Title

How important are phase 1 interventions for complex interpersonal trauma? A pilot randomised control trial of a group psychoeducational intervention

Running head

Psychoeducation for complex interpersonal trauma

Authors

[1, 2]Adam Mahoney, PhD

[3]Thanos Karatzias, PhD

[2]Kirsty Halliday, CPsychol

[3]Nadine Dougal, MSc

Affiliations

[1]Glasgow Caledonian University, Psychology Department

[2]HMP & YOI Cornton Vale, Psychology Department

[3]Edinburgh Napier University, School of Health & Social Science

[1]**Address for Correspondence:**
Department of Psychology
Glasgow Caledonian University
Cowcaddens Road
Glasgow
G4 0BA
Scotland, UK

Tel: (++44) 141 273 1804
Email: adam.mahoney@gcu.ac.uk

Word count (exc. figures/tables/references/abstract): 5,287

Conflict of interest statement:

The authors declare that they have no conflict of interest.

Title

How important are phase 1 interventions for complex interpersonal trauma? A pilot randomised control trial of a group psychoeducational intervention

Running head

Psychoeducation for complex interpersonal trauma

Abstract

Although psychoeducational interventions are recommended as phase 1 interventions for complex trauma there is limited evidence on their efficacy. This pilot randomised control trial (RCT) investigated the efficacy of a pure psychoeducational intervention for complex trauma. A brief 10 session intervention was delivered to n=44 female prisoners in a compressed format to accommodate short sentence lengths and was compared to 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 post intervention) for the main outcomes (Behavioural Assessment Checklist-Revised, β= 4.60 [95%CI, -1.60 to 10.88], p= .148; PTSD Checklist, β= -1.47 [95%CI, -4.30 to 1.36], p= .303). Post hoc Reliable Change analyses suggested twice the number of AD participants made progress in addressing PTSD symptoms compared to usual care (30.3% vs 17.6%, OR 2.03 [95%CI, .64 to 6.43]). Whilst further work in this area is required, initial results, overall, suggest that psychoeducational group based treatment modalities achieve only small effect sizes in comparison to usual care.

Key Practitioner Message:

- Small non-significant effect sizes are associated with group based psychoeducational interventions for complex interpersonal trauma.
- The RCT evidence base for the efficacy of phase 1 interventions designed to ameliorate interpersonal trauma in offender populations is not robust.
- Awareness of the potential for initial distress and the need for additional support as part of survivor's pathways to recovery is important.

- Interventions that address a client's primary symptom(s) may hold greater promise than brief general psychoeducational interventions.

Keywords:

complex trauma; PTSD; randomized control trial; group treatment; women

Introduction

Interpersonal trauma is associated with severe and life-long adverse physical and mental health outcomes (Cloitre, Khan, Mackintosh, Garvert, Henn-Haase, Falvey, & Saito, 2019; Felitti, Anda, Nordenberg, Williamson, Spitz, Edwards, Koss & Marks, 1998; López-Martínez, Serrano-Ibáñez, Ruiz-Párraga, Gómez-Pérez, Ramírez-Maestre & Esteve, 2016). It is also associated with poorer social, economic and criminal justice outcomes as well as the shortening of survivor's life spans (Bowen, Jarrett, Stahl, Forrester & Valmaggia, 2018; Bywaters, Bunting, Davidson, Hanratty, Mason, McCartan & Steils, 2016).

Experiences of interpersonal trauma also 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 out of, interpersonal trauma is therefore important if interventions are to be effective. This is particularly so for institutions, such as in women's prisons, where there are high rates of interpersonal trauma (Ney, Van Voohris & Lerner, 2011; Mahoney, 2019).

Various definitions of interpersonal trauma have been positioned (Briere & Elliot, 2003; Cloitre, Stolbach, Herman, van der Kolk, Pynoos, Wang & Petkova, 2009; Courtois & Ford, 2009). However, prior to the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, APA, 2013) and the inclusion of a dissociative subtype there has been little diagnostic recognition of complex manifestations of PTSD (APA, 2000). The recently released International Classification of Diseases, 11th version (ICD-11, World Health Organisation, 2018) has also sought to further define the symptoms of CPTSD.

Phase 1 psychoeducational interventions for interpersonal trauma

Various authors have advocated that trauma responsive interventions should adopt a phased based approach (Ford, Cortois, Steele, van der Hart & Mijenhuis, 2005; Harris & Fallot, 2001; Herman, 1992; van der Hart, Brown & van der Kolk, 1989). Phase 1 has been conceptualised as consisting of 'psychoeducation' with respect to providing both relevant information and actively assisting clients to develop coping and emotion regulation skills thereby promoting safety and stabilisation (Mahoney, Karatzias & Hutton, 2019). It has been advocated that this should occur prior to trauma memory processing (TMP) (Cloitre et al, 2012). Although phase based interventions appear to make clinical sense there is limited evidence of their efficacy (de Jongh, Resick, Zoellner et al, 2016).

To date there are only a limited number of studies evaluating what can be easily identified as psychoeducational group based interventions. Survive & Thrive (Ferguson, 2013), is a brief psychoeducational group based intervention for the stabilisation of symptoms associated with complex interpersonal trauma such as childhood sexual abuse (CSA). Survive & Thrive was developed in response to large numbers of survivors on waiting lists for mental health services and has since been widely delivered across Scotland. It has been considered both a useful and pragmatic intervention in that it promotes awareness of commonly experienced difficulties and how to manage these (NES, 2018; UKPTS, 2017).

Emerging evidence from non-randomised studies has suggested that Survive & Thrive may be useful to stabilise the mental health difficulties that survivors of childhood abuse and multiply traumatised female offenders in the community experience (Ball, Karatzias, Mahoney, Ferguson, & Pate, 2013; Karatzias, Ferguson, Chourliara, Gullone, Gosgrove & Douglas, 2014). However, a recent meta-analysis for group based interventions has suggested that only a small non-significant effect size may be apparent for psychoeducational interventions compared to usual care for PTSD symptoms once potential outliers have been accounted for (Mahoney et al, 2019).

Interpersonal trauma in female prison populations

Female offenders have particularly high rates of interpersonal trauma (Karatzias, Power, Woolston, Apurva, Begley et al, 2018). Therefore, not only is it imperative that otherwise effective treatment approaches are tested with this population but that specific consideration is given to the environmental constraints that may be apparent in institutions

such as prisons. To date no trauma based psychological intervention utilising a RCT methodology has produced unequivocal results with female offenders (Bradley & Follingstad, 2003; Cole, Sarlund-Heinrich & Browne, 2007; Ford, Chang, Levine & Zhang, 2013; Messina, Grella, Cartier & Torres, 2010; Zlotnick, Johnson & Najavtis, 2009).

Research Questions

In light of the limited evidence available the present study sought to investigate the efficacy of Survive &Thrive (S&T), a pure psychoeducation group based intervention for the stabilization (i.e. amelioration) of symptoms associated with complex interpersonal trauma in a female prison population. Specific research questions therefore were:

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

2. Will S&T be an efficacious intervention for reducing symptoms associated with PTSD in female survivors and compared to usual care in a prison setting?

3. Will S&T be an efficacious treatment for reducing general symptoms of psychopathology with female survivors compared to a wait list control group in a prison setting?

4. Will S&T be more efficacious treatment for those participants who receive ≥ 7 sessions ('adequate dose') compared to usual care in a prison setting?

Methods

Participants

Participants were recruited from 2 high security female prison establishments. Ethical approval was granted by the prison service, university and the National Health Service (NHS) Research and Ethics Service. Trial Registration: ISRCTN35772940 (https://doi.org/10.1186/ISRCTN35772940).

Procedure

All convicted women at the trial sites who reported a history of interpersonal violence and trauma and who were over the age of 18 years old were invited to participate in the study. Exclusion criteria were based on ensuring the safety and security of the individual and the establishment. It was emphasised that participation was entirely voluntarily and the multi-disciplinary mental health team (MDMHT) at each site provided independent clinical assurance to this effect. No incentives for participation were provided; the aims of both the intervention and the trial were provided prior to enrolment.

From July 2013 to August 2016, 139 women 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. Some of the women referred to the study, n= 17 (12.2%), declined to participate in the subsequent baseline assessments whilst others withdrew after completing baseline assessments, n= 18 (13%). In addition, n= 15 (10.8%) women were deemed by joint prison (MDMHT) and research treatment management procedures as requiring immediate assistance and for the purposes of this study were not considered eligible for randomisation (McDougall, Clarbour, Perry, & Bowles, 2009). A total of 86 women (61.9%) of the women who had been referred agreed to participate in baseline assessments Self-Report Instrument for Disorders of Extreme Stress (SIDES-SR) and Trauma Antecedents Questionnaire (TAQ, van der Kolk, 2002; 2010). Participants were then randomised to either the control (usual care) waiting list arm or the S&T arm in a 1:1 allocation. This was undertaken by the study's research assistant (RA) in conjunction with another member of the main trial site's psychology department who was not involved in the study. A computer generated randomisation list (randomizer.org) allocated participants to the study's arms. The flow of participants is presented in Figure 1.

The intention-to-treat (ITT) sample consisted of 86 female prisoners of whom n=44 were randomly allocated to S&T and n= 42 to the control waiting list / treatment as usual (WL/TAU) arm of the study. Due to prison based procedures linked to early release and transfer, as well as subsequent refusal, there was a considerable degree of attrition (19.1%) involving participants randomised to the control arm prior to 'pre' assessment i.e. first time point (T1). Further participant loss was also experienced at the 'post' (T2) and 'follow up' (T3) assessments.

FIGURE ONE ABOUT HERE

Participant characteristics are presented in Table 1. Participants were predominantly white and were not in a relationship. Demographics indicated social and economic marginalisation. Most participants were undertaking sentences for violent offending (71.4% and 65.9%, control and S&T arms respectively). The median sentence length for the control arm was 44.5 months (IQR= 102.3) and 25.5 months (IQR= 24.0) for the S&T arm. The difference between the two arms was partly explained by an additional 3.5% of Murder/Culpable Homicide category participants in the control arm and the considerably longer imprisonment terms associated with those sentences. Accumulatively, this resulted in the intervention arm having shorter sentences.

TABLE ONE ABOUT HERE

Experimental design and intervention

The study design utilised a control WL/TAU comparator group. Existing psychotherapeutic and pharmacological treatments (i.e. 'usual care') did not constitute a standardised intervention for interpersonal trauma. S&T (Ferguson, 2008, 2013) is a manualised psychoeducational intervention of 8-10 sessions. Each session focused on a separate symptom or rehabilitative concern often attributed to the distress caused by interpersonal trauma. A further description of S&T contents and associated cognitive, behavioural and affect skills as well as the amendments made for delivery in a prison setting are available in the online supplementary material. Sessions were delivered twice a week over a 5 week period. Approval was granted by the prison service to ensure the adjustments met necessary standards for delivery.

Sample size estimation

A priori power calculations were undertaken to ensure that the study would be able to detect clinically significant differences in PTSD symptomology as a result of the intervention. Power analysis was conducted on the PCL-C because of the absence of other available data. An expected small-medium effect size helped define the calculations using G*Power3 (Faul, Erdfelder, Lang, & Buchner, 2007). This indicated that a total sample of 70 participants would be needed with power set at 0.95 (1 - β) and α= 05. Anticipated intra-class

correlational coefficients (ICC) were also used to adjust the calculated sample size to account for outcome inflation by the group design effect. A lower ICC was anticipated as being the most relevant for psychoeducational interventions (Rutterford, Taljaard, Dixon, Copas, Eldridge, 2015). This 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).

Measures

Participant and sentencing characteristics. A self-report questionnaire was used for the assessment of demographic, forensic and mental health profiles of participants. In addition, basic information about index offences, previous convictions and involvement in other mental health interventions were collated.

Behavioural Assessment Checklist-Revised (BAC-R: as revised for this trial in 2013 from Nugent, Geohagan, & Travers, 2006). A staff/observer rated measure was chosen due to its potential as an objective procedure to rigorously assess behavioural stabilisation outcomes (Cima, 2003; Merckelbach, & Collaris, 2012; Rogers, 2018). Residential prison officers who had regular contact with participants received training to complete the BAC-R. The measure contains 54 items. The six subscales include: Belligerence, Withdrawal, Distress, Impulsivity, Ego-centricity and Problem Solving. Due to previously reported concerns over the construction and integrity of the BAC (McDougal et al, 2009, (pg. 14 and pg. 50) the usual 3 point Likert scale was changed to a new 5 point scale. The internal consistency of the revised measure was assessed at α= .91.

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 be relevant to multiple events in terms of generic 'stressful experiences' as evident in any population. The PCL-C can be scored by providing a total symptom severity score with higher scores indicating greater symptom severity. The subscales in this measure are: (a) Intrusion; (b) Avoidance and (c) Arousal.

Difficulties in Emotional Regulation Scale (DERS; Gratz & Roemer, 2004). The DERS is a 36-item measure 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. There is no official cut-off score. The DERS has excellent internal consistency and good construct validity (Fowler, Carak, Elhai, Allen, Frueh & Oldham, 2014).

Dissociative Experiences Scale (DES II, Bernstein & Putnam, 1986; Carlson & Putnam, 1993). This 28-item measure assesses the frequency of dissociative experiences and includes three subscales: (a) Depersonalisation; (b) Amnestic dissociation and (c) Absorption. Reliability and validity of the DES II has been well established including in offender populations (Mazzotti, Farina, Imperatori, Pruetti, Speranza & Barbaranelli, 2016; Ruiz, Poythress, Lilienfeld & Douglas, 2008).

Hospital Anxiety and Depression Scale (HADS: Zigmond & Snaith, 1983; Snaith, 2003). This brief 14 item 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. Each item is coded from 0 to 3 with increased scores indicating the severity of the symptom.

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. In correspondence with the author of the CCS (Tangney 1/15/2013) a further set of 8 items were made available which specifically introduced a new subscale associated with Reparation (to assuage feelings of guilt, for example, *I owe something to those hurt by my criminal actions*). The internal validity of the new adapted version was tested using at $\alpha = .74$.

Measurement schedule

All measures were administered for participants in the control and S&T arms at pre-treatment (T1), at the end of treatment, after 5 weeks of treatment (T2) and at 1 month following treatment (T3).

Allocation concealment

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 principle investigator (PI; first author), 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 after randomisation if they were on a waiting list.

Treatment fidelity

Supervision was provided by the first author during the trial who undertook the prison adaption of S&T. The first author was trained by the author of the original community version of S&T. Technical expertise to ensure fidelity and consistency to the overall model was provided as and when required.

A random selection of sessions was monitored for quality assurance and treatment integrity purposes by the author of the community version. 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.

Statistical analysis

A primary ITT analysis was conducted in which all participants were retained in the arms to which they were allocated. Treatment effect sizes between S&T versus WL/TAU (control) at T2 and T3 for all outcome measures were calculated using Cohen's *d* (Cohen, 1988). Linear Mixed Models (LMM) as available in SPSS (Version 23) was also used to increase the power and flexibility in analysing outcome data and computing more precise estimates of the differences between the two arms (Seltman, 2018).

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 with a hierarchy of levels, with the interaction being a nested term (Seltman, 2018). Akaike's Information Criterion (AIC) score was used as an indication of model fit (Matuschek, Kligl, Vasishth, Baayen & Bates, 2017). The dependent variable in each computation was the total or subscale scores from the outcome measures. Estimates of fixed effects and confidence intervals (CI) were used to help determine the difference between the arms at follow-up, adjusted for pre-

treatment. Results demonstrating the 95% CI to cross the point of no effect were used to understand the efficacy of S&T as compared to the control arm. Two models were used to analyse change. The first model used a linear scale to assess change across all time points. The second model utilised repeated measures as separate outcomes to determine the difference at 'post' (i.e. T2) and 'follow-up' (i.e. T3) (Heck, Thomas & Tabata, 2014, pg. 237).

To further investigate individual change post hoc Reliable Change Index/Clinically Significant Change (RCI/CSC), analyses were undertaken with adequate dose (AD) treatment participants and those assigned to the control arm. AD was defined as completing ≥7 sessions. This was computed using the procedures described by Jacobson and Truax (1991) and Jacobson, Roberts, Berns & McGlinchey (1999). All analyses were undertaken using the Leeds Reliable Change Indicator (Morley & Dowzer, 2014).

Missing data analysis

Whilst, results from Little's test indicated that data were missing at random (PCL-C, T1-T3: χ^2 = 10.63 (15), p= .778; BAC-R, T1-T3, χ^2 = 39.99 (36), p= .298) missing data within this study can best be understood as due to unforeseen circumstances connected to population management within the female prison estate. As noted in the CONSORT diagram participant attrition was most apparent for the WL/TAU arm at T1. 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 T1 scores (Jakobsen, Gluud, Wetterslev & Winkel, 2017).

Results

Summary of symptoms clusters associated with interpersonal trauma (SIDES-SR) and for trauma histories (TAQ) baseline assessments are presented in Tables 2 and 3.

TABLE TWO ABOUT HERE

It was noted that participants indicated a greater lifetime occurrence than current presence for each of the SIDES-SR symptom domains.

TABLE THREE ABOUT HERE

Treatment fidelity outcomes

An overall composite score was calculated from the quality assurance form. 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. Results suggested that treatment quality explained only a very small amount of the variance (R^2 of 0% - 9%) and did not significantly predict outcome. The association between quality assurance and dose was not significant (χ^2= 3.74(1), p= .053).

Comparative treatment effectiveness

Descriptive statistics and effect sizes for the two arms are presented in Table 4. ITT and AD statistics are presented in Table 5 in which the results from LMM analysis are also presented. The best fit for all analyses was a diagonal covariance matrix. This enabled subjects to demonstrate variation in their linear trajectories. See online supplement for subscale outcome data and analyses.

TABLE FOUR ABOUT HERE

Effects on behavioural stability

As noted in Table 5 there were few statistically significant differences across the majority of measures in the ITT and AD analyses. Small, non-statistically significant symptom increases for the S&T arm were evident in BAC-R. The BAC-R total score indicated that the symptom increase was predominantly observed 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 arms and results indicated a wide variance in participant's outcomes (β= 2.99 [95% CI, -10.97 to 16.96], p= .668).

Small to medium effect sizes (d= .08 to .55) noted symptom increases for the S&T arm across the BAC-R subscales in the AD analysis. This included a statistically significant increase for S&T in the BAC-R *Distress* subscale (β= 1.75, 95% CI .24 to 3.26, p= .023); with statistical significance being apparent after T2.

Effects on PTSD symptomatology

Small, non-statistically significant reductions was observed for both arms as reflected in the PCL-C total mean differences and effect sizes in Table 4. These mean differences are larger for the S&T arm. The AD analysis also reported an increase in effect sizes in favour of S&T participants although these were still small (T2: *d*= .18 and T3: *d*= .23). The interaction analysis also noted that there were no statistically significant differences between the study's arms in either the ITT or AD analyses.

The PCL-C and BAC-R total scores across the assessment time points highlight that reductions in the self-reported scores in the PCL-C do not appear to be systematically associated with staff reported outcomes in the BAC-R.

Effects on emotional regulation and psychopathology

An exception to the non-significant differences in outcomes between the arms was observed in the HADS Depression subscale (ITT analysis: β= .95, 95% CI .11 to 1.79, p= .027). This reflected a decrease in the WL/TAU arm's level of depression and a corresponding increase in the S&T arm. This statistically significant difference occurred after T2 and was also observed in the AD S&T analysis.

Another statistically significant finding was evident in the time x study arm interaction for the DERS Non-Acceptance subscale (ITT analysis: β= -1.65, 95% CI -3.22 to -.07, p= .041). This indicated a decrease in the S&T arm's emotional non-acceptance compared to the WL/TAU arm. The largest difference between the two study arms occurred after T2 which was also noted in the AD analysis.

It should be considered that in situations where multiple comparisons are made, statistically significant differences as described in these results 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 checked. The estimates and significance levels as reported in Tables 5 remained the same.

TABLE FIVE ABOUT HERE

Distributions of clinical significance and reliable change

A post hoc analysis indicated distributions of reliable change (RC) and clinical significant change (CSC), as presented in Table 6 were similar for both arms. These analyses were only completed for AD participants (n= 33) and participants in the WL/TAU arm (n= 34) who had completed T1 assessments. This produced a comparable number of participants for both arms and minimised missing data.

TABLE SIX ABOUT HERE

Within the analyses it was noted that more AD S&T participants than WL/TAU participants achieved positive RC at T2 for the PCL-C (30.3% vs 17.6%, OR 2.03 [95% CI, .64 to 6.43). In terms of clinical significant change this was still in favour of AD S&T participants but to a less extent. This comparative improvement for AD S&T participants was particular seen in the RC outcomes 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. However, for the PCL-C total differences for positive RC in favour of AD S&T participants were not evident by T3. Although somewhat more AD S&T participants had achieved positive RC in respect to the Arousal subscale (9.1% vs 2.94%, OR 3.19 [95% CI, .32 to 32.36]) at T3.

The RCI from the BAC-R results indicated that pathological behaviours were more reliably stabilised in the WL/TAU arm at T2 (BAC-R total score: 25.0% vs 12.0%, OR .41 [95 CI%, .09 to 1.80]). There were no substantive differences between the study's arms across the other outcome measures. The exception to this was the HADS Anxiety subscale where it was particularly noted that more AD S&T participants achieved positive reliable change (36.4% vs 2.9%, OR 18.86 [95% CI, 2.28 to 155.86]). Whilst there were comparable numbers of participants who achieved no RC on the HADS *Depression* subscale more people in the WL/TAU arm achieved positive RC (11.8% vs 0%).

Discussion

The purpose of the current study was to investigate the efficacy of S&T a phase 1 intervention for symptoms associated with complex interpersonal trauma with female offenders in a prison setting. The results from this RCT's ITT analysis suggest that psychoeducation group based treatment modalities achieve only small non-significant effect sizes in comparison to usual care.

As an intervention designed to ameliorate PTSD symptoms, S&T did not produce statistically or clinically significant levels of change compared to a WL/TAU control group. As a psychoeducational intervention S&T's limited impact on trauma symptomatology was expected. However, as an intervention designed to stabilise symptoms there are concerns about how effective this intervention might be in a prison setting. The evidence from previous RCTs in prison settings has also indicated unfavourable or very limited outcomes (Bradley & Follingstad, 2003; Cole et al, 2007; Ford et al, 2013; Messina et al, 2010; Zlotnick et al, 2009).

In this respect, psychoeducation group treatment modalities have been demonstrated to achieved predominantly (excluding outliers) small effect sizes when compared to usual care (Mahoney et al, 2019). However, direct comparisons undertaken by Mahoney et al (2019) between psychoeducational and TMP interventions did not established the superiority of the later. Similar, results were also found in other studies (Foa, McLean, Zang et al, 2018; Greger, Munder & Bath, 2014; Tran & Gregor, 2016). The most similar previous psychoeducation trial to this study was conducted by Ford et al (2013). These authors also noted a slight increase in negative mood for their experimental arm and no statistically significant difference in the main PTSD and affect regulation measures when compared to a non-trauma informed active support group.

Clinical implications

The comprehensive assessment and statistical analysis undertaken for this trial has helped to establish that small non-significant effect sizes are associated with phase 1 psychoeducational interventions. However, statistically significant increases in Depression (HADS subscale), Distress (BAC-R subscale) and decreases in Non-Acceptance (DERS subscale) were apparent in the intervention arm which had not previously been evidenced in other trials. Obviously, an important question that any clinician needs to consider is whether there is the potential for adverse effects and iatrogenic harm to be caused by the interventions they deliver (Berk & Palmer, 2009; Crawford, Thana, Farquharson, Palmer, Hancock, Bassett, Clarke & Parry, 2016). It is important also to give some consideration as to whether trauma informed psychoeducational interventions, lead to temporary rises in depression and negative affect more generally. Certainly, both arms continued to return

mean scores above the cut-off range considered clinically meaningful for both HADS subscales. Therefore, ensuring that participants undergoing potentially challenging interventions are supported appropriately by all staff involved in an individual's care is important (Harris & Fallot, 2001). It can be hypothesised that increases in emotional acceptance may be linked to increases in depression and that this may also be an important precursor to greater emotional regulation (Gratz & Roemer, 2004; Ratcliffe, 2018). Herman (1998) also describes the potential for difficult emotions to arise in the early stages of treatment: '…the descent into mourning is a necessary but dreaded part of the recovery process'. Further work is, therefore, needed to understand how low mood and depression could be associated with participation in trauma informed psychoeducational interventions. Similarly, as differences in depression were only significant one month after treatment, consideration should also be given to the importance of treatment endings and post intervention effects (Flory & Yehuda, 2015). It is also important to consider how brief psychoeducational interventions may cause particular iatrogenic harm through either not providing valid trauma informed or trauma focused assistance (Sweeney & Taggart, 2018; van Vilet, Huntjens, van Dilk & de Jongh, 2018).

The AD and the RC analyses noted improvements in the HADS Anxiety subscale for S&T. Arguably, this is because S&T like many treatment approaches for PTSD and complex trauma focuses on helping survivors down regulate symptoms of distress (Pai, Suris & North, 2017; Hyland, Shevlin, Fyvie & Karatzias, 2018). Whilst any improvements are welcomed it is important for a more complete range of improvements to be realised. A greater focus in relational functioning may also lead to wider and more robust outcomes (Ford et al, 2013, Mahoney & Karatzias, 2012).

This trial included a broad sample of female prisoners who had sought assistance with a range of interpersonal trauma symptoms and experiences. As such high levels of lifetime symptoms particularly with respect to alterations in relationships, self-perception, affect regulation and consciousness were reported. However, participants also reported a reduced level of current symptomology. The relative emotional and relational stability that prison provided for the women involved in this study may partially explain this. Such holistically orientated environments may of course not be present across all jurisdictions.

Future research and development

Whilst there are indications of reliable change for AD S&T participants such post hoc comparisons should be treated with caution. Similarly, symptom reduction, in term of reliable change, did not appear to be maintained one month post intervention. Nevertheless, future investigations may choose to peruse this approach if a more idiosyncratic analysis of participants' stabilisation is considered appropriate. Such approaches should also include establishing the symptom and treatment profiles of 'completers' and 'dropouts', for both study arms, thereby investigating any potential biases that might impact on trial outcomes (Imel, Laska, Jakcupcak & Simpson, 2013; Spieth, Kubasch, Penzlin et al, 2016).

Currently, from the ITT results of this study and the existent research it is the authors' views that new and innovative approaches need to be developed in order for interventions to more effectively meet the needs of survivors. It seems increasingly likely that new treatment models, beyond phased approaches, for treating complex interpersonal trauma need to be developed and tested (Karatzias & Cloiter, 2019). This may include developing interventions for specific populations or symptoms, as seen in Mahoney et al (2019) *Psychoeducational Plus* category, the inclusion of TMP protocols or a greater focus on interoceptive awareness skills (Bradley & Follingstad, 2003; Garland, Robert-Lewis, Tronnier, Graves & Kelley, 2016). Further, studies are also needed to investigated the efficacy of trauma-focused interventions with populations who have complicated forensic histories (López-Castro, Smith, Nicholson, Armas, & Hien, 2019; Mahoney, 2019).

Strengths and Limitations

The study was not sufficiently powered. As such given the dropout rate experienced in this study a larger sample size than was possible to recruit would be needed, particularly for a psychoeducational intervention, to demonstrate suitable public health gains (Ali, Rhodes, Moreea, McMillan, Gilbody, Leach, Lucock, Lutz & Delgadillio, 2017; Button, Ioannidis, Mokrysz, Nosek, Robinson & Munafo, 2013). Thus, it is imperative that results are replicated in a definitive trial with greater power.

The sporadic return rates of the BAC-R from staff presents a serious weakness in the interpretation of this measure and results should be interpreted with extreme caution. It is extremely difficult to rule out any potential bias in the administration of the BAC-R even with the support of the RA (McDougall et al, 2009). It is also recognised that having more clearly defined outcomes in terms of 'stabilisation' is important. However, studies with forensic populations may be particularly vulnerable to outcomes that represent a regression towards the mean emphasising the need for randomised control trials (Linden, 2013; Morton & Torgerson, 2003). The study was conducted within highly pragmatic parameters and data was analysed over a one month follow up period highlighting the trajectory of change. Both the LMM and RC/CSC analyses have measured participant change at the idiosyncratic level which is an important consideration given variations in symptom expression, chronicity and recovery.

Given the extensive analysis of subscales type 1 and other measurement errors should be considered (Barnett, van der Pols & Dobson, 2005). In both arms there is a lack of substantive change for the main outcome measures (BAC-R and PCL-C) across all assessment time points. 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.

Conclusions

On the basis of this trial and the existent RCT literature it is not possible to recommend S&T or other group based psychoeducational interventions as an effective stand-alone treatment for interpersonal trauma. Delivering a TMP based package of care may lead 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 brief psychological interventions seek to replicate some of the larger effect sizes as seen for more targeted *Psychoeducation Plus* interventions as described by Mahoney et al (2019). Developing protocols that address depression and other specific symptoms and expressions of interpersonal trauma may be more appropriate. It is therefore important that, as with any psychoeducational intervention, realistic expectations are set as to their effectiveness.

References

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. doi: 10.1016/j.brat.2017.04.006

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC, Author.

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. doi: 10.1080/14789949.2013.852233

Briere, J. & Elliot, D. M. (2003). Prevalence and psychological sequelae of self-reported childhood physical and sexual abuse in a general population sample of men and women. *Child Abuse & Neglect*, 27, 1205–1222. doi: 10.1016/j.chiabu.2003.09.008

Barnett, A., van der Pols, J. C. & Dobson, A. J. (2004). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*, 34, 215–220. doi: 10.1093/ije/dyh299

Berk, M. & Palmer, G. (2009). The elephant on the couch: side-effects of psychotherapy. *Australian and New Zealand Journal of Psychiatry*, 43, 787-794. doi: 10.1080/00048670903107559

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

Blanchard, E. B., Jones-Alexander, J., Buckley, T. C. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, 34, 669-673. doi: 10.1016/0005-7967(96)00033-2

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. doi: 10.1016/j.paid.2018.04.023

Bywaters, P., Bunting, L., Davidson, G.,Hanratty, J., Mason, W., McCartan, C. & Steils, N. (2016). The relationship between poverty, child abuse and neglect: an evidence review. Joseph Rowntree Foundation. Retrieved from: https://www.jrf.org.uk/report/relationship-between-poverty-child-abuse-and-neglect-evidence-review

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. doi: 10.1023/A:1024409817437

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 Neuroscience*, 14, 365-376. doi: 10.1038/nrn3502

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

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. doi: 10.1076/clin.17.2.235.16507

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. doi: 10.3402/ejpt.v4i0.2070

Cloitre, M., Khan, C., Mackintosh, M.-A., Garvert, D. W., Henn-Haase, C. M., Falvey, E. C., & Saito, J. (2019). Emotion regulation mediates the relationship between ACES and physical

and mental health. *Psychological Trauma: Theory, Research, Practice, and Policy*, 11(1), 82-89. doi: 10.1037/tra0000374

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. doi: 10.1002/jts.20444

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

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. doi: 10.1300/J229v08n02_07

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

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. doi: 10.1192/bjp.bp.114.162628

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. doi: 10.1002/da.22469

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. doi: 10.1191/1740774504cn006rr

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. doi: 10.3758/BF03193146

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

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

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., McLean, C. 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. doi: 10.1001/jama.2017.21242

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. doi: 10.1016/j.beth.2012.10.003

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. doi: 10.1002/jts.20051

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. doi: 10.1016/j.jpsychires.2014.07.029

Garland, E.L., Robert-Lewis, A., Tronnier, C. D., Graves, R., & Kelley, 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. doi: 10.1016/j.brat.2015.11.012

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. doi: 10.1017/S0033291714000853

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(1):406. doi: 10.1186/s13063-018-2735-z

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. doi: 10.1007/s10862-008-9102-4

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.

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

Herman, J. (1992). *Trauma and Recovery: The Aftermath of Violence – from Domestic Abuse to Political Terror*. New York: Basic Books.

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. doi: 10.1002/jts.22272

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 Clinical Psychology*, 81(3), 394–404. doi: 10.1037/ccp0000143

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. doi: 10.1037/0022-006X.67.3.300

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. doi: 10.1037/0022-006X.59.1.12

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. doi: 10.1186/s12874-017-0442-1

Karatzias, T. & Cloitre, M. (2019). Treating adults with complex posttraumatic stress disorder using a modular approach to treatment: rationale, evidence, and directions for future research. *Journal of Traumatic Stress*, 32 (6), 870-876. doi: 10.1002/jts.22457

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. doi: 10.1521/ijgp.2014.64.4.492

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. doi: 10.1002/cbm.2043

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. doi: 10.1016/j.biopsych.2012.11.020

Linden, A. (2013). Assessing regression to the mean effects in health care initiatives. *BMC Medical Research Methodology*, 13:119. doi: 10.1186/1471-2288-13-119

López-Castro, T., Smith, K. Z., Nicholson, R. A., Armas, A., & Hien, D. A. (2019). Does a history of violent offending impact treatment response for comorbid PTSD and substance use disorders? A secondary analysis of a randomized controlled trial. *Journal of Substance Abuse Treatment*, 97, 47–58. doi: 10.1016/j.jsat.2018.11.009

López-Martínez, A.E., Serrano-Ibáñez, E. R., Ruiz-Párraga, G. T., Gómez-Pérez, L., Ramírez-Maestre, C., & Esteve, R., 2016). Physical health consequences of interpersonal trauma: a systematic review of the role of psychological variables. *Trauma, Violence, & Abuse*, 19 (3), 305-322. doi: 10.1177/1524838016659488

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.

Mahoney, A. (2019). The complexity of 'complex' trauma. *The Psychologist*, December, 2019. Retrieved from: https://thepsychologist.bps.org.uk/complexity-complex-trauma.

Mahoney, A., Karatzias, T., & Hutton, P. (2019). A systematic review and meta-analysis of group treatments for adults with symptoms associated with complex post-traumatic stress disorder. *Journal of Affective Disorders*, 243, 305–321. doi: 10.1016/j.jad.2018.09.059

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. doi: 10.1080/14999013.2012.723664

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. doi: 10.1016/j.jml.2017.01.001

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. doi: 10.2147/NDT.S105110

MacDonald, K., Thomas, M.L., MacDonald, T. M., & Sciolla, A. F. (2015). A perfect childhood? clinical correlates of minimization and denial on the Childhood Trauma Questionnaire. *Journal of Interpersonal Violence*, 30(6) 988–1009. doi: 10.1177/0886260514539761

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.

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. doi: 10.1016/j.jsat.2009.09.004

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. doi: 10.1186/s13063-018-2733-1

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. doi: 10.1136/bmj.326.7398.1083

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. doi: 10.1249/01.MSS.0000126806.72453.1C

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.

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.

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.doi: 10.3390/bs7010007

Ratcliffe, M. (2018). The interpersonal structure of depression. *Psychoanalytic Psychotherapy*, 32(2), 122-139. doi: 10.1080/02668734.2018.1455729

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.

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. doi: 10.1016/j.jclinepi.2014.10.006

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. doi: 10.1177/1073191108315548

Seltman, H. J. (2018). Experimental Design and Analysis. Retrieved from: www.stat.cmu.edu/~hseltman/309/Book/Book.pdf

Shine, J. (2003). *General Principles for Video Monitoring. Guidance notes for Prison and Probation Service Treatment Managers* (Draft IV). London, HMPS & NPD.

Spieth, P. M., Kubasch, A. S., Penzlin, A. I. et al, (2016). Randomized controlled trials – a matter of design. Neuropsychiatric Disease and Treatment, 12, 1341–1349. doi: 10.2147/NDT.S101938.

Sweeney, A. & Taggart, D. (2018). (Mis)understanding trauma informed approaches in mental health. *Journal of Mental Health*, 27 (5), 383-387. doi: 10.1080/09638237.2018.1520973

Tangney, J.P., Meyer, P., Furukawa, E., Cosby, B. (2002). *The Criminogenic Cognitions Scale*. George Mason University; Fairfax, VA.

Tran, U. S. & Gregor, B. (2016). The relative efficacy of bona fide psychotherapies for post-traumatic stress disorder: a meta-analytical evaluation of randomized controlled trials. *BMC Psychiatry*, 16:266. doi: 10.1186/s12888-016-0979-2

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

van Vilet, N., Huntjens, R. J. C., van Dilk, M. K., & de Jongh, A. (2018). Phase-based treatment versus immediate trauma-focused treatment in patients with childhood trauma-related posttraumatic stress disorder: study protocol for a randomized controlled trial. *Trials*, 19:138. doi: 10.1186/s13063-018-2508-8

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. doi: 10.1002/jts.2490020404

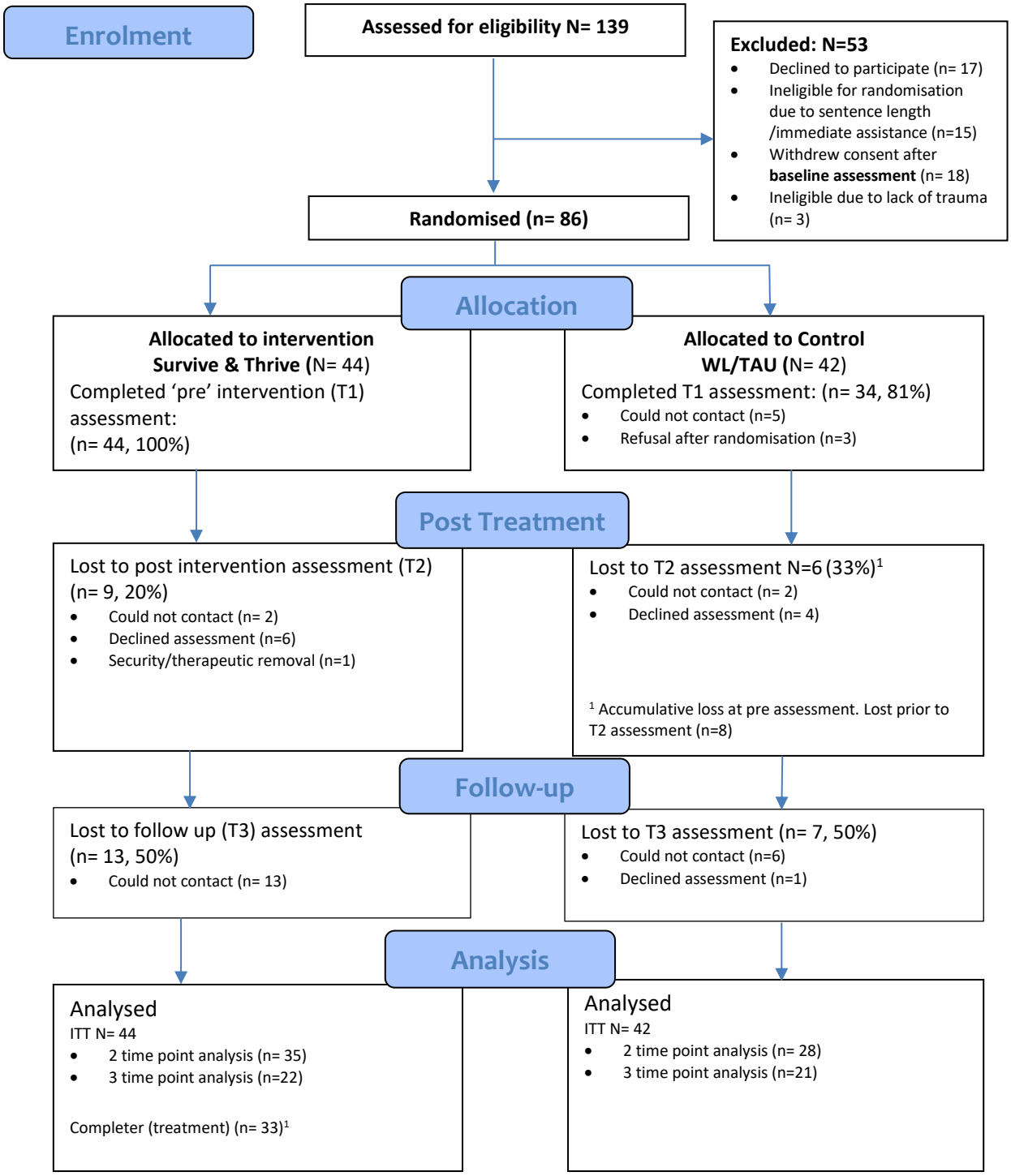
van der Kolk, B. (2002). Structured Interview for Disorders of Extreme Stress –Self Report (SIDES-SR). *Trauma Assessment Packet*. Brookline, MA., The Trauma Centre.

van der Kolk, B. (2010) Traumatic Antecedents Questionnaire (TAQ). *Trauma Assessment Packet*. Brookline, MA., The Trauma Centre.

WHO (2018). International Classification of Diseases-11. World Health Organisation. Retrieved from: https://icd.who.int8

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67, 361-370. doi: 10.1111/j.1600-0447.1983.tb09716.x

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. doi: 10.1016/j.beth.2008.09.004



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

Figure 1: CONSORT Diagram

Table 1. Sociodemographic, forensic and clinical characteristics of participants

	WL/TAU (n=42) M(SD) / N(%)	Survive & Thrive (n=44) M(SD) / N(%)
Age (years old)	33.17 (10.32)	33.53 (10.34)
Ethnicity		
White	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?[1]	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[2]	14 (33.3)	11 (25.0)
Violence[3]	16 (38.1)	18 (40.9)
Drug offence	6 (14.3)	9 (20.5)
Other offending[4]	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 Drug Offence?[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. All N% presented in the affirmative (i.e. 'Yes'). 1) Median and minim – maximum reported. 2) including Attempted Murder, 3) includes assault, fire arm offences and 'Schedule 1' offences of physical/emotional offences against a child. There were no sexual offences included within this sample. 4) includes shoplifting, theft, public order offence. As baseline measures were collected prior to randomisation statistical comparisons between the arms were not considered appropriate.

Table 2. Frequency and type of traumatic events at baseline

		Child (0-18 yrs) N (%)	Adult (≥ 19 yrs) N (%)	Lifetime N (%)
TAQ: Traumatic Events				
Neglect	WL/TAU	32 (76.2)	29 (69.0)	33 (78.6)
	S&T	34 (77.3)	37 (84.1)	41 (93.2)
Separation	WL/TAU	37 (88.1)	38 (90.5)	38 (90.5)
	S&T	38 (86.4)	37 (84.1)	42 (95.5)
Emotional Abuse	WL/TAU	30 (71.4)	32 (76.2	34 (80.1)
	S&T	31 (70.5)	33 (75.0)	36 (81.8)
Physical Abuse	WL/TAU	26 (61.9)	25 (59.5)	30 (71.4)
	S&T	30 (68.2)	28 (63.6)	38 (86.4)
Sexual Abuse	WL/TAU	24 (57.1)	15 (35.7)	27 (64.1)
	S&T	21 (47.7)	8 (13.2)	24 (54.5)
Witnessing	WL/TAU	35 (83.3)	31 (73.8)	35 (83.3)
	S&T	32 (72.7)	27 (61.4)	37 (84.1)
Other trauma	WL/TAU	33 (78.6)	36 (85.7)	38 (90.5)
	S&T	32 (72.7)	28 (63.6)	37 (84.1)
Alcohol / drugs	WL/TAU	34 (81.0)	35 (83.3)	38(90.5)
	S&T	34 (77.3)	33 (75.0)	35 (79.5)
Safety	WL/TAU	33 (78.6)	35 (83.3)	38 (90.5)
	S&T	37 (84.1)	35 (79.5)	40 (90.1)
Competence	WL/TAU	33 (78.6)	33 (78.6)	37 (88.1)
	S&T	37 (84.1)	32 (88.1)	40 (91.0)
No. of Trauma Categories[1]	WL/TAU	5.97 (1.86)	5.70 (1.70)	6.46 (1.53)
	S&T	5.74 (2.16)	5.21 (1.76)	6.48 (1.21)

Notes: WL/TAU n=42; S&T n=44. [1]Mean (SD) of trauma categories excluding Safety and Competence

Table 3. Summary of participant's lifetime occurrence, current presence and current severity of trauma symptom clusters at baseline.

		Lifetime N (%)	Current N (%)	Current Severity M (SD)
SIDES-SR: Trauma Symptoms Domains[1]				
Affect Dysregulation	WL/TAU	28 (66.7%)	22 (52.4%)	.80 (.74)
	S&T	32 (72.7%)	20 (45.5%)	.71 (.64)
Attention/Consciousness	WL/TAU	38 (90.5%)	25 (59.7%)	1.02 (.96)
	S&T	39 (88.6%)	23 (52.3%)	.97 (97)
Self-Perception	WL/TAU	35 (83.3%)	18 (42.9%)	.68 (.74)
	S&T	40 (90.9%)	19 (43.2%)	. 65 (.58)
Relations with Others	WL/TAU	39 (92.9%)	25 (59.5%)	.82 (.81)
	S&T	43 (97.7%)	27 (61.4%)	.79 (.70)
Somatisation	WL/TAU	21 (50.0%)	5 (11.9%)	.27(.48)
	S&T	15 (34.1%)	11 (25.0%)	.36 (.51)
Systems of Meaning	WL/TAU	35 (83.3%)	13 (31.0%)	.47 (.76)
	S&T	37 (84.1%)	20 (45.5%)	.64 (59)

Notes: WL/TAU= Wait Listed/ Treatment as Usual control arm (n=42); S&T= Survive & Thrive (n=44) intervention arm. SIDES-SR= Structured Interview for Disorders of Extreme Stress – Self Report. Lifetime= lifetime presence; Current= Meets current diagnostic criterial; Severity= current severity ratings scale 0-3. [1]All trauma symptom domains indicate adverse alterations in functioning.

Table 4. ITT and AD Survive & Thrive outcomes and the control arm for all measures across time points

		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*	
PCL-C Total	WL/TAU	49.35(16.19)	48.24(16.84)	47.11(18.04)	1.10 (9.41)	2.23 (10.57)		
	S&T ITT	50.60(16.56)	47.24(16.62)	45.26(14.41)	3.36 (11.74)	5.34 (10.24)	.06	.11
	S&T AD	50.38 (15.84)	45.31 (15.91)	43.54 (12.79)	5.07 (9.91)	6.85 (10.08)	.18	.23
BAC-R Total	WL/TAU	67.79 (20.23)	65.54 (20.81)	68.26 (15.31)	2.25 (14.05)	-.48 (24.59)		
	S&T ITT	69.94 (23.72	73.18(19.84)	72.64 (18.55)	-3.24 (16.34)	-2.7 (16.54)	.38	.26
	S&T AD	69.96 (19.64)	74.52 (20.26)	75.73 (18.96)	-3.76 (16.68)	-3.75 (16.26)	.44	.43
DERS Total	WL/TAU	106.33 (29.27)	99.95 (27.24)	94.14 (29.49)	6.39 (16.87)	12.20 (18.78)		
	S&T ITT	106.86(27.67)	103.83(28.47)	93.27 (31.23)	3.99 (22.10)	13.65 (22.26)	.11	.03
	S&T AD	106.83 (29.12)	100.46 (27.31)	90.04 (28.67	6.37 (21.3)	16.8 (20.37)	.02	.14
DES Total	WL/TAU	28.27(20.43)	29.68 (21.59)	28.56 (21.24)	-1.41 (11.31)	-.28 (10.97)		
	S&T ITT	29.41(20.21)	29.24 (19.41)	27.63 (16.21)	.17 (13.03)	1.78 (10.60)	.02	.05
	S&T AD	29.84 (20.94)	26.27 (19.11)	25.64 (14.36)	3.57 (12.05)	4.2 (9.16)	.17	.16
HADS Anxiety	WL/TAU	11.97(3.73)	11.95 (3.47)	11.50 (4.21)	.02 (2.02)	.46 (3.54)		
	S&T ITT	12.61(2.68)	11.58 (3.32)	11.64 (2.69)	1.02 (3.03)	.97(3.19)	.11	.04
	S&T AD	12.53 (2.62)	11.32 (3.25)	11.38 (2.33)	1.21 (3.02)	1.15 (3.08)	.19	.04
HADS Depression	WL/TAU	12.12(2.50)	11.92 (3.08)	11.87 (3.42)	.19 (2.61)	.24 (1.70)		
	S&T ITT	11.11(3.31)	12.30 (2.70)	12.15 (3.13)	-1.20 (2.53)	-1.05 (3.92)	.13	.09
	S&T AD	11.07 (3.34)	12.47 (2.76)	12.41 (3.00)	-1.40 (2.52)	-1.33 (3.99)	.18	.17
CCS Total	WL/TAU	70.71 (12.18)	71.19 (13.92)	73.13 (13.27)	-.48 (9.02)	-2.42 (5.66)		
	S&T ITT	72.89 (12.18)	73.81 (7.98)	73.02 (8.71)	-.92 (12.00)	-.13 (10.23)	.23	.01
	S&T AD	72.10 (12.06)	73.4 (7.41)	72.23 (8.36)	-1.29 (10.76)	-.13 (5.40)	.20	.08

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. AD= Adequate Dose. S&T AD n=33.

Table 5: ITT and AD 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
PCL Total	ITT	2.76	(-2.88 to 8.40)	.331	.25	(-5.15 to 5.65)	.927	-1.47	(-4.30 to 1.36)	.303
	AD	-.711	(-6.19 to 4.77)	.795	3.42	(-2.38 to 9.21)	.242	.173	(-4.91 to .90)	.173
BAC-R Total	ITT	2.99	(-10.97 to 16.96)	.668	-6.10	(-19.77 to 7.57)	.372	4.60	(-1.68 to 10.88)	.148
	AD	2.08	(-11.99 to 16.15)	.767	-7.65	(-21.15 to 5.84)	.258	5.15	(-1.15 to 11.42)	.107
DERS Total	ITT	8.14	(-2.89 to 19.18)	.145	6.19	(-5.19 to 17.57)	.281	-2.60	(-8.22 to 3.02)	.359
	AD	8.41	(-2.82 to 19.63)	.139	9.45	(-2.03 to 20.93)	.105	-4.38	(-10.05 to 1.28)	.127
DES Total	ITT	-.41	(-6.40 to 5.58)	.892	1.93	(-4.47 to 8.34)	.548	-.92	(-4.13 to 2.28)	.566
	AD	8.41	(-2.82 to 19.63)	.139	9.45	(-2.03 to 20.93)	.105	-4.38	(-10.05 to 1.28)	.127
HADS Anxiety	ITT	-.35	(-2.00 to 1.29)	.687	1.02	(-.67 to 2.72)	.223	-.67	(-1.50 to .16)	.113
	AD	-.32	(-1.98 to 1.34)	.699	1.22	(-.53 to 2.97)	.169	-.76	(-1.62 to .10)	.083
HADS Depression	ITT	-.24	(-1.94 to 1.47)	.783	-1.64	(-3.39 to -.11)	.066	.95	(.11 to 1.79)	.027
	AD	-.43	(-2.21 to 1.35)	.628	-2.03	(-3.83 to -.23)	.028	1.14	(.27 to 2.00)	.011
CCS Total	ITT	3.05	(-1.50 to 7.59)	.185	.59	(-4.21 to 5.38)	.808	-.62	(-3.19 to 1.95)	.630
	AD	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; S&T= Survive & Thrive experimental intervention arm; ITT= Intent to Treat; AD= Adequate Dose (≥ 7 sessions). 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.

Table 6: Reliable and Clinical Change for AD S&T and control participants

Measure		RC/ CSC outcome categories at T2			
		No reliable change	Deteriorated	Improved	Recovered (CSC)
PCL Total	AD S&T	66.7 %	3%	30.3%	15.2%
	WL/TAU	73.5%	8.8%	17.6%	11.8%
BAC-R Total	AD S&T	68%	20%	12%	0%
	WL/TAU	67.9%	7.1%	25%	0%
DERS Total	AD S&T	69.7%	9.1%	21.2%	9.1%
	WL/TAU	76.5%	2.9%	20.6%	2.9%
DES	AD S&T	84.8%	3%	12.1%	6%
	WL/TAU	88.2%	8.8%	2.9%	2.9%
HADS Anxiety	AD S&T	57.6%	6.1%	36.4%	6.1%
	WL/TAU	97.1%	0%	2.9%	0%
HADS Depression	AD S&T	81.8%	18.2%	0%	0%
	WL/TAU	76.5%	11.8%	11.8%	0%
CCS Total	AD S&T	90.9%	6.1%	3%	0%
	WL/TAU	94.1%	2.9%	2.9%	0%

Notes: AD S&T= Adequate Dose Survive & Thrive experimental intervention arm. [1]AD S&T: N= 33; Control: N= 34