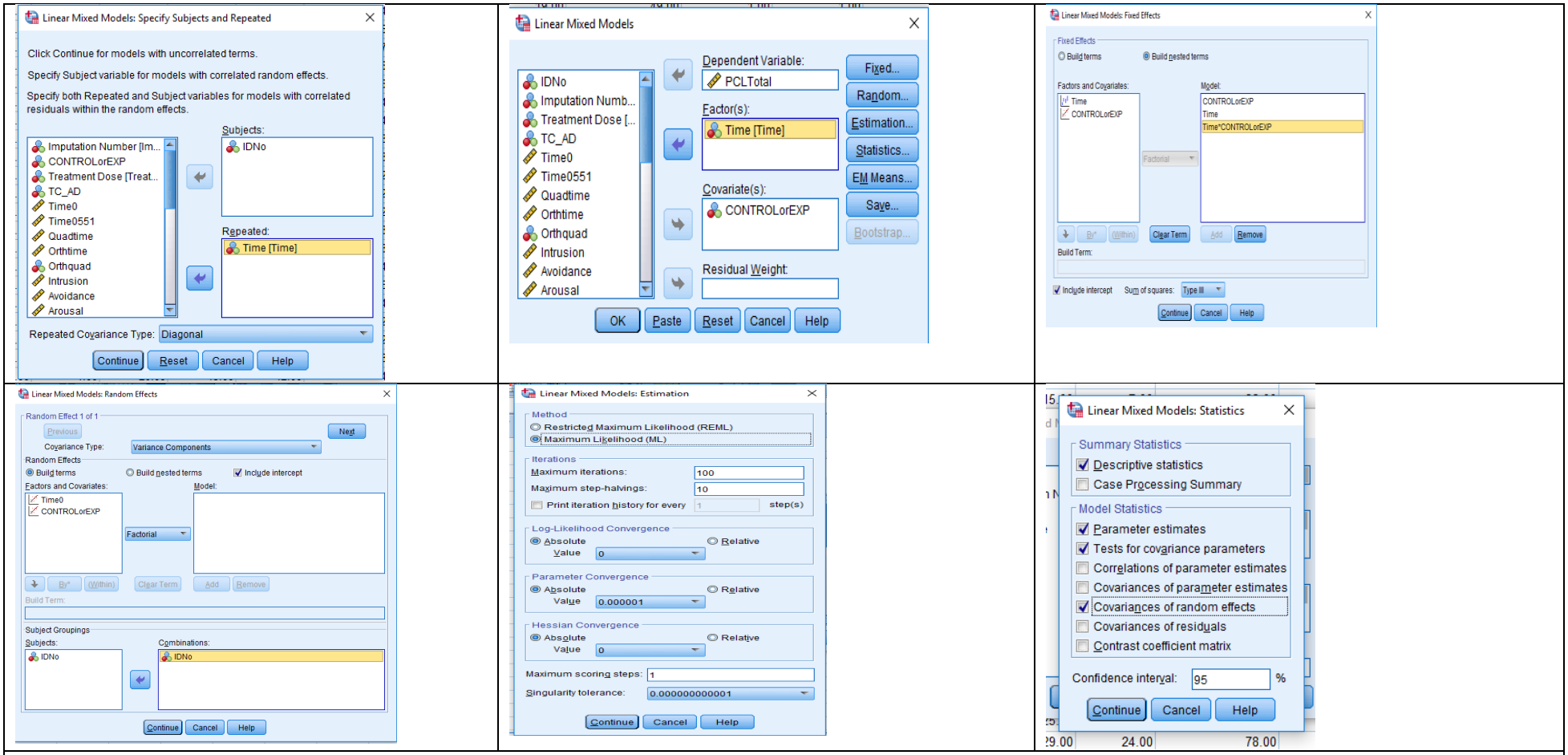


Figure 2: SPSS LMM procedures used to investigate the interaction between time and treatment arm interaction where time is a categorical variable (referd to as Modle B).

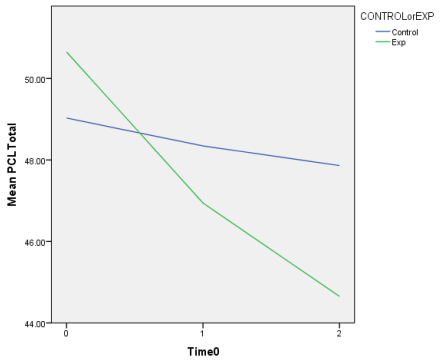
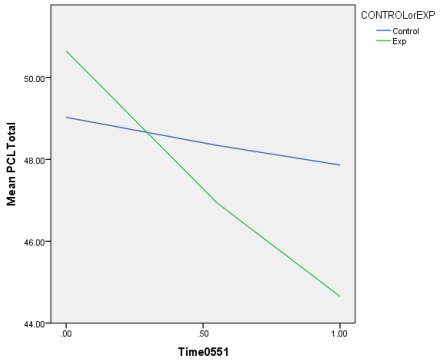
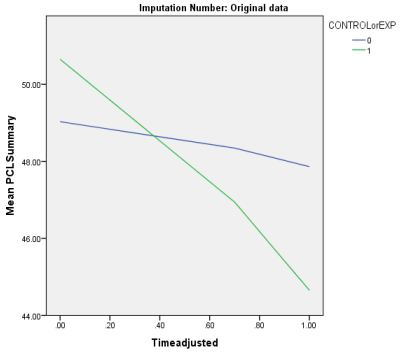
LMM time coding

Objective: To investigate whether model fit improved through the way in which change over time was quantified. Different time coding structures were investigated to find the most appropriate fit using the AIC model fit indices (see Figure 3). These have been graphed below to demonstrate the impact on PCL-C Total scores. Although there were only three time points a quadratic component was initially investigated in respect ascertaining the rate of change (e.g. accelerating or decelerating) as specified by Heck et al (2014 pg. 187). A curvilinear trend (i.e. nonlinear and non-incremental) was also investigated which show the greatest model fit (lowest AIC).

It was concluded that linear time coding (i.e. T0, T1, T2) appeared to be most monotonic and best fitted data from only three time points. In this respect none of the non-linear time coding from visual examination of the line graphs in Figure 3 appeared to have satisfactorily controlled for change between the time points. Indeed, whilst the non-linear time coding which emphasised T2 (i.e. 'post') treatment outcomes such as T0, T.80, T1 resulted in lower AIC scores they also skewed the overall model and resulted in a loss of monotonicity (i.e. away from a linear trajectory).

Whilst there was some improvement in model fit with a slight variation in the linear time coding T0, T.55, T1 it was subsequently found that this was not found for all measures and subscales. Therefore, a standardised linear approach was utilised as an important basis from which to interpret results. It may be that the time coding T0, T.55, T1 does reflect the slightly faster rate of change between the first two time points and any variations in statistical significance were checked where appropriate.

Whilst the graphs presented below in Figure 3 demonstrate a time by study arm interaction (as the lines representing the rates of change for each arm cross over each other) it is noted that this not a statically significant as the WL/TAU arm is also decreasing albeit at a slower rate.

Linear T1=0, T2=1, T3=2 (AIC: 1461.838) Time* Arm: (β= -1.47, 95% CI= -4.30 to 1.36, p= .303) 	**Time nonlinear 0.55** T1=0, T2=.55, T3=1 (AIC: 1461.421) Time* Arm: (β= -3.00, 95% CI= -8.63 to 2.63, p= .291) 	**Time nonlinear 0.70** T1=0, T2=.70, T3=1 (AIC: 1460.434) Time* Arm: (β= -3.03, 95% CI= -8.46 to 2.40, p= .269) 
Time nonlinear .80 T1=0, T2=.80, T3=1 (AIC: 1460.033) Time* Arm: (β= -2.94, 95% CI= -8.16 to 2.28, p= .265)	**Quadtime** T1=0, T2=1, T3=4 (AIC: 1464.155) Time* Arm: (β= -.56, 95% CI= -1.93 to .80, p= .408)	

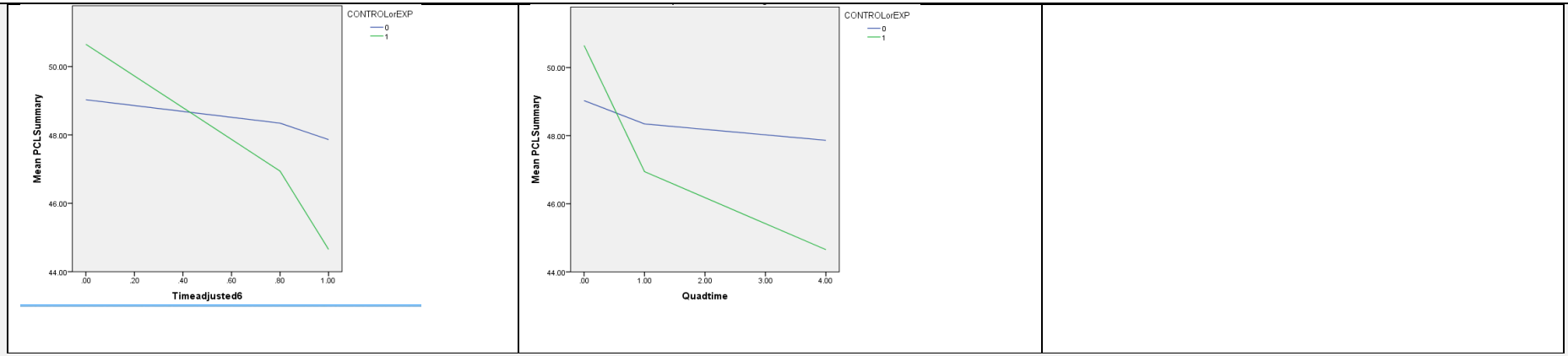


Figure 3: Investigation of different time coding procedures on PCL-C total scores.

2.3.3. Interpretation of LMM output

(Model A)

Estimation of fixed effects

Imputation Number	Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Original data	Intercept	49.564414	2.783796	91.456	17.805	.000	44.035117	55.093712
	CONTROLorEXP	.164109	3.731407	91.411	.044	.965	-7.247424	7.575641
	Time0	-.806932	1.022809	58.477	-.789	.433	-2.853951	1.240088
	CONTROLorEXP * Time0	-1.468265	1.413630	60.176	-1.039	.303	-4.295776	1.359245

Intercept: predicted outcome score (estimate) at the initial status for participants in the WL/TAU condition.

Time: predicted outcome score across all participants. Predicted rate of decrease (-.81) for every unit increase in time. In this case it was not a

significant predictor in the model

CONTROLorEXP [Treatment Arms]: The difference at the initial status in means (between WL/TAU and S&T); would expect it to be non-significant as randomly assigned

CONTROLorEXP*Time0: The interaction effect i.e. the difference between treatment and WL/TAU groups. The WL/TAU group being the reference category (coded as 0). In this case for every unit increase in the interaction term there is a decrease (-1.47) in difference between groups as time moves from one time point to another. This highlights whether (or not) the cross level interaction in treatment is a predictor in the level of symptom reduction. This is a non-significant predictor and therefore there was no significant difference over time in outcome measures between the study arms.

Estimates of covariance parameters

Imputation Number	Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Original data	Repeated Measures	Var: [Time0=0]	60.600994	16.097464	3.765	.000	36.005991	101.996371
		Var: [Time0=1]	52.605782	14.710095	3.576	.000	30.409757	91.002642
		Var: [Time0=2]	39.159110	14.155732	2.766	.006	19.281005	79.530910
	Intercept [subject = IDNo]	Variance	218.193174	39.228485	5.562	.000	153.392969	310.367949

Repeated measures: Significant differences in the residuals are evident across the time points at the various measurement occasions for the outcome measure.

Wald Z: indicates whether the explanatory predicors are significant and whether they add something to the model and whether they should be retained. See http://www.statisticshowto.com/wald-test/

Intercept: Level 2; grand intercept mean significantly accounts almost all of the variance in the model.

3. Appendix for results: Randomised control trial

Contents

		page
3.1.	Complete results for the Trauma Antecedent Questionnaire (TAQ)	269
3.2.	Complete results for the SIDES –Self Report	270
3.3.	Missing data analysis	273
3.4.	Investigation into quality assurance of treatment facilitation	278
3.5.	Differences between intervention arms for outcome measures	279
3.5.1	PTSD Checklist –Civilian Version	279
3.5.2.	Behavioural Assessment Checklist-Revised	282
3.5.3.	Difficulties in Emotional Regulation Scale	286
3.5.4.	Dissociative Experiences Scale	289
3.5.5.	Hospital Anxiety Depression Scale	292
3.5.6.	Criminal Cognitions Scale	293
3.6.	Differences between adequate dose and non-adequate dose participants	296
3.6.1	PTSD Checklist –Civilian Version	296
3.6.2	Behavioural Assessment Checklist-Revised	299
3.7.	Differences between adequate dose and WL/TAU participants	303
3.7.1	PTSD Checklist –Civilian Version	303

3.7.2	Behavioural Assessment Checklist-Revised	305
3.7.3	Difficulties in Emotional Regulation Scale	307
3.7.4	Dissociative Experiences Scale	310
3.7.5.	Hospital Anxiety Depression Scale	312
3.7.6.	Criminal Cognitions Scale	313
3.8.	Plots for reliable change scores	216

Abbreviations used:

ITT	Intent to Treat
MI	Multiple imputation
T1	Time Point 1 or 'pre' intervention measure
T2	Time Point 2 or 'post' intervention measure
T3	Time Point 3 or 'FU' intervention measure at one moth post intervention
S&T	Survive & Thrive: experimental intervention arm
AD	Adequate dose participants
Non-AD	Non-adequate dose participants

3.1. Complete results for the Trauma Antecedent Questionnaire (TAQ)

Table 1: TAQ mean scores for positive and negative events across age categories and lifetime presence										
	0-6 yrs		7-12 yrs		13-18 yrs		> 19 yrs (Adult)		Lifetime	
	M (SD)	N (%)	M (SD)	N (%)	M (SD)	N (%)	M (SD)	N (%)	M (SD)	N (%)
Competence										
WL/TAU	1.47(1.40)	23(54.8)	2.22(1.14)	32(76.2)	2.30(0.90)	36 (85.7)	2.07(1.09)	33(78.6)	8.05(3.17)	37(88.1)
S&T	1.97(1.31)	31(70.5))	2.59(0.58)	40(90.9)	2.31(0.89)	37 (84.1)	1.86(1.11)	32(72.7)	8.74(2.41)	40(91.0)
Safety										
WL/TAU	2.43(1.03)	35(83.3)	2.47(0.92)	36(89.2)	2.24(1.07)	32 (83.8)	2.39(1.01)	35(83.3)	9.53(3.12)	38 (90.5)
S&T	2.53(1.02)	35(86.8)	2.43(1.01)	37(86.8)	2.67(0.53)	39(97.4)	2.40(1.00)	35(79.5)	10.04(2.59)	40(97.4)
Neglect										
WL/TAU	1.03(1.35)	17(40.5)	1.48(1.36)	24(57.1)	2.07(1.14)	31 (73.8)	2.02(1.27)	29(69.0)	6.61(4.14)	33(78.6)
S&T	0.99(1.33)	17(38.6)	1.47(1.30)	25(56.8)	2.01(1.20)	31(70.5)	2.37(0.92)	37(84.1)	6.84(3.14)	41(93.2)
Separation										
WL/TAU	1.48(1.47)	21(50.0)	2.14(1.26)	31(73.8)	2.56(0.93)	36(85.71)	2.80(0.69)	38(90.5)	8.98(3.20)	38 (90.5)
S&T	1.53(1.48)	23(52.1)	2.29(1.22)	34(77.3)	2.58(0.93)	37(84.1)	2.64(0.93)	37(84.1)	9.04(3.03)	42 (95.5)
Emotional Abuse										
WL/TAU	0.89(1.33)	14(33.3)	1.55(1.40)	22(52.3)	1.93(1.25)	30(71.4)	2.05(1.16)	32(76.2)	6.42(4.10)	34 (80.1)
S&T	1.00(1.34)	16(36.4)	1.73(1.30)	28(63.6)	1.85(1.25)	29(59,1)	2.09(1.15)	33(75.0)	6.66(4.00)	36(81.8)
Physical Abuse										
WL/TAU	0.76(1.28)	13(31.0)	1.21(1.37)	19(45.2)	1.58(1.37)	25(59.5)	1.73(1.33)	25 (59.5)	5.27(4.23)	30(71.4)
S&T	0.64(1.17)	12(27.3)	0.90(1.23)	17(38.6)	1.57(1.34)	26(59.1)	1.73(1.32)	28(63.6)	4.84(3.21)	38(86.4)
Sexual abuse										
WL/TAU	0.68(1.23)	12(28.6)	0.93(1.30)	16(38.1)	1.23(1.38)	19 (45.2)	0.90(1.33)	15(35.7)	3.73(3.86)	27(64.1)
S&T	0.36(0.94)	9(20.5)	0.82(1.25)	14(31.8)	0.64(1.19)	13 (29.5)	0.34(0.90)	8(13.2)	2.32(2.58)	24(54.5)
Witnessing										

WL/TAU	1.29(1.44)	19(45.2)	1.96(1.25)	29(69.0)	2.33(1.09)	34(81.0)	2.10(1.18)	31(73.8)	7.68(3.75)	35(83.3)
S&T	1.37(1.42)	23(52.3)	1.73(1.37)	28(63.6)	1.74(1.47)	26(59.1)	1.74(1.37)	27(61.4)	6.49(3.83)	37(84.1)
Other trauma										
WL/TAU	1.04(1.44)	16(38.1)	1.43(1.44)	20(46.6)	1.66(1.37)	25(59.5)	2.53(0.94)	36(85.7)	6.70(3.20)	38(90.5)
S&T	0.86(1.37)	13(29.5)	1.23(1.41)	22(50.0)	1.71(1.36)	27(61.4)	1.79(1.34)	28(63.6)	5.60(3.91)	37(84.1)
Alcohol/drug abuse										
WL/TAU	1.24(1.46)	19(45.2)	1.80(1.44)	26(61.9)	2.36(1.01)	34(81.0)	2.49(1.37)	35(83.3)	7.89(4.06)	38(90.5)
S&T	1.21(1.45)	19(43.2)	1.49(1.45)	23(52.3)	2.20(1.20)	33(75.0)	2.21(1.20)	33(75.0)	7.11(4.24)	35(79.5)

3.2. Complete results for the SIDES –Self Report

Table 2: SIDES-SR mean scores for lifetime occurrence, current presence and current severity

SIDES symptom clusters	WL/TAU (n= 42)			S&T (n= 44)		
	Lifetime Occurrence[1] (N/%)	Current Presence[1] N/%)	Current Severity M(SD)	Lifetime Occurrence[1] (N/%)	Current Presence[1] N/%)	Current Severity M(SD)
Alterations in affect regulation	**28 (66.7)**	**22 (52.4)**	**0.80 (0.74)**	**32 (72.7)**	**20 (45.5)**	**0.71 (.64)**
a) Affect regulation	33 (78.6)	27 (64.3)	1.61 (1.36)	33 (75.0)	23 (52.3)	1.36 (1.16)
b) Anger modulation	21 (50.0)	16 (38.1)	0.90 (1.16)	25 (56.8)	16 (36.4)	.88 (1.05)
c) Self-destructiveness	28 (66.7)	16 (38.1)	0.93 (1.21)	33 (77.3)	13 (29.5)	0.82 (1.17)
d) Suicidal preoccupation	25 (59.5)	3 (7.1)	0.19 (0.71)	23 (52.3)	2 (4.5)	0.12 (0.54)
e) Sexual preoccupation	31 (72.7)	15 (35.7)	0.93 (1.23)	30 (68.2)	13 (29.5)	0.75 (1.06)
f) Excessive risk taking	23 (54.8)	4 (9.5)	0.24 (0.79)	22 (56.8)	6 (13.6)	0.32 (0.76)
Alterations in Consciousness	**38 (90.5)**	**25 (59.5)**	**1.02 (0.96)**	**39 (88.6)**	**23 (52.3)**	**0.97 (0.97)**
a) Amnesia	28 (66.7)	14 (33.3)	0.80 (1.15)	32 (72.7)	11 (25.0)	0.68 (1.09)
b) Dissociative episodes	35 (83.6)	22 (52.4)	1.24 (1.18)	39 (88.6)	21 (47.7)	1.25 (1.17)

Alterations in self-perception	**35 (83.3)**	**18 (42.9)**	**0.68 (0.74)**	**40 (90.9)**	**19 (43.2)**	**0.65 (0.58)**
a) Ineffectiveness	22 (52.4)	9 (21.4)	0.46 (0.87)	24 (54.5)	9 (20.5)	0.47 (0.86)
b) Permanently damaged	27 (64.3)	17 (40.5)	1.05 (1.28)	30 (68.2)	19 (43.2)	1.19 (1.26)
c) Guilt	32 (76.2)	15 (35.7)	0.95 (1.29)	34 (77.3)	14 (31.8)	0.85 (1.15)
d) Shame	22 (52.4)	12 (28.6)	0.73 (1.17)	14 (31.8)	8 (18.2)	0.49 (0.94)
e) Nobody understands	26 (61.9)	9 (21.4)	0.56 (1.08)	28 (63.6)	12 (27.3)	0.69 (1.06)
f) Minimising risk	18 (42.9)	7 (16.7)	0.36 (0.82)	17 (38.6)	3 (6.8)	0.18 (0.61)
Alterations in relationships	**39 (92.9)**	**25 (59.5)**	**0.82 (0.81)**	**43 (97.7)**	**27 (61.4)**	**0.79 (0.70)**
a) Inability to trust	38 (90.5)	23 (54.8)	1.40 (1.29)	43 (97.7)	25 (56.8)	1.53 (1.21)
b) Re-victimisation	29 (69.0)	12 (28.6)	0.83 (1.31)	25 (56.8)	11 (25.0)	0.78 (1.23)
c) Victimising others	16 (38.1)	5 (11.9)	0.24 (0.65)	15 (34.1)	2 (4.5)	0.11 (0.42)
Somatization	**21 (50.0)**	**5 (11.9)**	**0.27 (0.48)**	**15 (34.1)**	**11 (25.0)**	**0.36 (0.51)**
a) Digestive system	11 (26.2)	5 (11.9)	0.27 (0.73)	8 (18.2)	7 (15.9)	0.37 (0.79)
b) Chronic pain	16 (38.1)	10 (23.8)	0.56 (1.01)	19 (43.2)	13 (29.5)	0.70 (0.99)
c) Cardiopulmonary	13 (31.0)	3 (7.1)	0.15 (0.52)	12 (27.3)	5 (11.4)	0.29 (0.78)
d) Conversion systems	17 (40.5)	4 (9.5)	0.22 (0.68)	13 (29.5)	4 (9.1)	0.20 (0.58)
e) Sexual systems	8 (19.5)	2 (4.8)	0.12 (0.55)	6 (13.6)	2 (4.5)	0.10 (0.42)
Alterations in meaning	**35 (83.3)**	**13 (31.0)**	**0.47 (0.76)**	**37 (84.1)**	**20 (45.5)**	**0.64 (0.59)**
a) Despair	31 (73.8)	11 (26.2)	0.64 (1.08)	33 (75.0)	13 (29.5)	0.75 (1.06)
b) Loss of beliefs	28 (66.7)	5 (11.9)	0.29 (0.80)	29 (65.9)	9 (20.5)	0.52 (0.96)
Note: 1. Lifetime occurrence and current presence = figures represent the affirmative (i.e. 'yes'). Symptom data in **bold** denotes meeting SIDES-SR diagnostic criteria for symptom domain						

Baseline demographic and trauma history differences between the arms were initially tested and found not to be significant. As described by de Boer, Waterlander, Kujper, Steenuis & Twisk (2015) randomisation accounted for equal variance between the two

arms and therefore presentation of baseline testing was not a logical requirement of the CONSORT statement and removed.

- de Boer, M. R., Waterlander, W. E., Kuijper, L. D. J., Steenhuis, I. H. M. & Twisk, W. R. (2015) Testing for baseline differences in randomised control trials: an unhealthy research behaviour that is hard to eradicate. *Physical Activity*, 12:4

Table 3: Number of participants meeting the current diagnostic criteria for symptom domains in SIDES-SR

	WL/TAU		**S&T**	
No. of Symptom Domains	N	%	N	%
0	13	31.0	10	22.73
1	3	7.14	4	9.09
2	7	16.67	10	22.73
3	6	14.29	8	18.18
4	8	19.05	6	13.64
5	3	7.14	4	9.09
6*	2	4.76	1	2.27

Note: based on MI data * Full diagnostic criteria as proposed by Luxenberg, Spinazzola & van der Kolk (2001, pg 374) for Disorders of Extreme Stress, Not Otherwise Specified (DESNOS)

- Luxenberg, T., Spinazzola J. & van der Kolk, B. A. (2001). Complex Trauma and Disorders of Extreme Stress (DESNOS) Diagnosis, Part One: Assessment. Directions in Psychiatry, 21, (25), 373- 395.

3.3. Missing data analysis

Table 4. Comparisons of demographic and baseline differences for outcome completers and missing data participants for the PCL-C										
	Post (T2)					**Follow Up (T3)**				
Variables *Missing* *Completer*	Arms		Comparisons[1]			Arms		Comparisons[1]		
	WL/TAU M (SD) / N (%)	S&T M(SD) / N(%)	Arms	Completion	Arms x Completion	WL/TAU M(SD) / N(%)	S&T M(SD) / N(%)	Arms	Completion	Arms x Completion
Demographics and Sentencing Characteristics										
Age *Missing* *Completer* (years)	28.93 (8.29) 35.29 (10.71)	30.56(10.25) 34.56 (10.37)	.015(1, 82) .902	3.98(1,82) .049	.27(1, 82) .605	28.70(7.43) 37.23(11.03)	34.48 (10.96) 32.65 (9.90)	.078 (1,82) .781	2.42(1,82) .124	5.77 (1,82) .019
Length of Sentence *Missing* *Completer* (months)	27.43(20.11) 91.75(70.54)	56.33(61.74) 48.14(51.95)	.274(1, 82) .602	3.98(1,82) .049	6.67 (1,82) .012	30.75 (20.94) 106.27 (72.42)	36.29 (43.40) 62.17 (59.44)	2.80(1,82) .098	19.33(1,82) .000	4.63(1,82) .034
No. of convictions *Missing* *Completer* (previous convictions)	15.14 (23.65) 5.79 (11.74)	5.44(4.92) 6.09(15.42)	1.52(1,82) .221	1.31(1,82) .256	1.72(1, 82) .193	11.30(20.54) 6.73 (13.05)	3.95 (4.51) 7.78 (18.73)	.883(1,82) .350	.012(1,82) .912	1.57(1,82) .218
Sentence Type (Index categories)	Numerous	Numerous	4.66(5) .459	1.54(5) .909		Numerous	Numerous	18.03(5) .003	1.54(1) .909	
Drug conviction[2] Missing Completer	5 (35.7%) 6 (21.4%)	2 (22.2) 10 (28.6%)	.985(1) .321	.013(1) .910		6 (30.0) 5 (22.7)	6(28.6) 6 (26.1)	.287(1) .592	.013(1) .910	

Violence conviction[2] Missing Completer	10 (71.4%) 21 (75.0%)	6 (66.7%) 26 (74.3%)	.218(1) .640	.013(1) .910		15 (75.0) 16 (72.7)	12 (57.1) 20 (87.0)	2.191(1) .139	.013(1) .910	
Mental Health[2] Missing Completer	10(71.4) 21 (75.0)	7 (77.8) 19 (54.3)	.819(1) .366	2.08(1) .149		14 (70.0) 17 (77.3)	15 (71.4) 11 (47.8)	.695(1) .404	2.083(1) .149	
Facilitation and Participation: S&T arm only										
No. of sessions Missing Completer (sessions attended)		4.17(3.0) 8.51(1.93)		22.01(1,39) .000			7.00(3.39) 8.57 (1.50)		3.96(1,39) .054	
Quality Assurance Missing Completer (Total Score)		3.42 (1.02) 3.59 (1.02)		.188(1,42) .667			3.45 (1.59) 3.65 (.97)		.411(1,42) .525	
Baseline Measures										
TAQ Child Total[2] Missing Completer	5.27 (1.97) 6.32 (1.79)	4.80 (1.97) 5.98 (1.79)	.000(1,75). 994	3.744(1,75) .343	.182(1,75) .671	5.91 (1.85) 6.03 (1.92)	5.58 (1.98) 5.88 (1.92)	.017(1.75) .898	.002(1,75) .966	.014(1,75) .906
TAQ Adult Total[2] Missing Completer	5.11 (2.20) 5.99 (1.43)	4.49 (1.30) 5.39 (1.84)	.783(1,75) .379	.889(1,75) .349	.158(1,75) .693	5.38(1.93) 5.98(1.47)	5.21(1.73) 5.21(1.47)	1.309(1,75) .256	.136(1,75) .714	.997(1,75) .321
TAQ Lifetime Total[2] Missing Completer	5.70 (1.90) 6.84 (1.31)	5.42 (1.22) 6.79 (1.22)	.452(1,75) .504	.910(1,75) .343	1.21(1,75) .274	6.19 (1.73) 6.70 (1.36)	6.30 (1.09) 6.64 (1.32)	.093(1,75) .761	.414(1,75) .522	.370(1,75) .545
SIDES Current										

Total[2] Missing Completer	.923 (.393) 1.71(.267)	1.00(.535) 1.36(.246)	.131(1,77) .719	2.33|(1,77) .131	.319(1,77) .574	1.42 (1.22) 1.50 (1.63)	1.26 (1.28) 1.33 (1.56)	.255(1,77) .615	.054(1,77) .817	.000(1,77) .989
SIDES Lifetime Total[2] Missing Completer	2.15(.57) 3.14(.38)	2.14 (.77) 3.06 (.36)	.007(1,77) .932	.004(1,81) .085	.004(1,77) .948	2.74 (1.97) 2.91 (2.31)	2.47 (1.72) 3.29 (2.17)	.015(1,77) .902	1.145(1,81) .288	.484(1,81) .489

Notes: 1. Comparison statistics F(df) p value / χ^2 (df) p value. 2. Answers presented in the affirmative (i.e. Yes)

Table 5. Comparisons of demographic and baseline differences for outcome completers and missing data participants for the BAC-R staff rated measure										
	Post (T2)					**Follow Up (T3)**				
Variables *Missing* *Completer*	Arms		Comparisons[1]			Arms		Comparisons[1]		
	WL/TAU M (SD) / N (%)	S&T M (SD) / N (%)	Arms	Completion	Arms x Completion	WL/TAU M (SD) / N (%)	S&T M (SD) / N(%)	Arms	Completion	Arms x Completion
Demographics and Sentencing Characteristics										
Age Missing Completer	29.05 (7.86) 36.91 (11.91)	34.12 (11.59) 33.15 (9.68)	.09 (1,82) .767	2.45(1,82) .121	4.02 (1, 82) .048	30.21 (9.31) 37.11 (10.53)	33.78 (10.81) 33.12 (9.85)	.01 (1,82) .925	1.96 (1,82) .166	2.87 (1,82) .094
Length of Sentence Missing Completer	42.65(54.53) 95.45(66.38)	31.00 (27.10) 61.67 (62.39)	3.38 (1,82) .070	11.40(1,82).001	.80(1,82) .373	40.63 (47.98) 109.89 (66.84)	32.70 (31.45) 77.00 (69.17)	3.06(1,82) .084	23.69(1,82) .000	1.15 (1,82) .288
No. of convictions Missing Completer (previous convictions)	6.65(12.70) 10.95(20.18)	4.35 (4.94) 6.96 (17.35)	.85 (1,82) .358	1.10 (1,82) .312	.06 (1,82) .804	9.58 (19.08) 8.00 (14.17)	3.78 (4.26) 9.41 (21.62)	.415(1,82) .521	.35 (1,82) .554	1.12 (1,82) .293
Sentence Type (Index categories)	Numerous	Numerous	.707(1) 0.400	2.17 (4) .704		Numerous	Numerous	18.90(5) .002	1.54 (5) .909	
Drug conviction[2] Missing Completer	7 (35.0) 4 (18.2)	5(29.4) 7 (25.9)	1.53 (1) .216	.13 (1) .717		7 (29.2) 4 (22.2)	8 (29.6) 4 (23.5)	.257(1) .6127	.01 (1) .910	
Violence conviction[2] Missing Completer	14 (70,0) 17 (77.3)	9 (52.9) 23 (85.2)	4.079(1) .043	.01 (1) .910		17 (70.8) 14 (77.8)	15 (55.6) 17 (100)	7.07 (1) .008	.01 (1) .910	
Mental Health[2] Missing Completer	13 (65.0) 18(81.8)	11 (64.7) 15 (55.6)	1.53 (1) .216	2.08 (1) .149		15 (62.5) 16 (88.9)	16 (59.3) 10 (58.8)	3.705(1) .054	2.08 (1) .149	

Facilitation and Participation										
No. of sessions Missing Completer (sessions attended)		7.36 (3.41) 8.15 (2.07)		.86 (1,39) .360			7.50 (3.13) 8.41(1.46)		1.24 (1,39) .272	
Quality Assurance Missing Completer (Total Score)		3.45 (1.07) 3.62(.97)		.310 (1,42) .581			3.59 (1.18) 3.50 (.65)		.09 (1,42) .765	
Baseline Measures										
TAQ Child Total[2] Missing Completer	5.91 (1.88) 6.03 (1.88)	5.44 (2.12) 5.93 (2.22)	.048(1,71) .828	.048(1,71) .828	.215(1,71) .644	5.78 (1.78) 6.22 (1.99)	5.53 (2.21) 6.06 (214)	.005 (1,71) .941	.240(1,71) .626	.018 (1,71) .895
TAQ Adult Total[2] Missing Completer	5.36 (1.81) 6.00 (1.66)	5.07 (1.89) 5.30 (1.73)	1.318(1,71) .255	.018(1,71) .895	.034(1,71) .105	5.30(1.86) 6.22(1.51)	5.34 (1.64) 5.00 (1.93)	1.86 (1,71) .177	.001 (1,71) .981	1.842 (1,81) .179
TAQ Lifetime Total[2] Missing Completer	6.06 (1.53) 6.82 (1.56)	6.12 (1.19) 6.71 (1.24)	.039(1,71) .844	.014 (1,71) .907	.051 (1,71) .820	6.00 (1.67) 7.05 (1.39)	6.36 (.98) 6.67 (1.50)	.044 (1,71) .835	.146(1,71) .704	.592 (1,71) .444
SIDES Current Total[2] Missing Completer	1.63 (1.38) 1.31 (1.49)	1.21 (1.25) 1.34 (1.52)	.352(1,77). 555	.077 (1,77) .783	.460 (1,77) .499	1.65 (1.40) 1.22 (1.48)	1.26 (1.25) 1.35(1.66)	.164 (1,77) .681	.275 (1,77) .657	.657 (1,77) .420
SIDES Lifetime Total[2] Missing Completer	2.95 (2.15) 2.73 (2.16)	2.50 (1.74) 3.12 (2.10)	.004 (1,77) .950	.175 (1,77) .677	.780 (1,77) .380	2.91 (2.07) 2.72 (2.27)	2.70 (1.72) 3.18 (2.32)	.064(1,77) .800	.097(1,77) .757	.518 (1,77) .474

Notes: 1. Comparison Statistics F(df) p value / χ^2 (df) p value. 2. Answers presented in the affirmative (i.e. yes)

3.4. Investigation into quality assurance of treatment facilitation

Table 6. Regression analysis of treatment quality scores to ascertain significant predictors of outcome scores at T2							
Measure	**R^2**	**F**	**df**	**p**	**B**	**t**	**p**
PCL	.012	.409	1,33	.527	1.604	.706	.480
HADS	.000	.013	1,33	.908	-.119	-.218	.828
DERS	.018	.588	1,33	.449	3.251	.839	.402
DES II	.064	2.251	1,33	.143	6.088	1.911	.057
CCS	.093	3.365	1,33	.076	1.978	1.781	.075
BAC	.087	2.385	1,25	.135	5.466	1.888	.059
Notes: Pooled MI data used where possible; not possible for F tests.							

Simple linear regression was used to assess whether treatment quality significantly predicted outcome in the above measures at T2. The results of the regression suggested that treatment quality explained only 1% of the variance, R^2= .01, *F* (,33)= .409, *p*= .527. Treatment quality does not significantly predict outcome on the PCL-C, *B*= 1.604, *t*= .706, *p*= .408

3.5. Differences between intervention arms for outcome measures

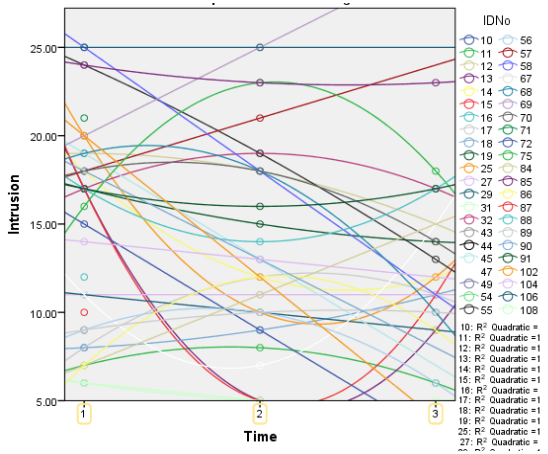
3.5.1. PTSD Checklist –Civilian Version (PCL-C)

Table 7. Raw mean scores for PCL-C						
	WL/TAU M (SD)			S&T M (SD)		
Subscales / N	34	29	21	43	35	22
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Intrusion	15.85(5.48)	15.34(5.62)	13.73(5.83)	14.81(6.00)	13.54(5.75)	12.43(4.29)
Avoidance	18.71(6.67)	18.31(6.77)	20.18(7.94)	20.12(6.86)	18.89(7.26)	17.87(6.61)
Arousal	14.47(5.74)	14.69(5.74)	13.95(6.20)	15.72(5.44)	14.51(6.36)	14.34(5.01)
Total PCL-C	49.03 (16.19)	48.34 (16.85)	47.86 (18.04)	50.56 (16.56)	46.94 (16.62)	44.65 (14.41)

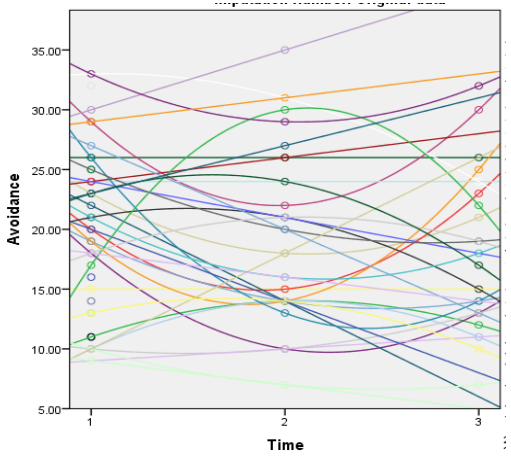
Table 8. Intent-to-Treat (Multiple Imputation) mean scores for PCL-C						
	WL/TAU M (SE)			**S&T** M (SE)		
Subscales / N	42	42	42	44	44	44
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Intrusion	15.92(.82)	15.33(.89)	13.47(.87)	14.82(.89)	14.04(1.00)	12.57(.92)
Avoidance	18.80(.94)	18.32(.91)	19.68(.96)	20.10(1.02)	18.66(1.00)	18.34(.81)
Arousal	14.62(.82)	14.58(.79)	13.97(.75)	15.68(.81)	14.54(.80)	14.35(.66)
Total PCL-C	49.35(2.26)	48.24(222)	47.11(2.05)	50.60(2.47)	47.24(2.27)	45.26(1.67)

Note: SE only provided for pooled MI data

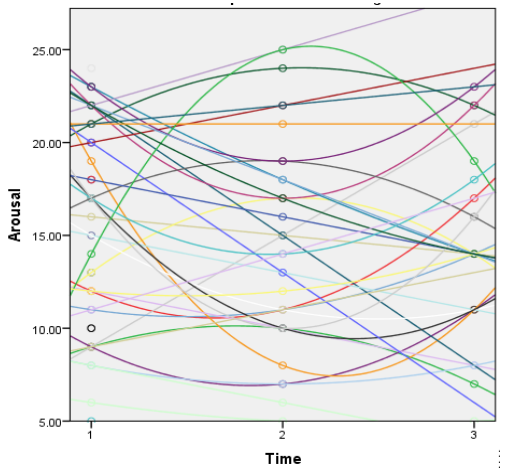
Intrusion



Avoidance



Arousal



Total PCL-C

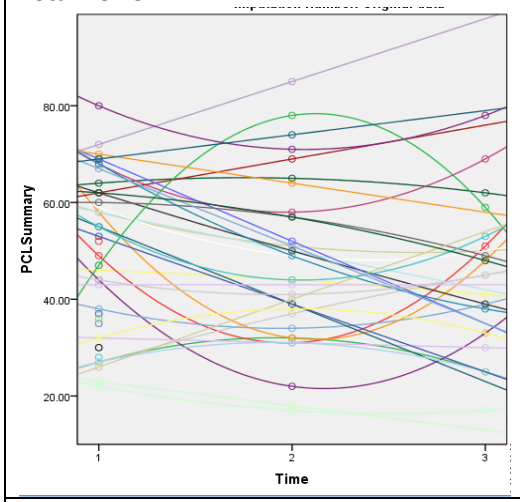


Figure 1: Individual S&T arm participants outcome trajectories for PCL-C total and subscales investigating quadratic growth trends

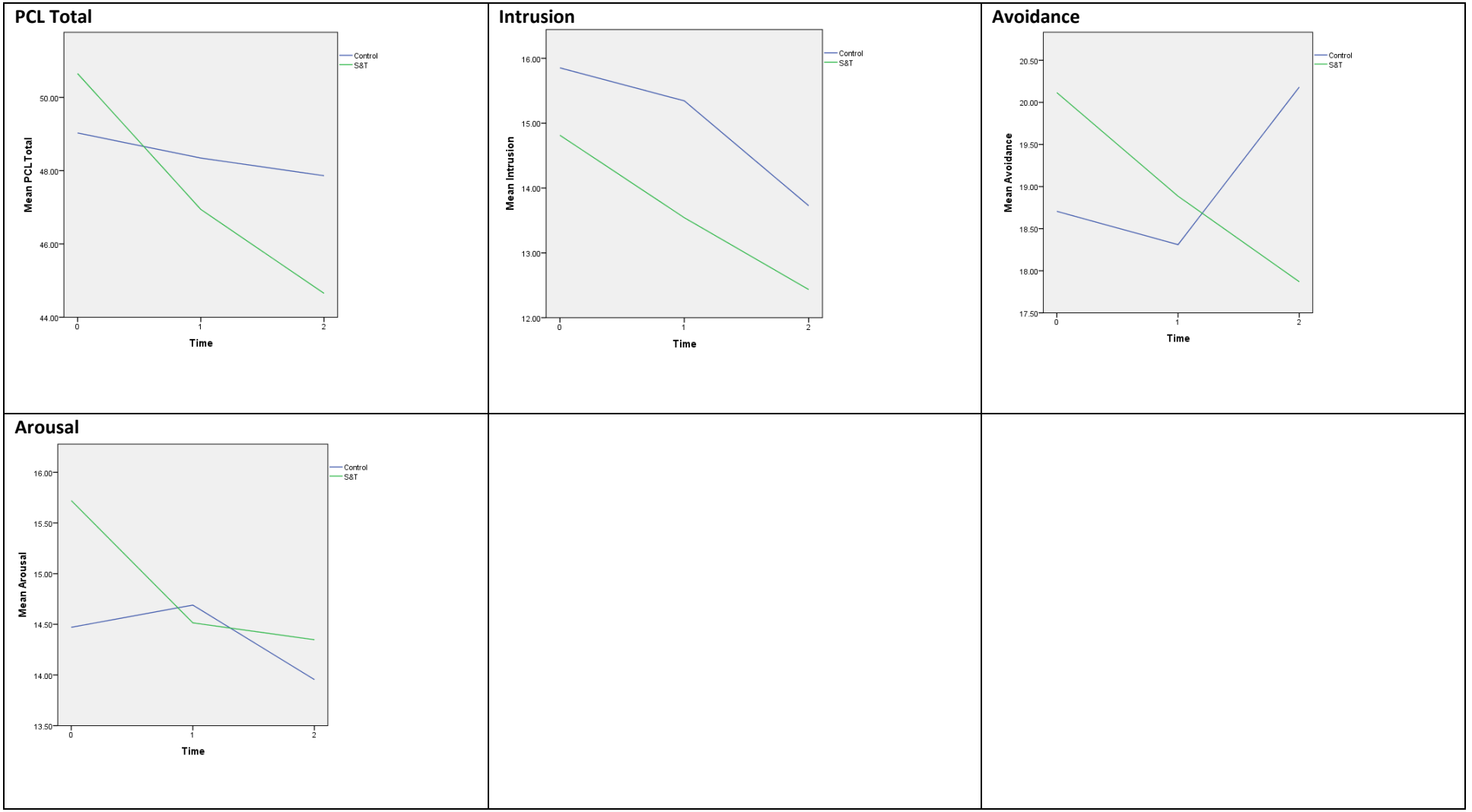


Figure 2: ITT means for PCL-C total and subscales for outcomes at T1 – T3.

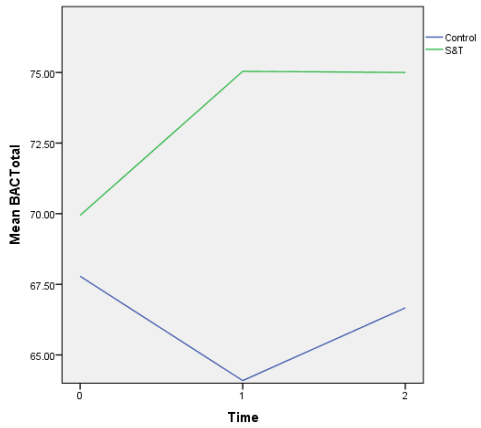
3.5.2. Behavioural Assessment Checklist- Revised (BAC-R)

Table 9. Raw mean scores for BAC-R						
	WL/TAU M (SD)			**S&T** M (SD)		
Subscales / N	28	22	18	35	25	16
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Belligerence	6.61(5.62)	6.05 (5.36)	6.17(4.98)	8.46(6.96)	9.80 (7.17)	8.68 (5.96)
Distress	7.96 (4.32)	6.77 (4.15)	6.78(3.08)	7.34 (4.76)	7.68 (3.65)	8.94 (4.01)
Withdrawal	6.89(2.92)	6.50 (3.66)	6.44 (2.66)	6.69 (3.14)	6.48 (1.78)	6.88 (3.90)
Impulsivity	14.25(7.92)	12.41 (7.69)	13.17 (5.80)	15.15 (8.47)	17.80 (9.20)	16.25 (7.24)
Egocentricity	8.54(5.37)	8.27 (5.98)	7.83 (4.69)	8.71 (5.50)	10.88 (6.88)	10.44 (5.42)
Problem solving	23.54 (6.12)	24.09 (5.47)	26.27 (4.93)	24.03 (6.84)	22.40 (6.16)	23.81 (5.47)
Total BAC	67.79 (20.23)	64.09 (20.81)	66.67 (15.31)	69.94 (23.72)	75.04 (19.84)	75.00 (18.55)

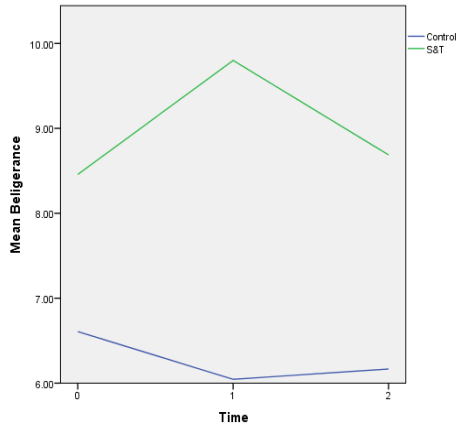
Table 10. Means (SD) prior to Cronbach's alpha correction						
	WL/TAU M (SD)			**S&T** M (SD)		
Subscales / N	28	22	18	34	28	16
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Withdrawal	8.11(3.39)	7.45(4.34)	8.22(2.71)	8.25(3.08)	7.44(2.15)	8.76(3.93)
Problem solving	24.89(5.63)	26.09(5.13)	28.39(3.94)	25.23(6.86)	24.00(6.42)	25.00(5.53)
Total BAC-R	73.96(22.21)	69.04(19.75)	72.33(17.78)	74.69(22.69)	76.30(18.46)	81.82(18.18)

Table 11. Intent-to-treat (multiple imputation) mean scores for BAC-R						
	WL/TAU M (SE)			**S&T** M (SE)		
Subscales / N	28	28	28	35	35	35
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Belligerence	6.61 (1.06)	7.00 (1.41)	7.02 (1.21)	8.45 (1.17)	8.73 (1.44)	7.99 (1.51)
Distress	7.96 (.82)	6.83 (.73)	7.14 (.64)	7.34 (.80)	7.68 (.59)	8.38 (.54)
Withdrawal	6.89 (.55)	6.35 (.66)	6.51 (.49)	6.69 (.53)	6.57 (.40)	6.56 (.54)
Impulsivity	14.25 (1.49)	13.05 (1.38)	13.66 (.96)	14.71 (1.48)	17.17 (1.37)	15.45 (.97)
Egocentricity	8.54 (1.02)	8.45 (1.05)	8.32 (.78)	8.71 (.93)	10.31 (1.03)	9.74 (.73)
Problem solving	23.53 (1.16)	23.87 (.97)	25.61 (.90)	24.03 (1.16)	22.72 (.94)	24.52 (.76)
Total BAC	67.79 (3.82)	65.54 (3.55)	68.26 (2.43)	69.94 (4.01)	73.18 (2.94)	72.64 (2.21)

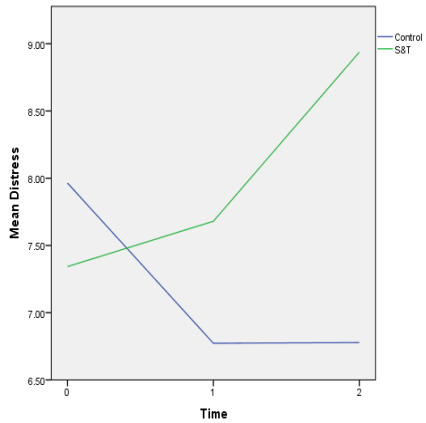
BAC Total



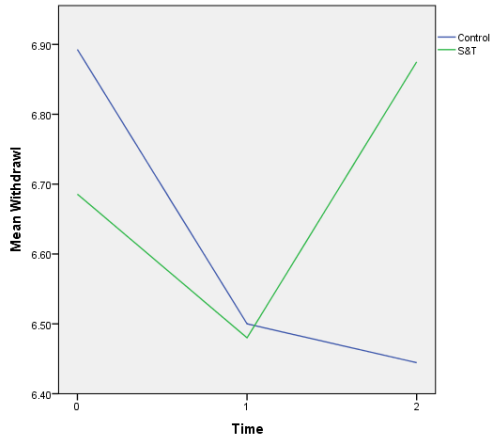
Belligerence



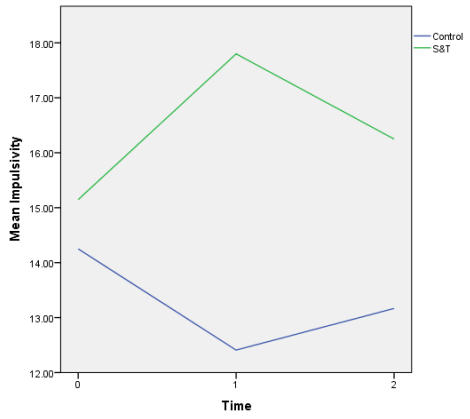
Distress



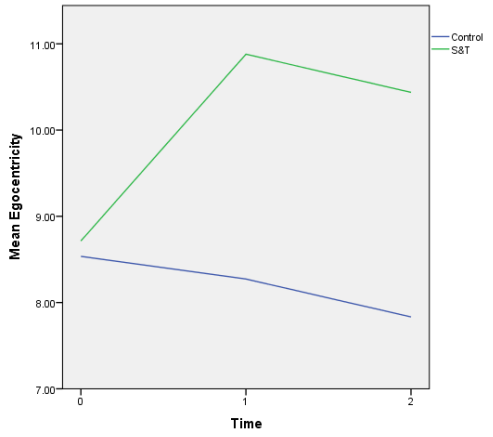
Withdrawal



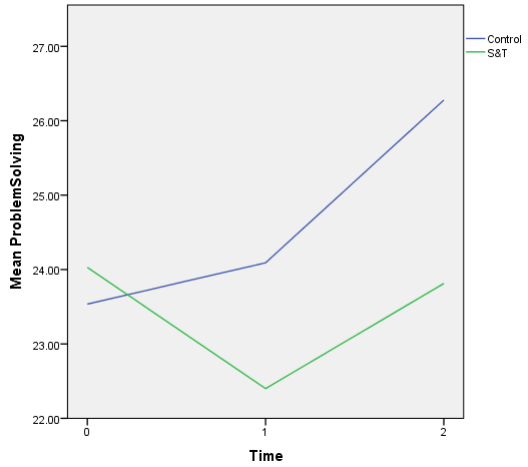
Impulsivity



Egocentricity



Problem Solving



Impulsivity Non-Linear **(T1=0, T2=.80, T3=1)**

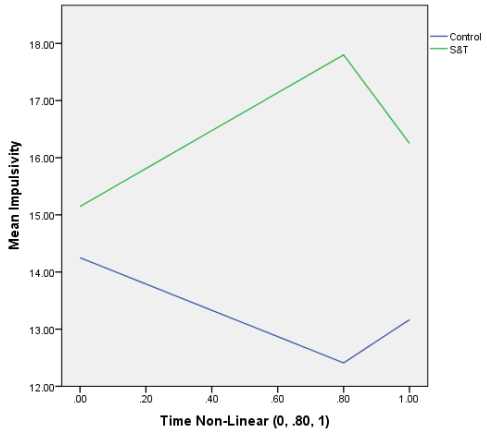


Figure 3: ITT means for BAC total and subscales for outcomes at T1 – T3.

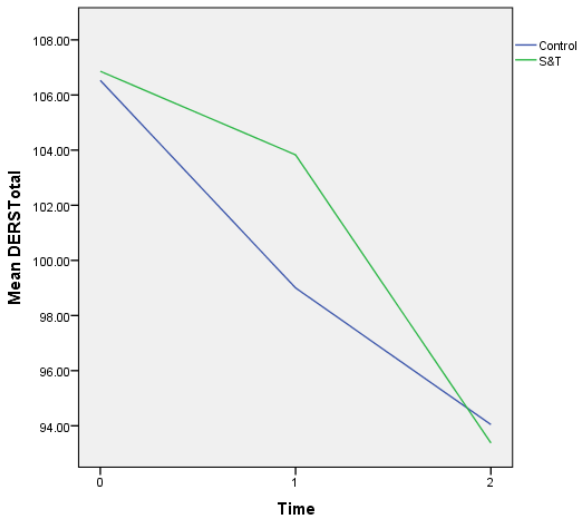
3.5.3. Difficulties in Emotional Regulation Scale (DERS)

Table 12. Raw mean scores for DERS						
	WL/TAU Mean (SD)			**S&T** Mean (SD)		
Subscales / N	34	29	22	43	35	24
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Non Accept	17.56(6.22)	17.72(6.89)	17.91(6.78)	18.72(7.1)	16.97(6.53)	16.00(7.30)
Goals	16.65(5.50)	14.56(5.53)	14.00(5.03)	16.40(4.70)	15.66(5.12)	14.33(5.98)
Impulse	15.56(7.99)	14.55(5.23)	13.45(6.28)	16.67(6.30)	14.88(6.01)	14.13(6.65)
Aware	19.97(6.97)	18.45(5.94)	17.27(5.86)	17.77(5.98)	19.49(8.62)	17.04(6.23)
Strategies	23.41(8.05)	21.17(8.13)	19.73(8.34)	23.40(7.82)	22.66(8.25)	20.33(8.86)
Clarity	13.38(4.21)	12.55(3.61)	11.68(4.56)	13.91(4.61)	14.17(4.86)	11.54(4.46)
Total DERS	106.53(29.27)	99.00(27.24)	94.05(29.49)	106.86(27.67)	103.83(28.45)	93.38(31.23)

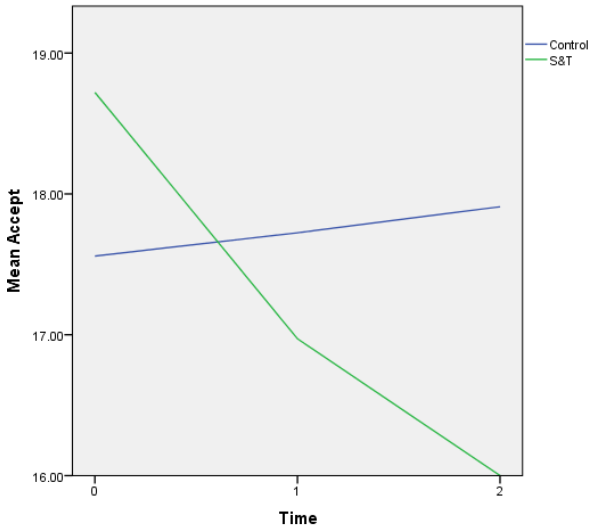
Table 13. Intent-to-treat (multiple imputation) mean scores for DERS						
	WL/TAU Mean (SE)			**S&T** Mean (SE)		
Subscales / N	42	42	42	44	44	44
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Non Accept	17.78(1.08)	18.16 (1.24)	17.46 (1.46)	18.73(1.08)	16.65 (1.10)	16.08 (1.31)
Goals	16.55(.79)	14.60 (.77)	14.05 (.66)	16.43(.71)	15.49 (.75)	14.42 (.72)
Impulse	15.58(1.15)	14.57 (1.03)	13.62 (.83)	16.65(.94)	14.88 (.84)	14.06 (.82)
Aware	19.73(.89)	18.49 (.80)	17.34 (.69)	17.77 (.90)	19.34 (1.19)	17.06 (.80)
Strategies	23.24 (1.14)	21.36 (1.08)	19.96 (1.00)	23.37 (1.17)	22.59 (1.27)	20.21 (1.03)
Clarity	13.44 (.63)	12.78 (.53)	11.69 (.60|)	14.00 (.69)	13.97 (.68)	11.44 (.57)

Total DERS	106.33 (4.11)	99.95 (3.55)	94.14 (3.31)	106.33 (4.11)	99.95 (3.55)	93.27 (3.50)

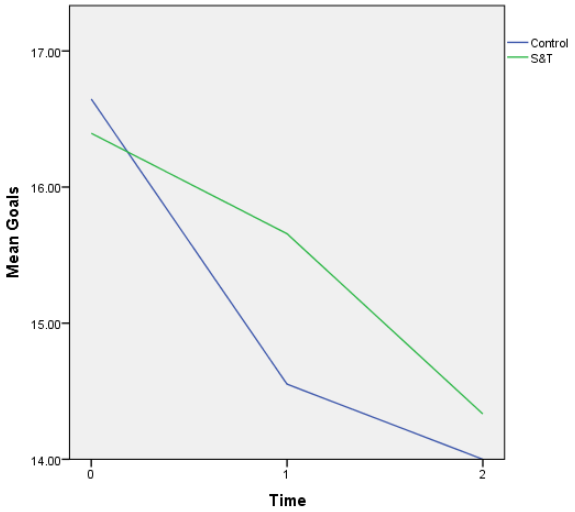
DERS Total



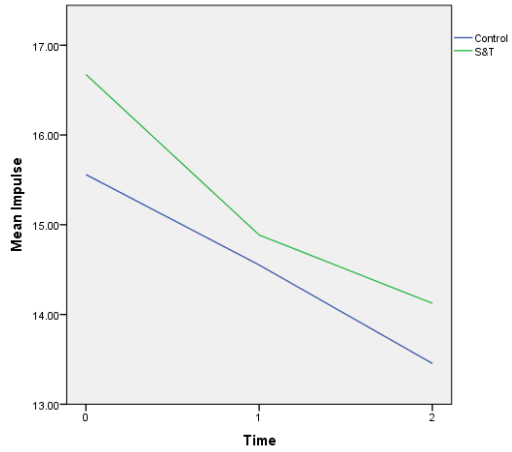
Non Accept



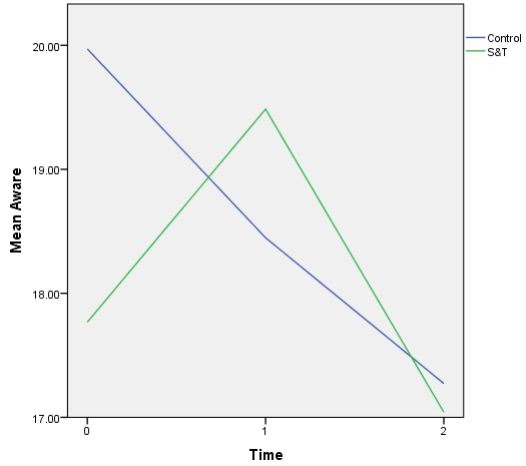
Goals



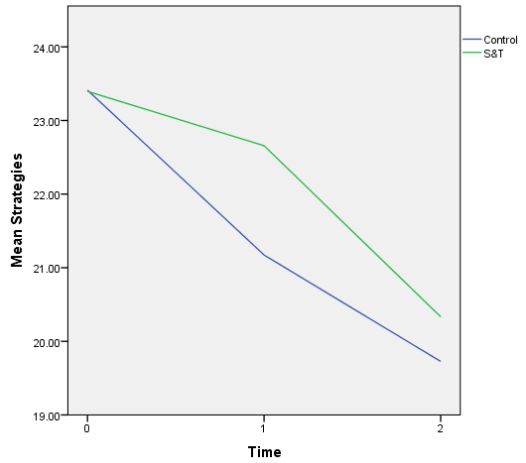
Impulsive



Aware



Strategies



Clarity

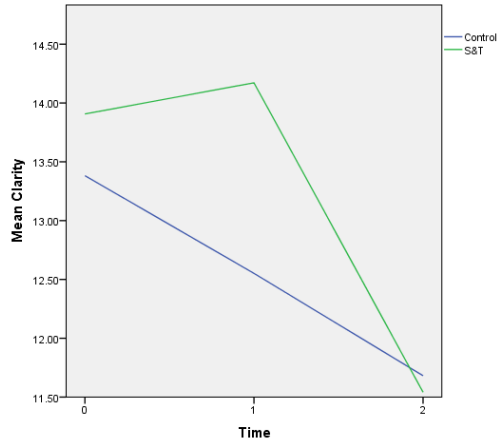


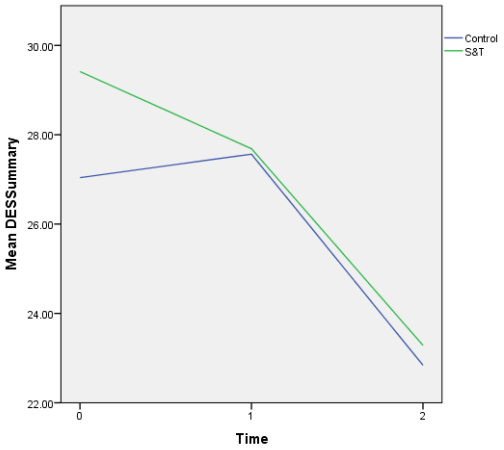
Figure 4: ITT means for DERS total and subscales for outcomes at T1 – T3.

3.5.4. Dissociative Experiences Scale (DES II)

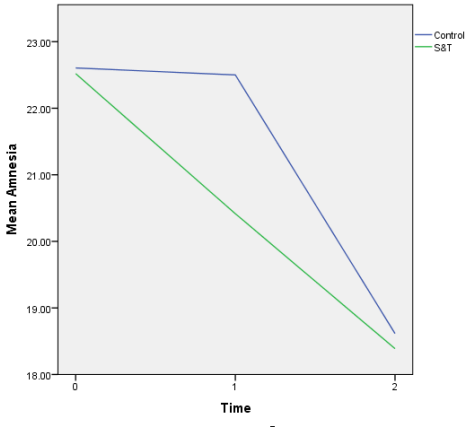
Table 14. Raw mean scores for DES II						
	WL/TAU Mean (SD)			**S&T** Mean (SD)		
Subscales / N	34	28	21	44	35	22
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Amnesia	22.61(19.58)	22.50(19.89)	18.62(19.36)	22.52(18.74)	20.00(15.12)	18.39(15.03)
Absorption	30.54(20.86)	31.31(19.77)	25.67(20.11)	32.82(20.12)	31.21(19.79)	25.24(14.78)
Depersonalisation	17.39(18.06)	18.11(18.50)	15.53(19.60)	20.80(19.82)	19.72(18.13)	17.46(14.99)
Total DES	27.04(20.43)	27.56(21.59)	22.84(21.24)	29.41(20.21)	27.69(19.41)	23.28(16.21)

Table 15. Intent-to-treat (multiple imputation) mean scores for DES II						
	WL/TAU Mean (SE)			**S&T** Mean (SE)		
Subscales / N	42	42	42	44	44	44
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Amnesia	22.77(3.23)	24.05(2.84)	24.76(3.54)	22.52(2.83)	21.49(2.56)	21.66(2.86)
Absorption	31.57(3.91)	33.58(3.63)	30.02(4.07)	32.82(3.03)	32.43(3.24)	29.12(3.02)
Depersonalisation	18.66(3.04)	20.27(3.13)	20.85(3.56)	20.80(2.99)	21.23(2.90)	22.03(2.98)
Total DES	28.27(3.61)	29.68(3.32)	28.56(3.82)	29.41(3.05)	29.24(3.15)	27.63(3.13)

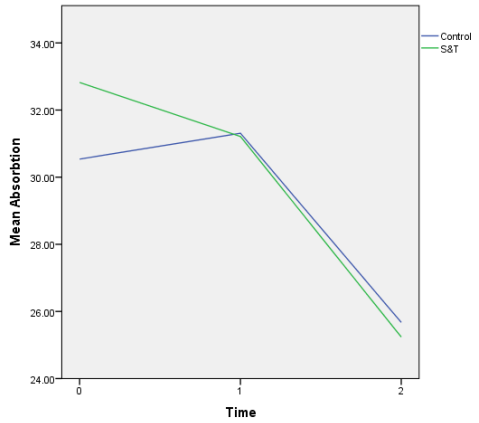
DES Total



Amnesia



Absorption



Depersonalisation

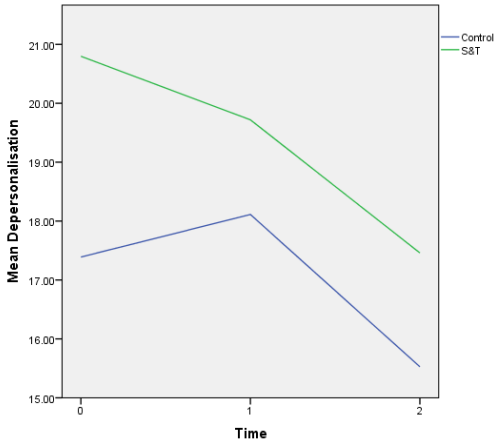


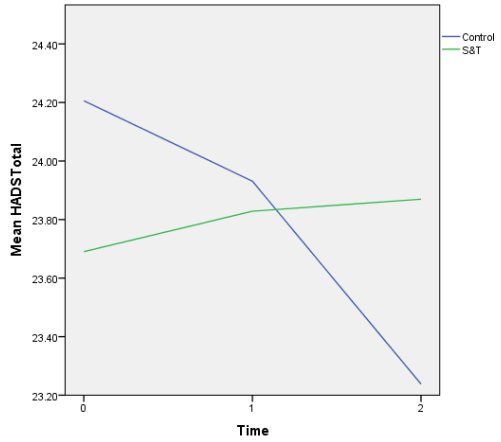
Figure 5: ITT means for DES II total and subscales for outcomes at T1 – T3.

3.5.5. Hospital Anxiety Depression Scale (HADS)

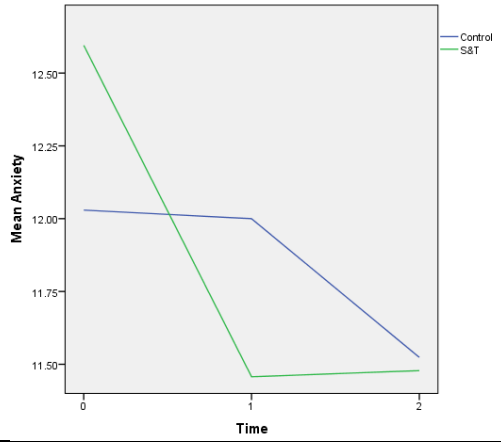
Table 16. Raw mean scores for HADS						
	WL/TAU Mean (SD)			**S&T** Mean (SD)		
Subscales / N	34	29	21	42	35	23
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Anxiety	12.03(3.73)	12.00(3.47)	11.52(4.21)	12.60(2.68)	11.46(3.32)	11.48(2.69)
Depression	12.18(2.50)	11.93(3.08)	11.71(3.42)	11.10(3.31)	12.37(2.70)	12.39(3.13)
Total HADS	24.21(2.42)	23.93(3.42)	23.24(3.35)	23.69(2.88)	23.83(3.70)	23.87(3.40)

Table 17. Intent-to-treat (multiple imputation) mean scores for HADS						
	WL/TAU M (SE)			**S&T** M (SE)		
Subscales / N	42	42	42	44	44	44
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Anxiety	11.97(.61)	11.95(.64)	11.50(.63)	12.61(.42)	11.58(.51)	11.64(.58)
Depression	12.12(.41)	11.92(.47)	11.87(.47)	11.11(3.31)	12.30(.40)	12.15(.42)
Total HADS	24.09(.37)	23.87(.52)	23.38(.45)	23.72(2.88)	23.88(.52)	23.78(.55)

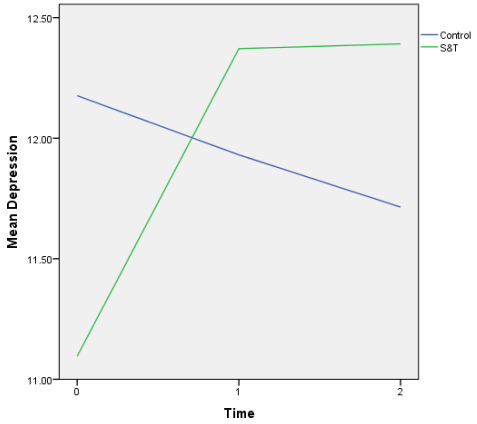
HADS Total



Anxiety



Depression



Anxiety Non-Linear (T1=0, T2=.80, T3=1)

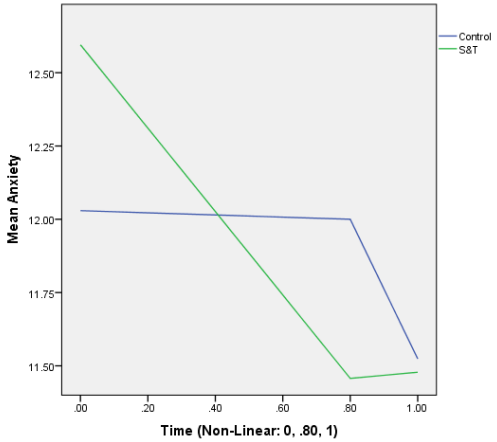


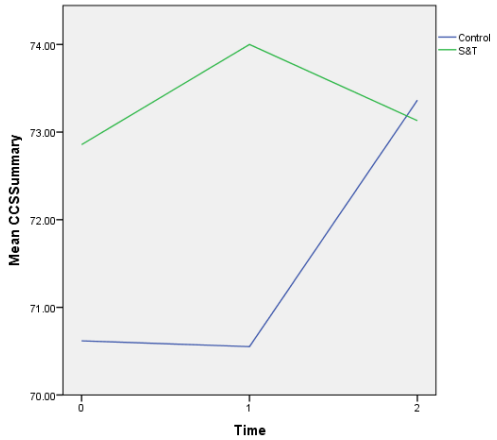
Figure 6: ITT means for HADS total and subscales for outcomes at T1 – T3

3.5.6. Criminal Cognitions Scale (CCS)

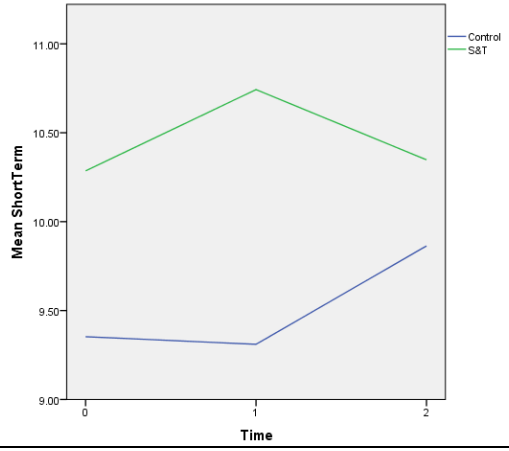
Table 18. Raw mean scores for CCS						
	WL/TAU Mean (SD)			**S&T** Mean (SD)		
Subscales / N	34	29	22	42	35	23
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Short Term	9.35(2.41)	9.31(2.89)	9.86(2.46)	10.29(2.74)	10.74(2.00)	10.35(2.50)
Entitlement	8.74(2.43)	8.90(3.05)	9.14(2.57)	9.19(2.97)	9.40(2.13)	9.00(2.26)
Responsibility	11.41(2.95)	11.24(3.00)	11.23(3.39)	11.81(2.86)	11.80(2.77)	11.57(2.19)
Negative	10.71(2.69)	10.86(2.95)	11.32(3.12)	11.45(2.93)	11.29(1.98)	11.17(2.37)
Insensitivity	8.79(2.91)	8.41(3.46)	8.64(3.02)	9.45(2.89)	8.77(2.77)	9.13(2.56)
Guilt	21.91(5.28)	21.83(6.03)	23.18(5.51)	20.74(4.44)	22.00(3.79)	21.91(4.07)
Summary CCS	70.61(12.18)	70.56(13.93)	73.36(13.27)	72.86(12.18)	74.00(7.98)	73.13(8.71)

Table 19. Intent-to-treat (multiple imputation) mean scores for CCS						
	WL/TAU M (SE)			**S&T** M (SE)		
Subscales / N	42	42	42	44	44	44
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Short Term	9.56(.58)	9.66 (.83)	9.79 (.64)	10.40(.45)	10.79 (.54)	10.20 (.82)
Entitlement	8.88 (.42)	8.92 (.43)	9.23 (.35)	9.19 (.45)	9.40 (.36)	9.04 (.43)
Responsibility	11.47 (.46)	11.33 (.47)	11.29 (.50)	11.81 (.44)	11.66 (.41)	11.33 (.33)
Negative	10.61 (.45)	10.93 (.40)	11.29 (.49)	11.51 (.45)	11.19 (.31)	11.15 (.34)
Insensitivity	8.88 (.45)	8.47 (.51)	8.52 (.45)	9.38 (.44)	8.74 (.36)	9.00 (.40)
Guilt	21.60 (.66)	22.06 (.75)	23.18 (.70)	20.69 (.66)	22.02 (.55)	22.30.(.61)
Summary CCS	70.71 (1.74)	71.19 (1.80)	73.13 (1.52)	72.89 (1.80)	73.81 (1.15)	73.02 (1.12)

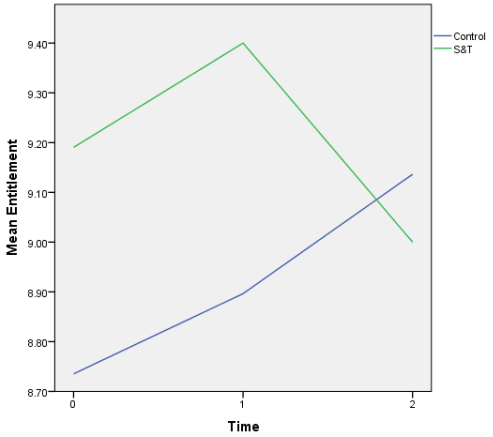
CCS Summary



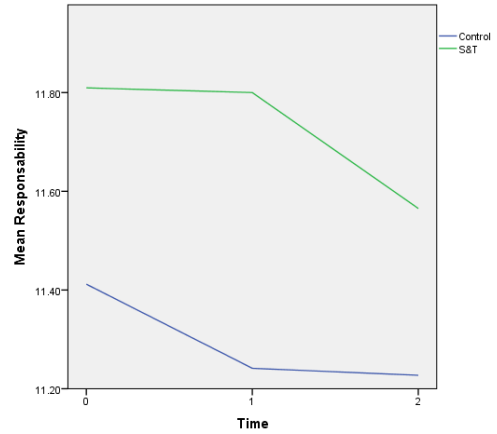
Short Term



Entitlement

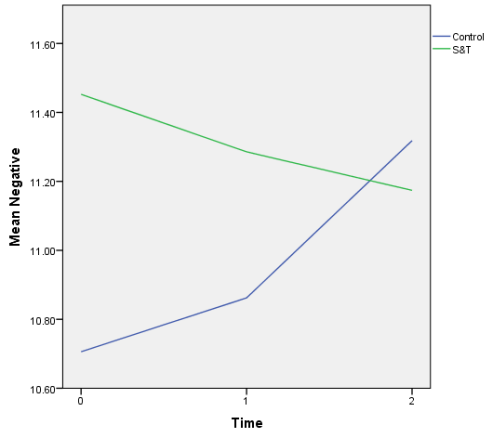


Responsibility

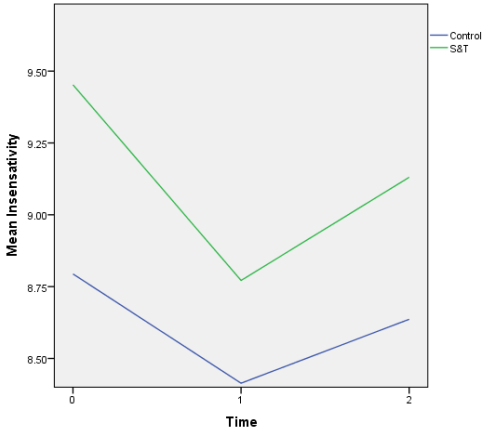


Guilt

Negative



Insensitivity



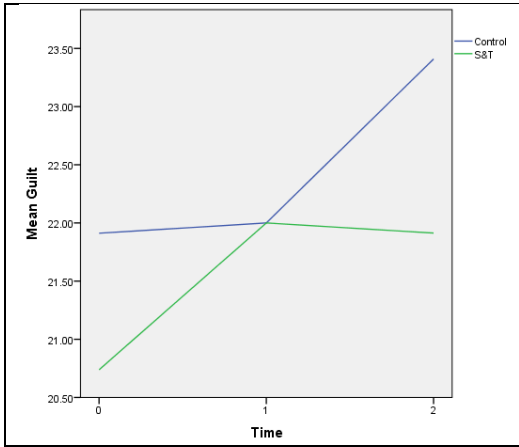


Figure 7: ITT means for CCS total and subscales for outcomes at T1 – T3

3.6. Differences between Adequate Dose and Non-Adequate Dose Survive & Thrive Participants

3.6.1. PTSD Checklist –Civilian Version (PCL-C)

Table 20. Raw mean scores for PCL-C outcomes for AD and non-AD participants in the S&T arm

	Non-adequate dose M (SD)			**Adequate dose** M (SD)		
Subscales / N	11	4	3	32	31	20
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Intrusion	14.82 (6.63)	18.25 (5.85)	19.33 (3.21)	14.81 (5.87)	12.94 (5.54)	11.40 (3.41)
Avoidance	20.27 (7.86)	25.25 (5.50)	23.67 (7.64)	20.06 (6.61)	18.06 (7.12)	17.00 (6.19)
Arousal	16.18 (6.49)	18.00 (5.77)	18.67 (4.51)	15.56 (5.13)	14.06 (5.34)	13.70 (4.85)
Total PCL-C	51.27 (19.32)	61.50 (16.78)	61.67 (15.18)	50.44 (15.84)	45.06 (15.91)	42.10 (12.79)

Table 21. ITT (MI) mean scores for PCL-C outcomes for AD and non-AD participants in the S&T arm

	Non-adequate dose M (SE)			**Adequate dose** M (SE)		
Subscales / N	11	11	11	33	33	33
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Intrusion	14.82 (2.00)	16.77 (2.93)	15.20 (1.81)	14.82 (1.01)	13.13 (.96)	11.69 (.87)
Avoidance	20.27 (2.37)	20.48 (1.69)	19.83 (1.69)	20.04 (1.14)	18.05 (1.21)	17.84 (.92)
Arousal	16.18 (1.96)	15.78 (1.59)	15.41 (1.04)	15.51 (.89)	14.13 (.91)	14.01 (.79)
Total PCL-C	51.27 (5.83)	53.04 (3.83)	50.43 (3.38)	50.38 (2.72)	45.31 (2.70)	43.54 (1.83)

Table 22. Linear mixed model analysis of AD and non-AD participants x time interaction for PCL-C outcomes									
	Post (T2)			**FU1 (T3)**			**Post and FU1 (T2 and T3)**		
Subscales	β	(95% CI)	p	β	(95% CI)	p	β	(95% CI)	p
Intrusion	.15	(-4.06 to 4.36)	.942	-4.74	(-8.81 to -.66)	.024	2.13	(0.14 to 4.13)	.037
Avoidance	4.54	(-.92 to 1.00)	.100	-2.24	(-7.35 to 2.88)	.378	1.55	(-1.02 to 4.13)	.229
Arousal	2.24	(2.15 to 6.63)	.303	-.54	(-5.27 to 4.20)	.820	-.36	(-2.04 to .2.77)	.762
PCL Total	7.62	(-3.25 to 18.48	.162	-6.56	(17.61 to 4.50)	.236	-3.15	(-2.46 to 8.77)	.260

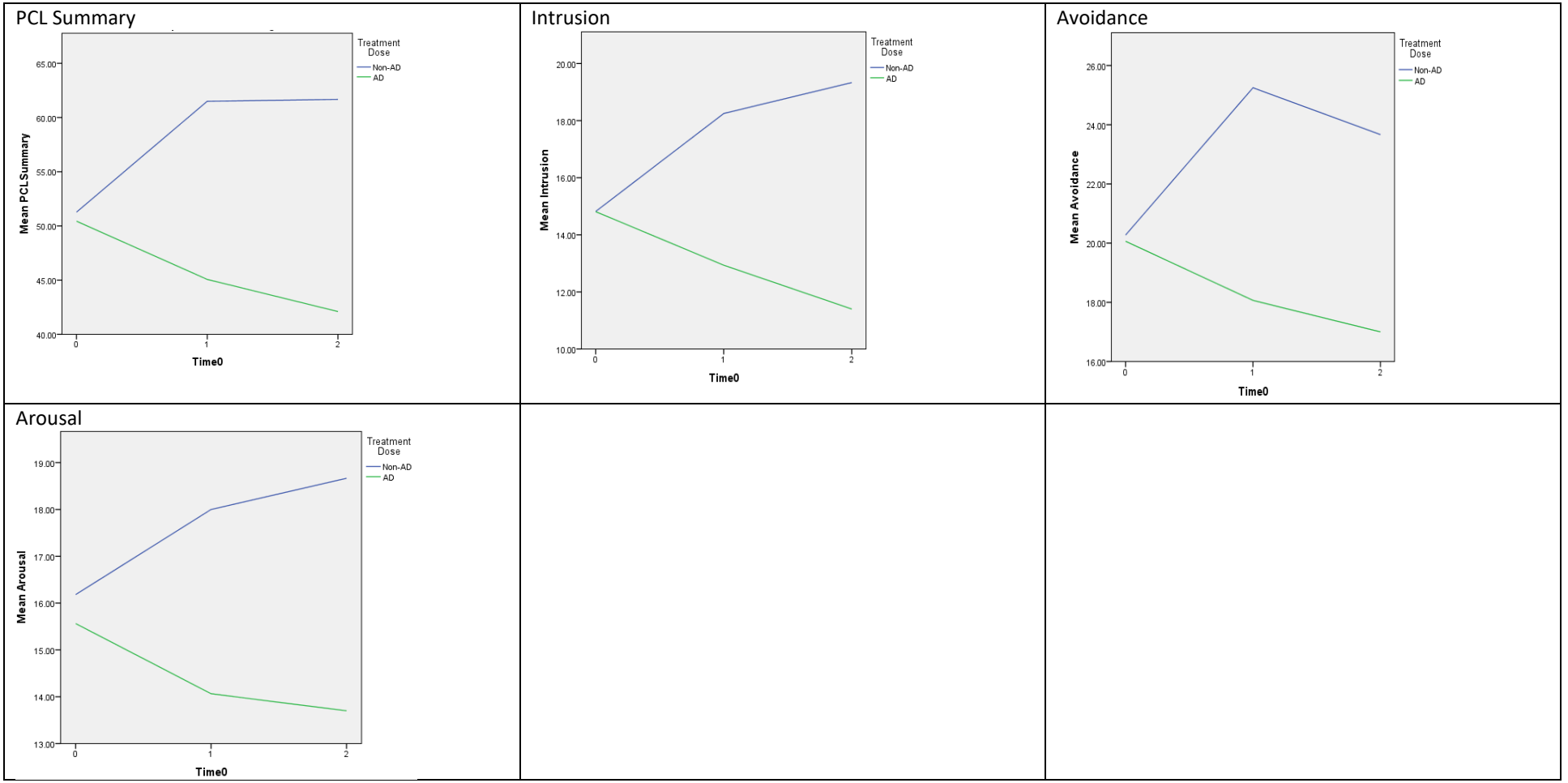


Figure 8: ITT means for PCL-C total and subscale outcomes of AD and non-AD participants within the S&T arm at T1 – T3

3.6.2. Behavioural Assessment Checklist-Revised

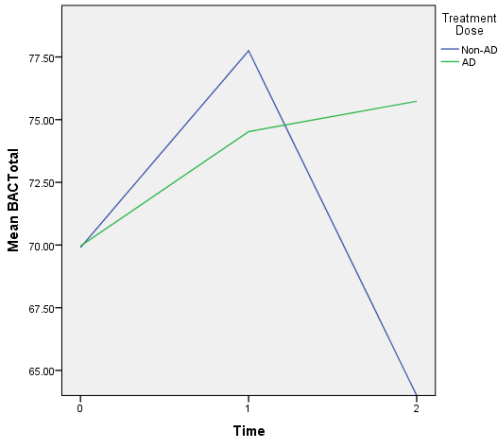
Table 23. Raw mean scores for BAC-R outcomes for AD and non-AD participants in the S&T arm						
	Non-adequate dose M (SD)			**Adequate dose** M (SD)		
Subscales / N	10	4	1	25	21	15
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Belligerence	8.60 (7.47)	11.75 (5.50)	7.00 (-)	8.40 (6.91)	9.43 (7.50)	8.80 (6.16)
Distress	8.30 (5.66)	8.25 (5.50)	6.00 (-)	6.96 (4.42)	7.57 (3.36)	9.13 (4.07)
Withdrawal	6.50 (3.63)	6.00 (2.00)	3.00 (-)	6.76 (3.00)	6.57 (1.78)	7.13 (3.89)
Impulsivity	17.78 (9.60)	18.25 (5.56)	17.00 (-)	14.20 (8.02)	17.71 (9.84)	16.20 (7.49)
Egocentricity	9.20 (5.90)	13.00 (8.28)	9.00 (-)	8.52 (5.45)	10.48 (6.74)	10.53 (5.59)
Problem Solving	21.30 (7.07)	20.50 (6.14)	22.00 (-)	25.12 (6.56)	22.76 (6.25)	23.93 (5.64)
Total BAC-R	69.90 (33.11)	77.75 (19.99)	64.00 (-)	69.96 (19.64)	74.52 (20.26)	75.73 (18.96)

Table 24. ITT (MI) mean scores for BAC-R outcomes for AD and non-AD participants in the S&T arm						
	Non-adequate dose M (SE)			**Adequate dose** M (SE)		
N	10	10	10	25	25	25
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Belligerence	Multiply imputing data with large amounts of missing data not appropriate.			8.40 (1.06)	9.06 (1.69)	8.42 (1.48)
Distress				6.96 (.88)	7.54 (.68)	8.74 (.68)
Withdrawal				6.76 (.601)	6.61 (.43)	6.85 (.70)
Impulsivity				14.20 (1.61)	17.34 (1.83)	15.46 (1.25)
Egocentricity				8.52 (1.09)	10.25 (1.27)	9.93 (.95)
Problem Solving				25.12 (1.31)	22.91 (1.17)	24.31 (.94)

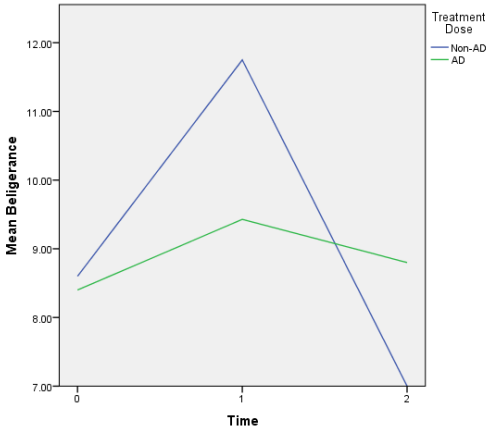
Total BAC-R				69.96 (3.93)	73.72 (3.76)	73.71 (3.01)

Table 25. Linear mixed model analysis of AD and non-AD participants x time interaction for BAC-R outcomes									
	Post (T2)			**FU1 (T3)**			**Post and FU1 (T2 and T3)**		
Subscales	β	(95% CI)	p	β	(95% CI)	p	β	(95% CI)	p
Belligerence	-2.11	(-10.62 to 14.85)	.733	.77	(-11.50 to 13.04)	.896	1.13	(-2.94 to 5.20)	.574
Distress	-4.99	(-3.69 to 13.66)	.245	6.27	(-2.46 to 15.01)	.150	-2.44	(-5.92 to 1.04)	.166
Withdrawal	-2.99	(-5.10 to 11.09)	.446	-3.52	(-4.71 to 11.76)	.381	-.99	(-3.42 to 1.44)	.417
Impulsivity	-4.60	(-9.50 to 18.71)	.506	-9.37	(-3.86 to22.61)	.153	-3.94	(-9.36 to 1.48)	.149
Egocentricity	-6.09	(-3.86 to 16.05)	.218	5.12	(-4.07 to 14.31)	.256	-.77	(-4.60 to 3.06)	.686
Problem Solving	3.10	(-15.07 to 8.86)	.595	4.35	(-16.03 to 7.34)	.442	1.06	(-3.62 to 5.74)	.650
Total BAC-R	-17.01	(-18.76 to 52.79)	.335	23.54	(-10.82 to 57.91)	.165	-8.51	(-21.38 to 4.37)	.186

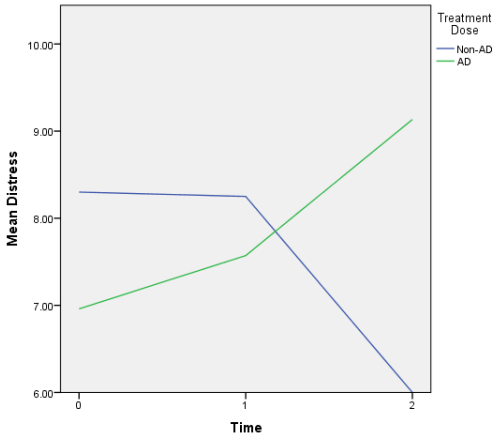
BAC-R Total



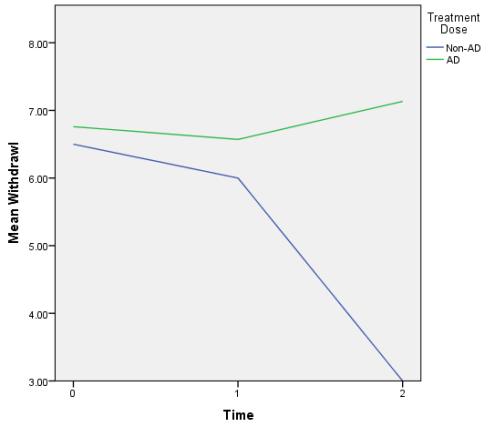
Belligerence



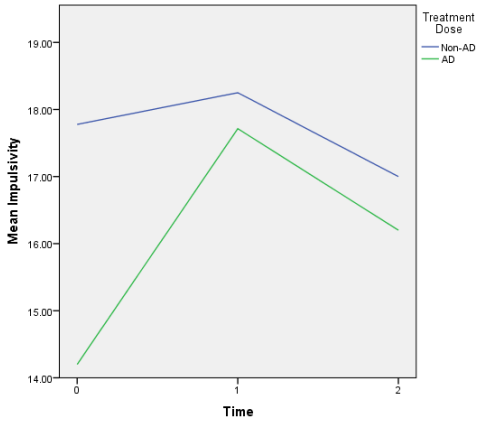
Distress



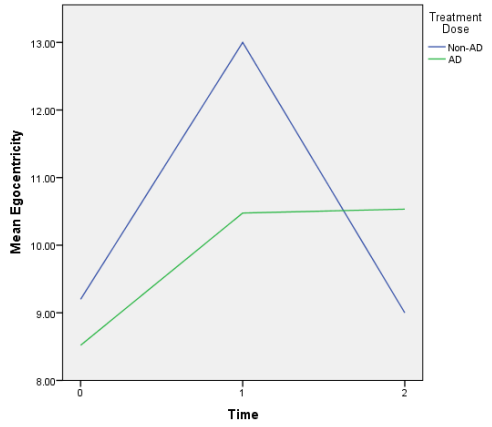
Withdrawal



Impulsivity



Egocentricity



Problem Solving

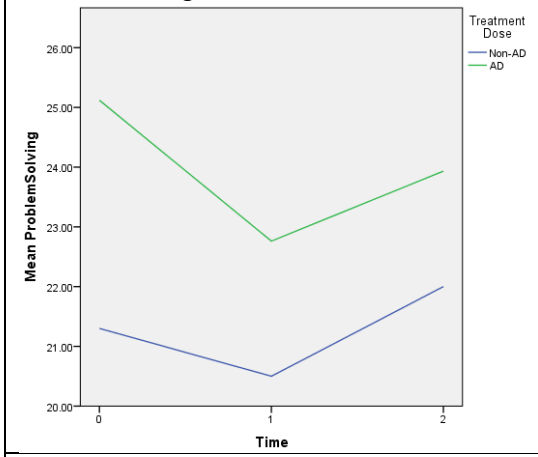


Figure 9: ITT means for BAC-R total and subscale outcomes of AD and non-AD participants within the S&T arm at T1 – T3

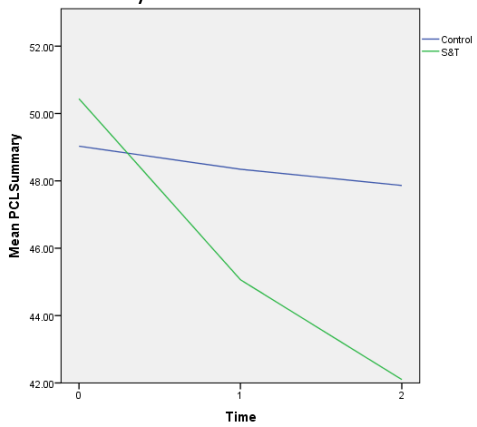
3.7. Differences between adequate dose and WL/TAU participants

3.7.1. PTSD Checklist: Civilian Version (PCL-C)

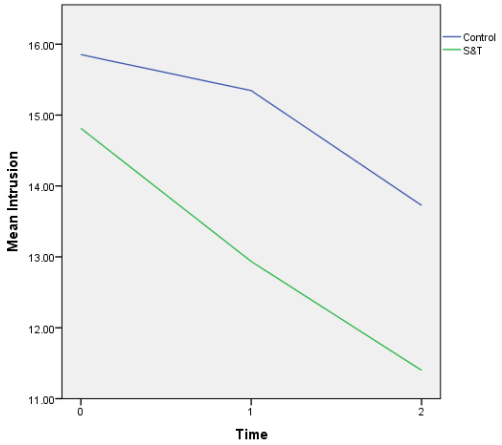
Table 26. Raw and ITT (MI) mean scores for PCL-C outcomes for AD participants in the S&T arm

	Raw data M (SD)			**ITT multiple imputation** M (SE)		
Subscales / N	32	31	20	33	33	33
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Intrusion	14.81 (5.87)	12.94 (5.54)	11.40 (3.4)	14.82 (1.01)	13.13 (.96)	11.69 (.87)
Avoidance	20.06 (6.61)	18.06 (7.12)	17.00 (6.19)	20.04 (1.14)	18.05 (1.21)	17.84 (.92)
Arousal	15.56 (5.13)	14.06 (5.34)	13.70 (4.85)	15.51 (.88)	14.13 (.91)	14.01 (.79)
Total PCL-C	50.44 (15.84)	45.06 (15.91)	42.10 (12.79)	50.38 (2.72)	45.31 (2.70)	43.54 (1.83)

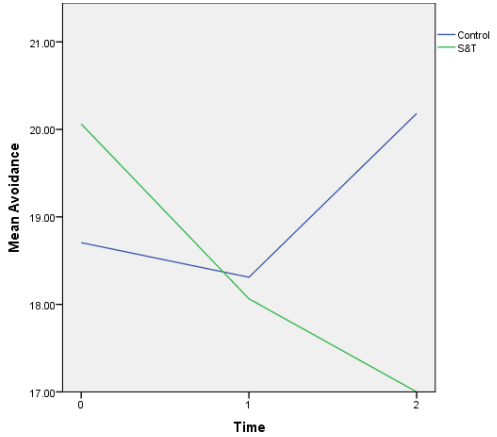
PCL Summary



Intrusion



Avoidance



Arousal

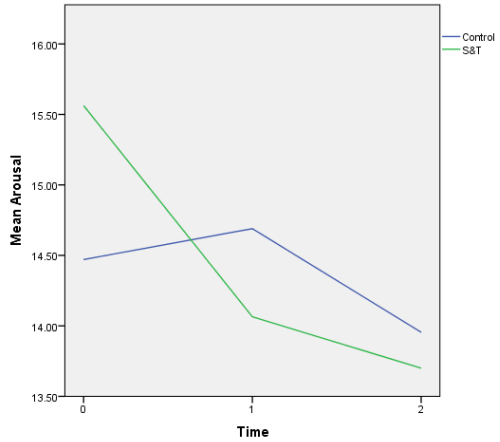
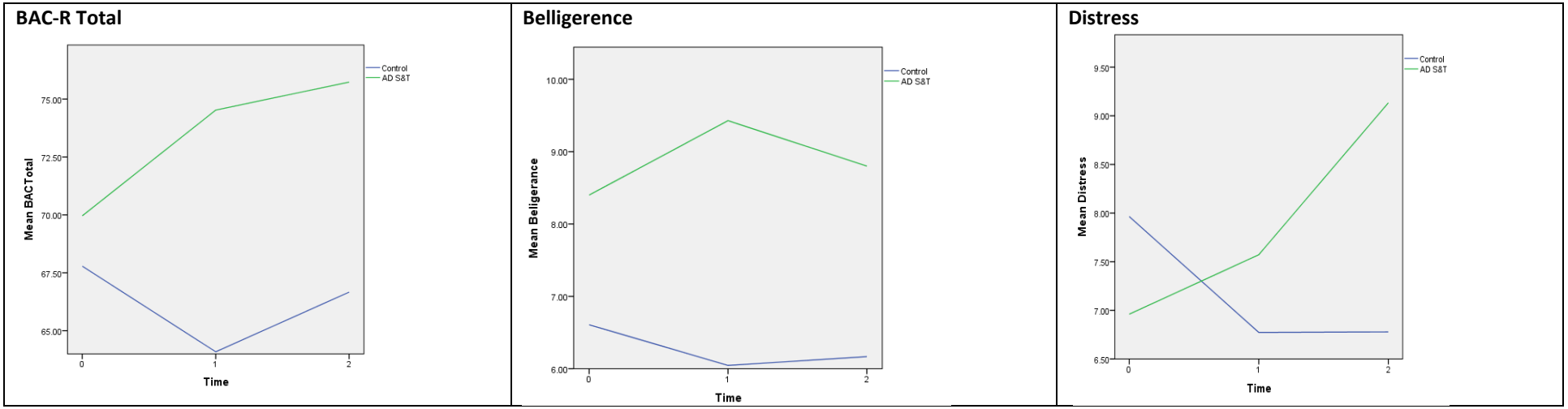


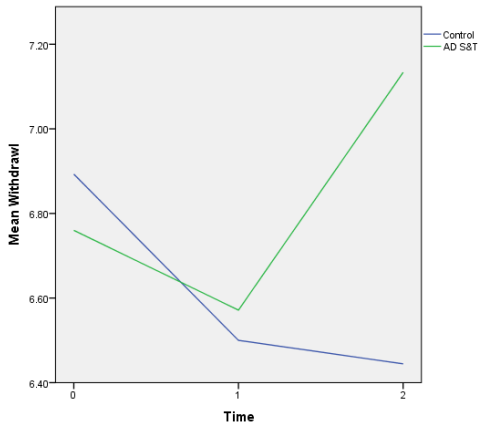
Figure 9: ITT means for PCL-C total and subscales of WL/TAU and AD participants within the S&T arm at T1 – T3

3.7.2. Behavioural Assessment Checklist - Revised

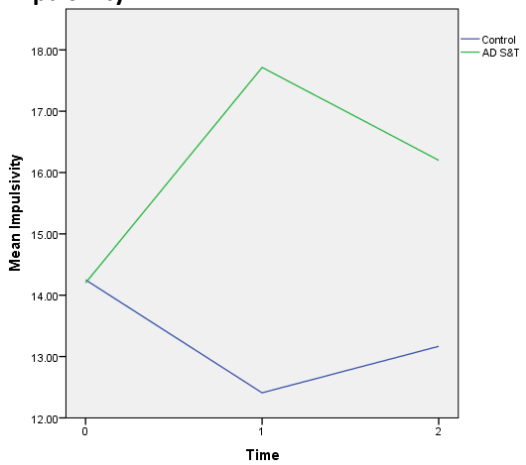
See Tables 23and 24 for AD S&T participant data



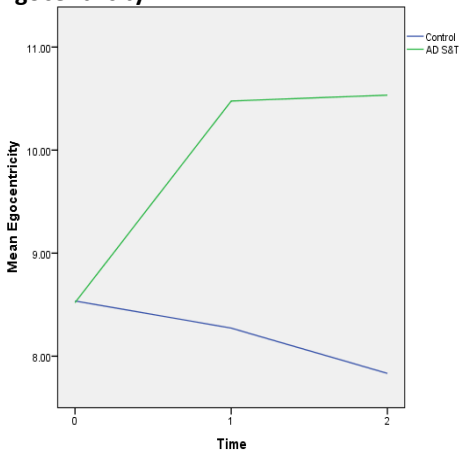
Withdrawal



Impulsivity



Egocentricity



Problem Solving

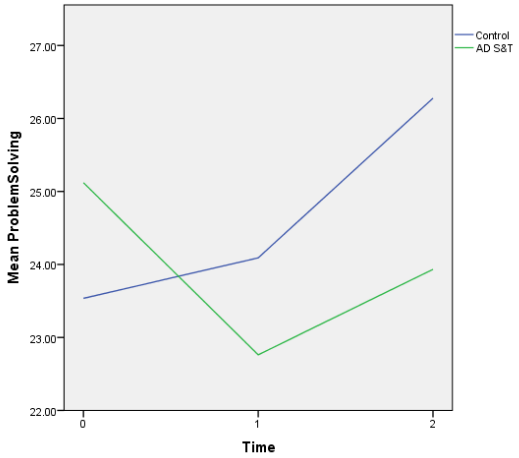
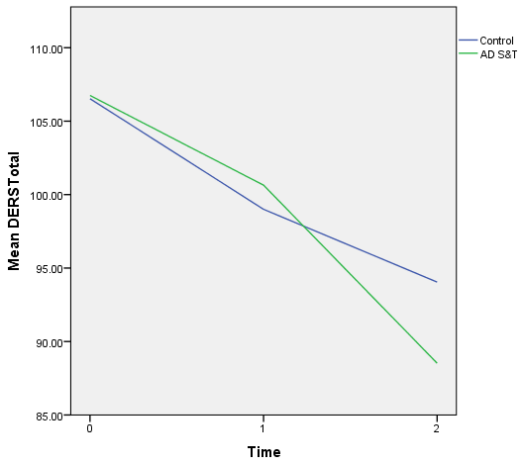


Figure 10: ITT means for BAC-R total and subscales of WL/TAU and AD participants within the S&T arm at T1 – T3

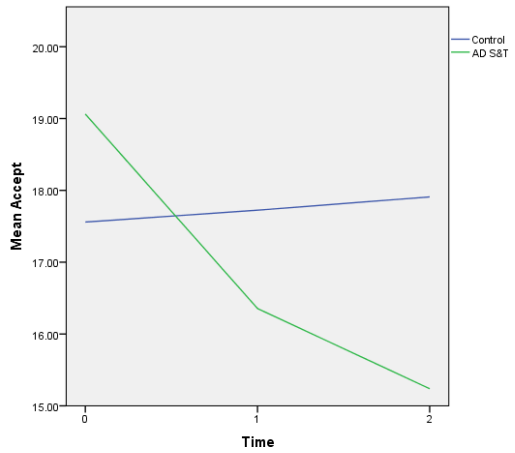
3.7.3. Difficulties in Emotional Regulation Scale (DERS)

Table 27. CCA and ITT (MI) mean scores for DERS outcomes for AD participants in the S&T arm						
	Raw data M (SD)			**Multiple imputation** M (SE)		
Subscales / N	32	31	21	33	33	33
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Non-Accept	19.06 (6.96)	16.35 (6.29)	15.24 (7.19)	19.06 (1.12)	16.23 (1.12)	15.46 (1.39)
Goals	16.38 (4.92)	15.29 (5.13)	13.76 (5.89)	16.43 (.86)	15.29 (.87)	14.06 (.87)
Impulse	16.50 (6.54)	14.39 (5.83)	13.38 (6.29)	16.47 (1.13)	14.44 (1.01)	13.57 (.92)
Aware	17.31 (5.86)	19.45 (9.13)	16.24 (6.06)	17.37 (1.01)	19.36 (1.55)	16.47 (.92)
Strategies	23.53 (8.09)	21.74 (7.90)	19.33 (8.28)	23.49 (1.39)	21.76 (1.34)	19.58 (1.22)
Clarity	13.97 (4.82)	13.42 (4.42)	10.57 (3.84)	14.03 (.83)	13.37 (.75)	10.88 (.62)
DERS Summary	106.75 (29.12)	100.64 (27.31)	88.52 (28.67	106.83 (4.50)	100.46 (4.61)	90.04 (4.02)

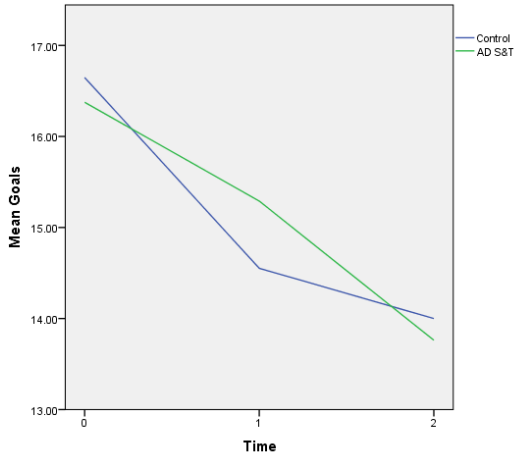
DERS Total



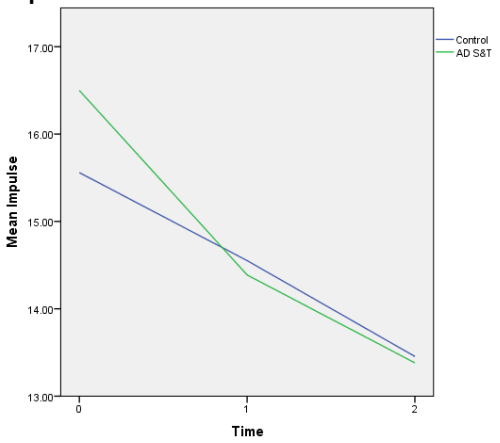
Accept



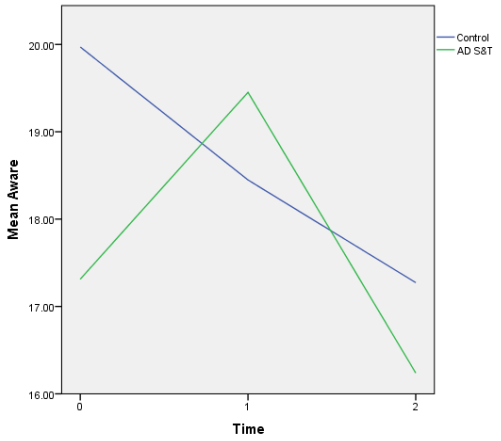
Goals



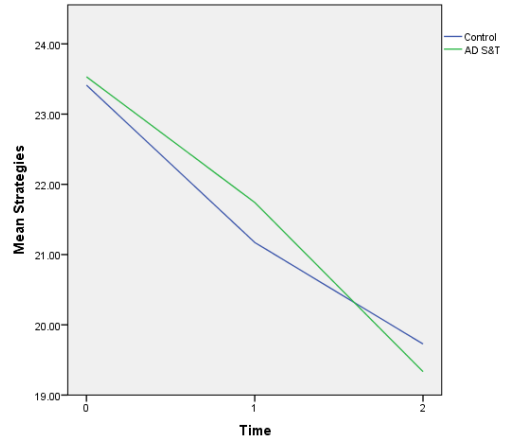
Impulse



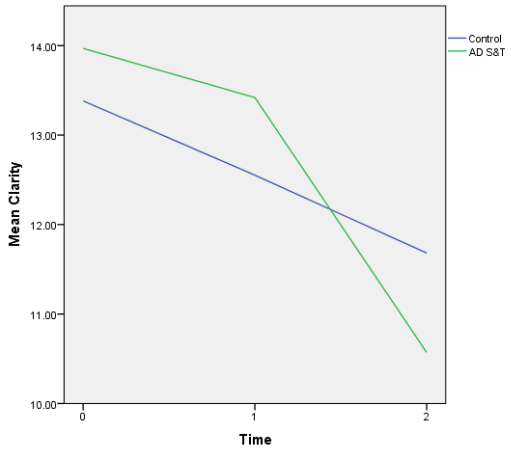
Aware



Strategies



Clarity



Impulse Non-Linear

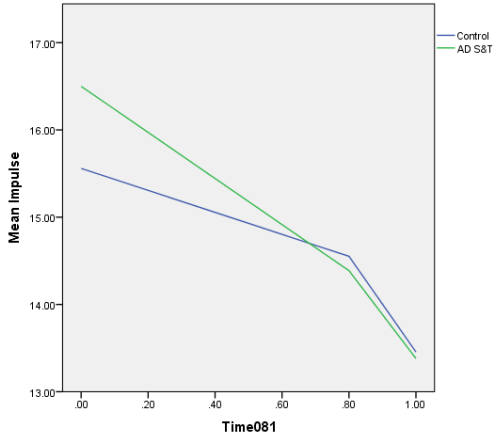
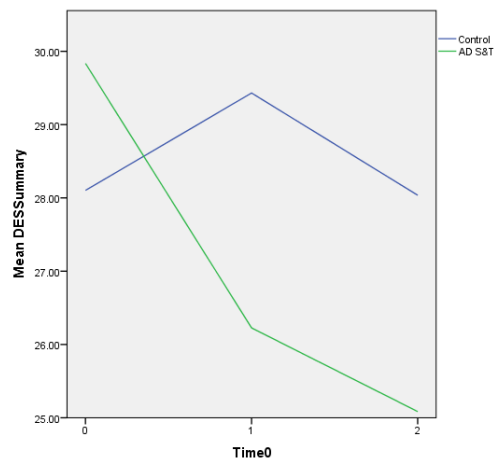
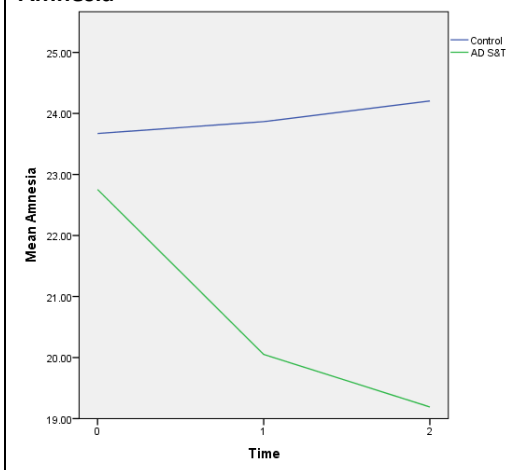


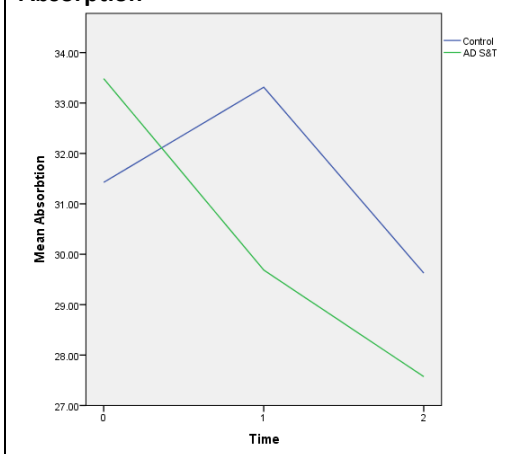
Figure 11: ITT means for DERS total and subscales of WL/TAU and AD participants within the S&T arm at T1 – T3

3.7.4. Dissociation Experiences Scale (DES II)

Table 28. Raw and ITT (MI) mean scores for DES II outcomes for AD participants in the S&T arm						
	Complete case analysis M (SD)			**ITT multiple imputation** M (SE)		
Subscales / N	33	31	19	33	33	33
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Amnesia	22.75 (19.27)	19.79 (15.84)	18.62 (19.11)	22.75 (3.36)	20.10 (2.83)	19.58 (3.13)
Absorption	33.48 (20.69)	29.79 (19.87)	23.08 (14.40)	33.48 (3.60)	29.67 (3.45)	28.09 (3.36)
Depersonalisation	22.26 (20.94)	17.94 (17.09)	14.41 (12.43)	22.26 (3.65)	18.61 (3.00)	20.09 (3.20)
Total DES II	29.84 (20.94)	26.01 (19.11)	20.29 (14.36)	29.84 (3.64)	26.27 (3.33)	25.64 (3.39)







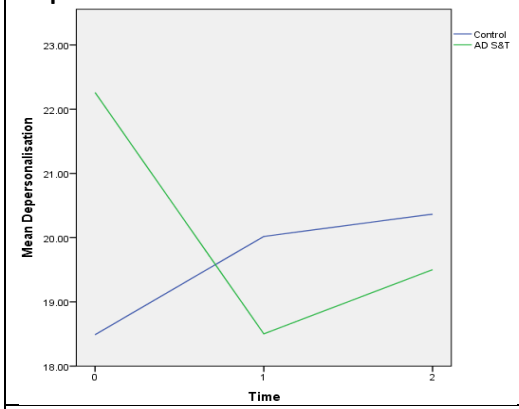


Figure 12: ITT means for DES total and subscales of WL/TAU and AD participants within the S&T arm at T1 – T3

3.7.5. Hospital Anxiety Depression Scale (HADS)

Table 29. Raw and ITT (MI) mean scores for HADS outcomes for AD participants in the S&T arm						
	Raw data M (SD)			**ITT multiple imputation** M (SE)		
Subscales / N	31	31	20	33	33	33
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Anxiety	12.52 (2.62)	11.23 (3.25)	10.95 (2.33)	12.53 (.48)	11.32 (.58)	11.38 (.56)
Depression	11.06 (3.34)	12.52 (2.76)	12.85 (3.00)	11.07 (.58)	12.47 (.47)	12.41 (.48)
Total HADS	23.58 (2.81)	23.74 (3.71)	23.80 (2.65)	23.61 (.48)	23.79 (.64)	23.79 (.63)

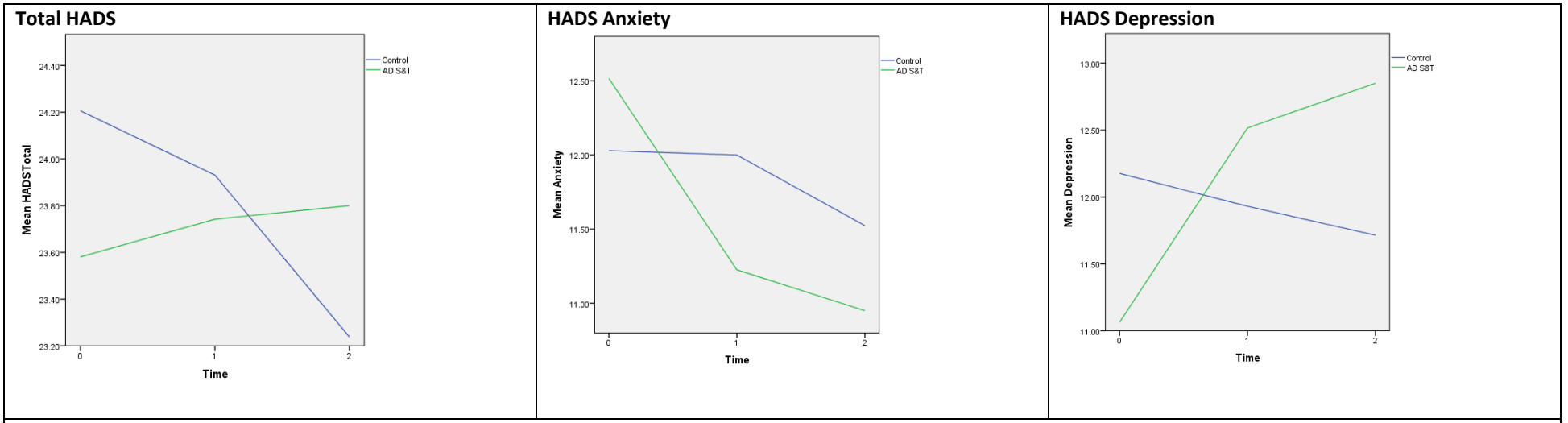
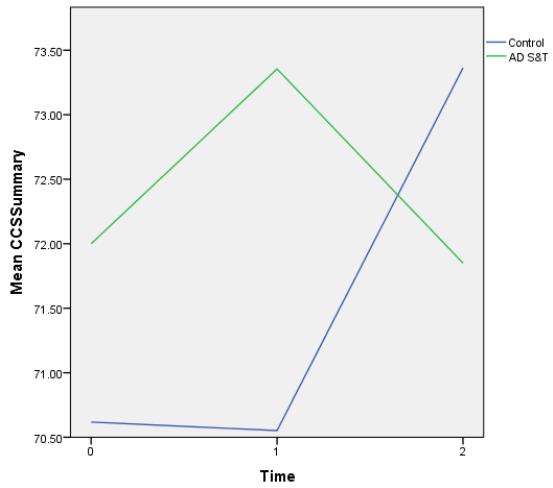


Figure 13: ITT means for HADS total and subscales of WL/TAU and AD participants within the S&T arm at T1 – T3

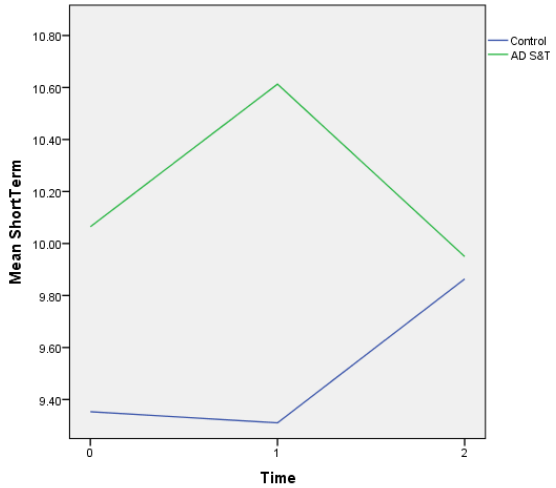
3.7.6. Criminal Cognitions Scale (CCS)

Table 30. Raw and ITT (MI) mean scores for CCS outcomes for AD participants in the S&T arm						
	Raw data M (SD)			**ITT multiple imputation** M (SE)		
Subscales / N	31	31	20	33	33	33
	Pre (T1)	Post (T2)	FU1 (T3)	Pre (T1)	Post (T2)	FU1 (T3)
Short Term	10.06 (2.67)	10.61 (1.98)	9.95 (2.28)	10.23 (.53)	10.70 (.43)	10.00 (.98)
Entitlement	9.12 (2.66)	9.45 (2.00)	8.85 (2.37)	9.13 (.48)	9.46 (.35)	8.93 (.55)
Responsibility	11.74 (2.67)	11.68 (2.77)	11.50 (1.96)	11.75 (.48)	11.62 (.48)	11.33 (.36)
Negative	11.29 (2.92)	10.94 (1.71)	10.85 (2.18)	11.37 (.52)	10.95 (.30)	10.88 (.35)
Insensitivity	9.03 (2.61)	8.77 (2.23)	9.00 (2.66)	8.96 (.46)	8.80 (.40)	8.92 (.47)
Guilt	20.84 (4.31)	21.90 (3.83)	21.70 (4.13)	20.76(.74)	21.85 (.66)	22.15 (.64)
Summary CCS	72.00 (12.06)	73.35 (7.41)	71.85 (8.36)	72.10 (2.04)	73.40 (1.30)	72.23 (1.31)

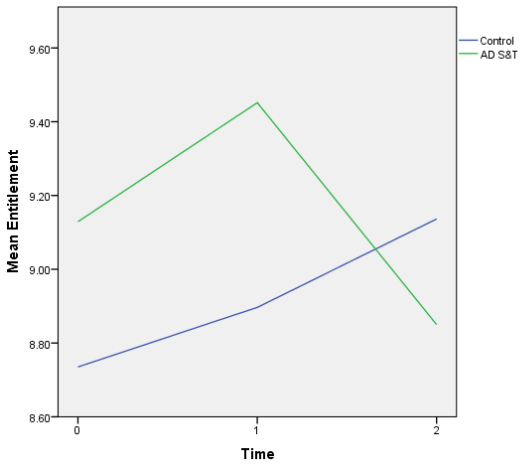
CCS Summary



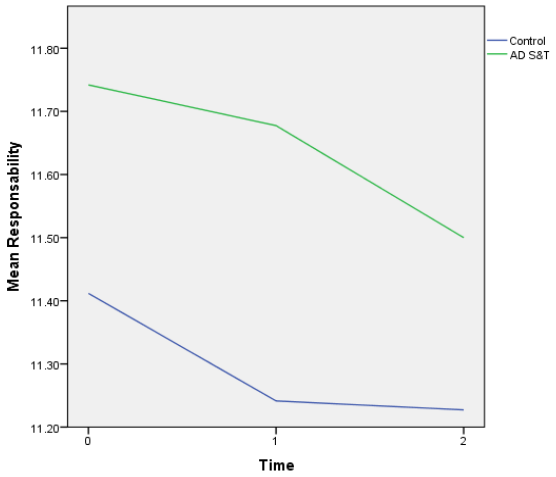
Short Term



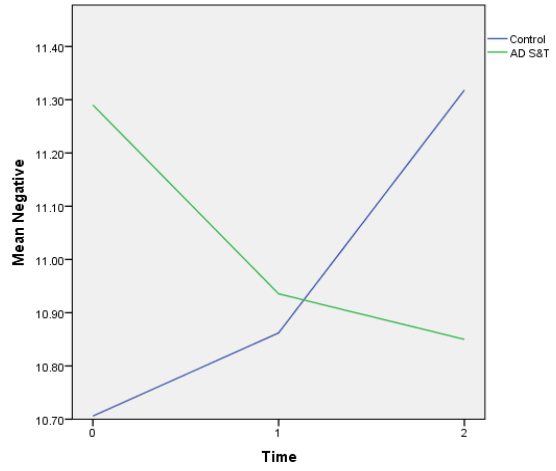
Entitlement



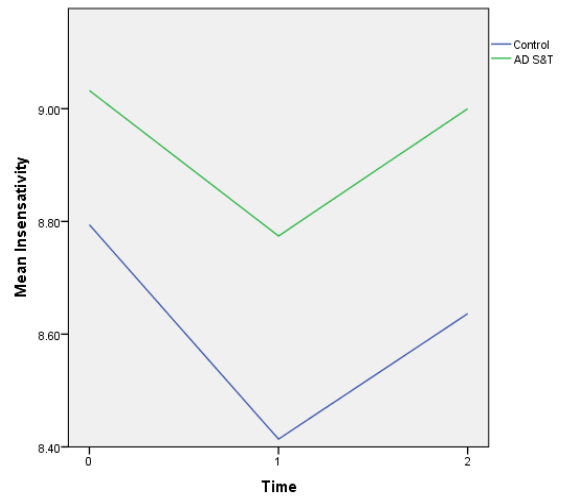
Responsibility



Negativity



Insensitivity



Guilt

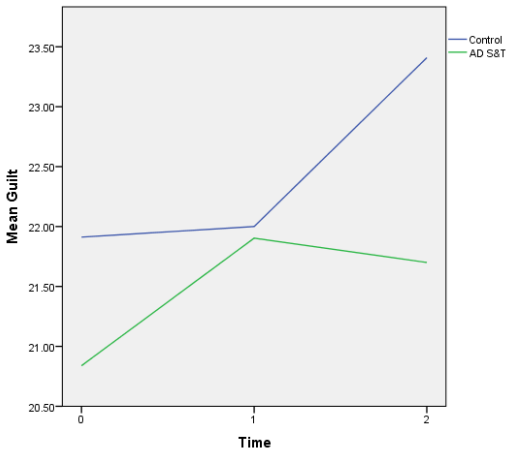


Figure 14: ITT means for CCS total and subscales of WL/TAU and AD participants within the S&T arm at T1 – T3

3.8. Plots for reliable change scores at T2

Figure 15. Plot for PCL-C total outcome for AD S&T partcipants at T2

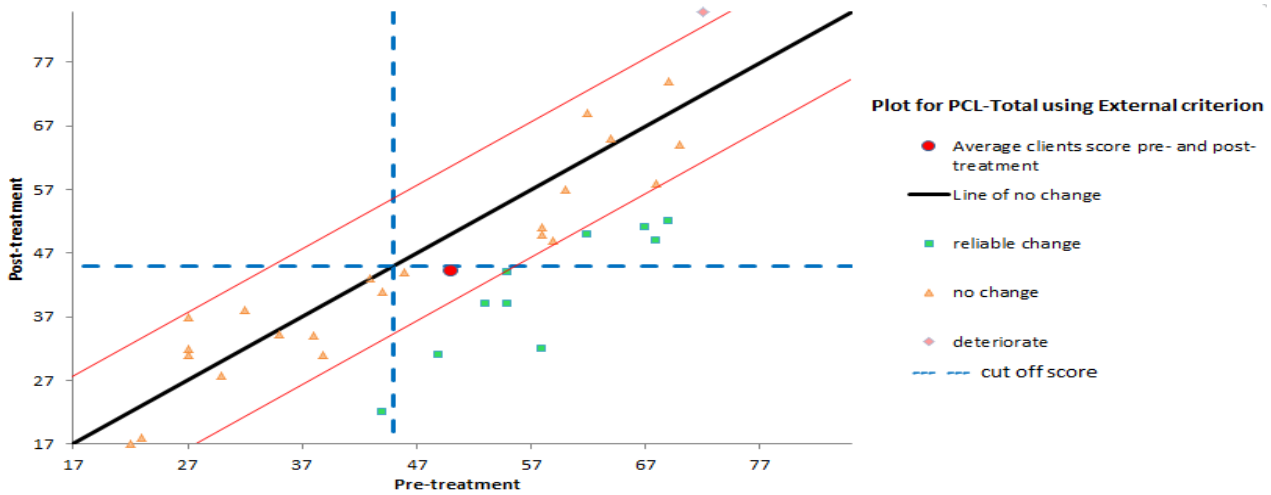


Figure 16. Plot for BAC-R total outcome for AD S&T partcipants at T2

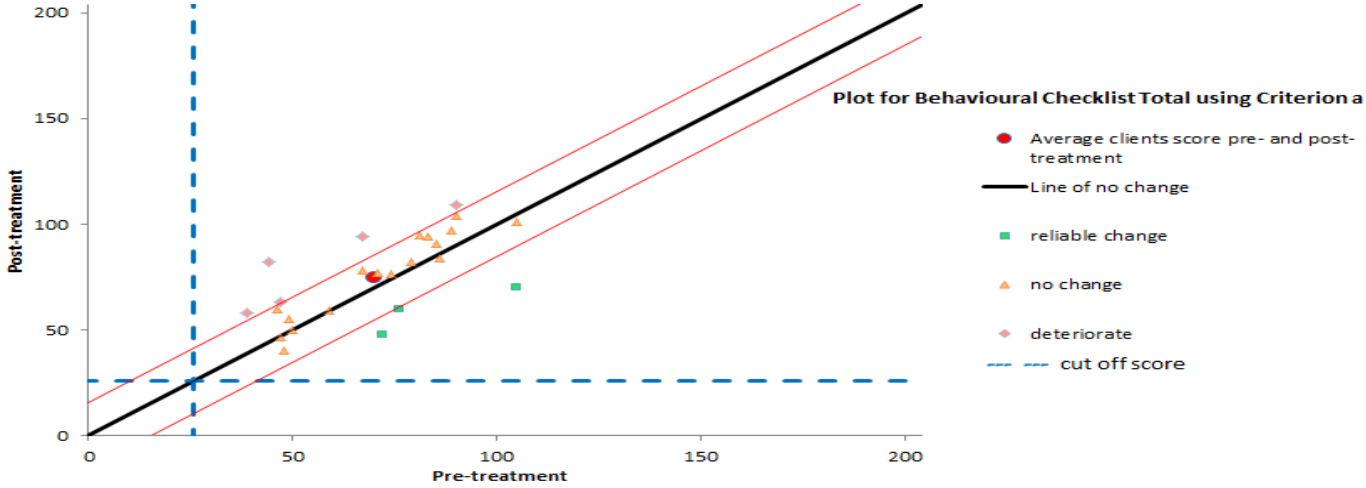


Figure 17. Plot for DERS total outcome for AD S&T partcipants at T2

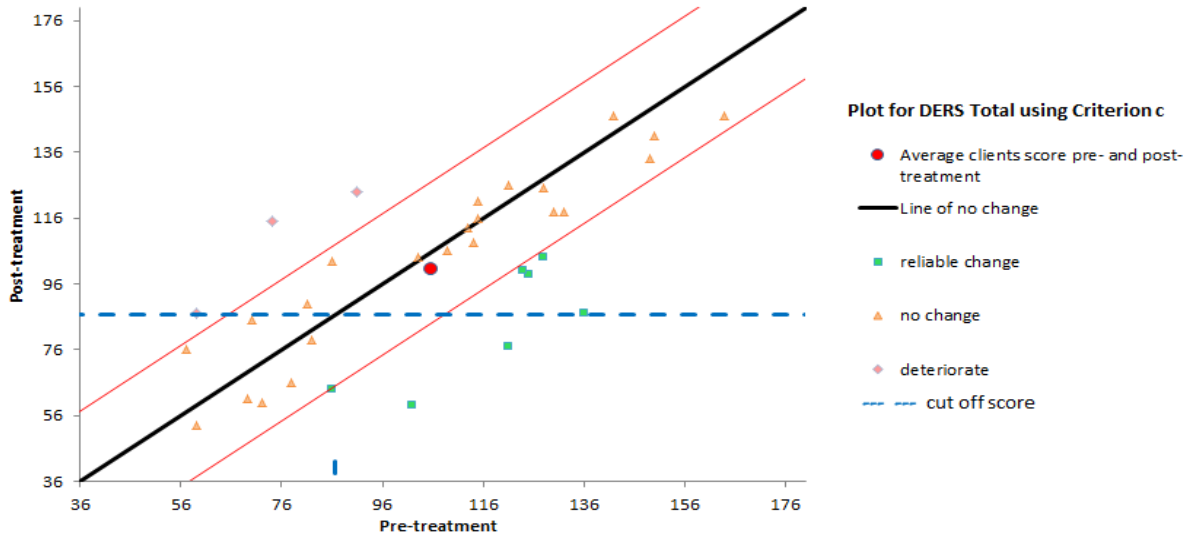


Figure 18. Plot for DES II total outcome for AD S&T partcipants at T2

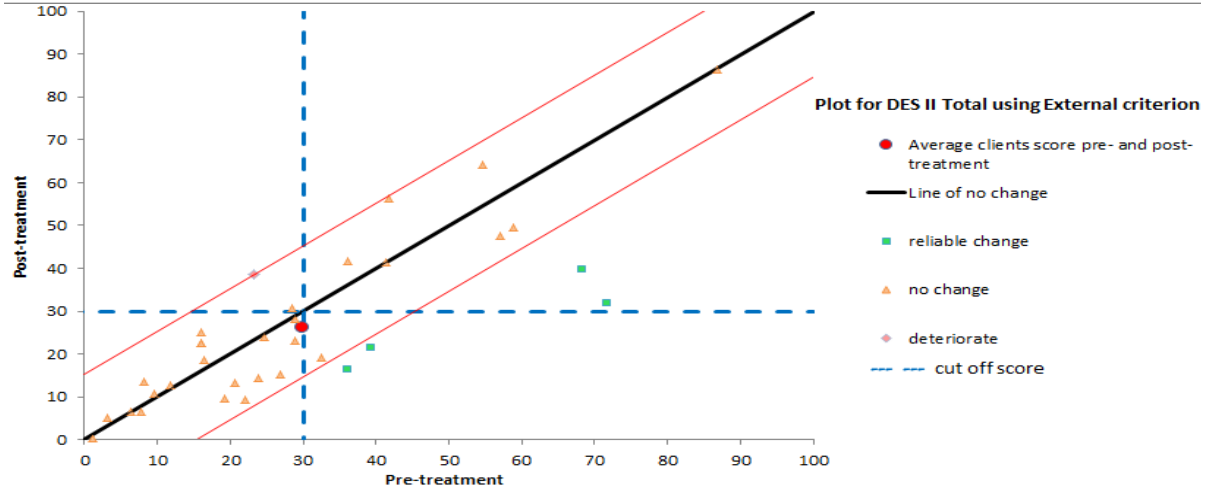


Figure 19. Plot for HADS total outcome for AD S&T partcipants at T2

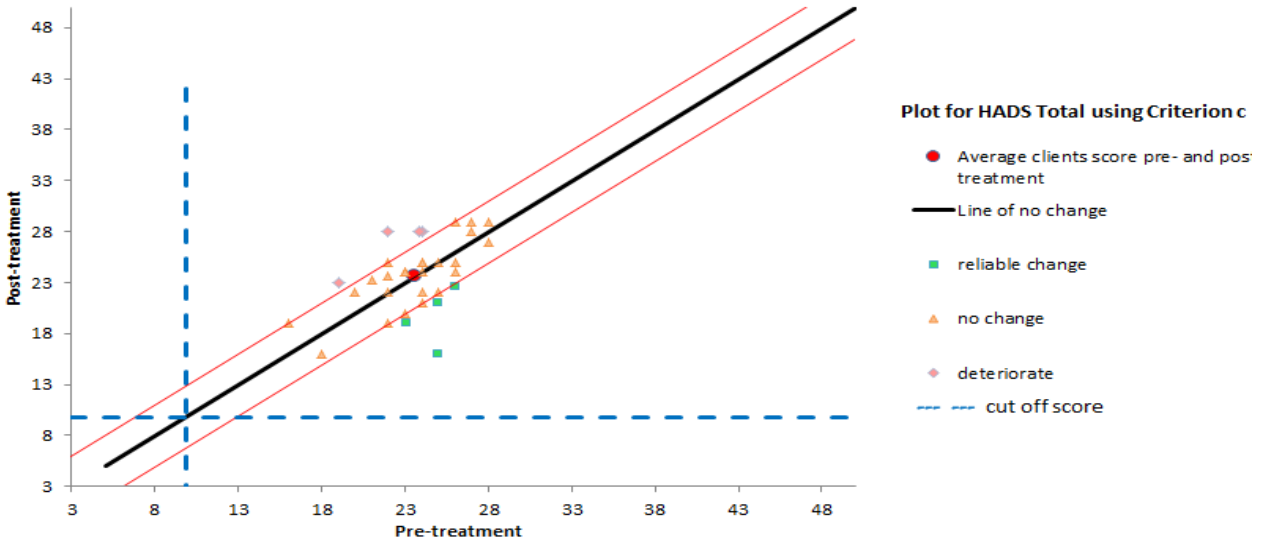
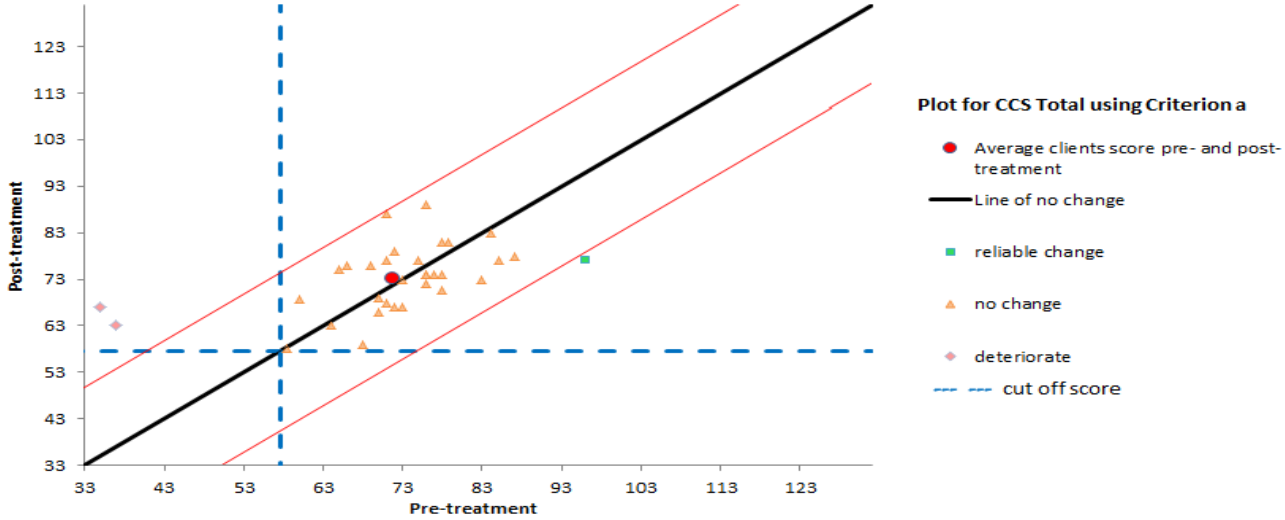


Figure 20. Plot for CCS total outcome for AD S&T partcipants at T2



Plots for Reliable change scores at T3 for PCL-C

Figure 21. Plot for PCL-C total outcome for AD S&T partcipants at T3

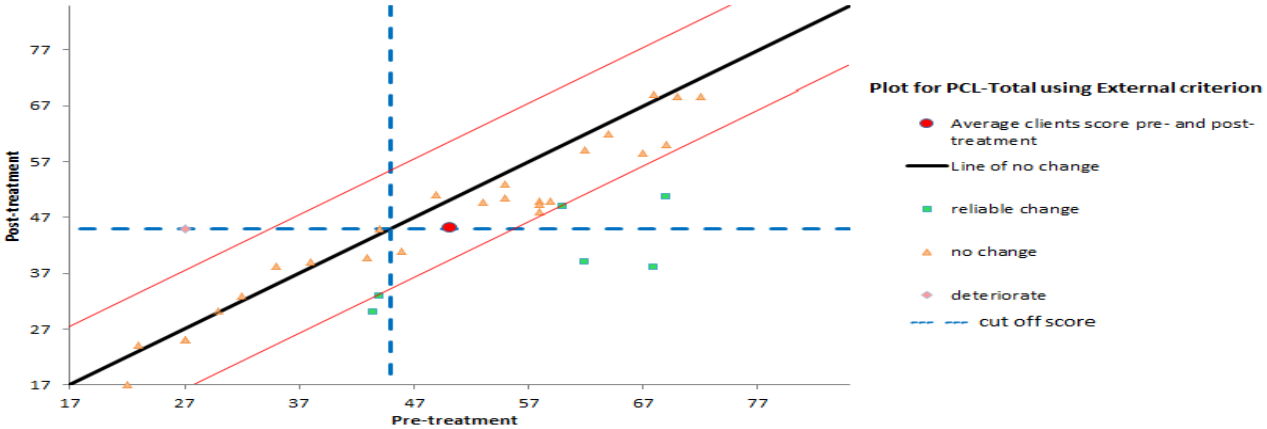


Figure 22. Plot for CCS total outcome for WL/TAU Contro arm partcipants at T3