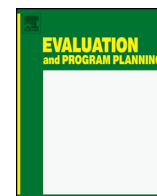




ELSEVIER

Contents lists available at ScienceDirect

Evaluation and Program Planning

journal homepage: www.elsevier.com/locate/evalprogplan

Wilderness adventure therapy effects on the mental health of youth participants

Daniel J. Bowen^{a,*}, James T. Neill^a, Simon J.R. Crisp^b^a Centre for Applied Psychology, Locked Bag 1, University of Canberra, ACT 2601, Australia^b Neo Psychology, 439a Riversdale Road, Hawthorn East, Victoria, 3123, Australia

ARTICLE INFO

Article history:

Received 23 June 2015

Received in revised form 30 April 2016

Accepted 8 May 2016

Available online 13 May 2016

Keywords:

Adventure therapy

Mental health

Youth

Program evaluation

Youth development

ABSTRACT

Adventure therapy offers a prevention, early intervention, and treatment modality for people with behavioural, psychological, and psychosocial issues. It can appeal to youth-at-risk who are often less responsive to traditional psychotherapeutic interventions. This study evaluated Wilderness Adventure Therapy (WAT) outcomes based on participants' pre-program, post-program, and follow-up responses to self-report questionnaires. The sample consisted of 36 adolescent out-patients with mixed mental health issues who completed a 10-week, manualised WAT intervention. The overall short-term standardised mean effect size was small, positive, and statistically significant (0.26), with moderate, statistically significant improvements in psychological resilience and social self-esteem. Total short-term effects were within age-based adventure therapy meta-analytic benchmark 90% confidence intervals, except for the change in suicidality which was lower than the comparable benchmark. The short-term changes were retained at the three-month follow-up, except for family functioning (significant reduction) and suicidality (significant improvement). For participants in clinical ranges pre-program, there was a large, statistically significant reduction in depressive symptomology, and large to very large, statistically significant improvements in behavioural and emotional functioning. These changes were retained at the three-month follow-up. These findings indicate that WAT is as effective as traditional psychotherapy techniques for clinically symptomatic people. Future research utilising a comparison or wait-list control group, multiple sources of data, and a larger sample, could help to qualify and extend these findings.

© 2016 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Engagement and treatment of adolescents is a key challenge for mental health clinicians. Mental health disorders affect a greater proportion of young people than other age groups (Department of Health and Ageing, 2013; World Health Organization, 2001), with 26% of 16–24 years old in Australia experiencing a personal mental illness in the previous 12 months (Australia Bureau of Statistics [ABS], 2007). Teenagers are susceptible to vulnerabilities arising from social and cultural changes, including unstructured home environments, growth in single-parent families, and media saturated with sex, violence, and pleasure-seeking (Barrett & Ollendick, 2004; Dowd, Singer, & Wilson, 2006; Villani, 2001). Adolescents face a myriad of other challenges including learning

problems, disengagement from education, family issues, homelessness, delinquency, substance abuse, and unemployment (Perkins & Borden, 2003).

There are a wide range of treatment options which aim to minimise the duration, severity, and recurrence of mental disorders or, at least, to maximise the duration between episodes (Mrazek & Haggerty, 1994). Meta-analytic reviews of psychotherapeutic treatments for youth have found a range of short-term effects including:

- psychotherapy for psychopathology (0.34; Weisz, McCarty, & Valeri, 2006)
- cognitive behavioural therapy for anxiety (0.98; James, Cowdrey, Soler, & Choke, 2013), depression (0.53; Klein, Jacobs, & Reinecke, 2007), anger (0.67; Sukhodolsky, Kassinove, & Gorman, 2004), violent behaviour (0.09; Özabacı, 2011), antisocial behaviour (0.48; Bennett & Gibbons, 2000), and criminal offending (0.84; Landenberger & Lipsey, 2005); pharmacotherapy for depression (0.19; Michael & Crowley, 2002) and anxiety (0.69; Ipson, Stein, Hawkrigde, & Hoppe, 2010)

* Corresponding author.

E-mail addresses: daniel@danielbowen.com.au (D.J. Bowen), james.neill@canberra.edu.au (J.T. Neill), simon@neopsychology.com.au (S.J.R. Crisp).

- multi-systemic therapy for antisocial behaviour and psychiatric symptoms (0.55; [Curtis, Ronan, & Borduin, 2004](#))
- rational emotive therapy for anxiety (0.48), disruptive behaviours (1.15), and self-concept (0.38; [Gonzalez et al., 2004](#))
- family-based crime prevention programs for offending outcomes (0.22), delinquency outcomes (0.32), and antisocial behaviour outcomes (0.20; [Farrington & Welsh, 2003](#))
- residential treatment programs for internalising problem behaviour (0.45) and externalising problem behaviour (0.60; [Knorth, Harder, Zandberg, & Kendrick, 2008](#))
- art therapy for anxiety (0.49; [Campbell, 2010](#)) and music therapy for mental illness (0.53; [Gitman, 2009](#)) and psychopathology (0.61; [Gold, Voracek, & Wigram, 2004](#)).

Although there are a wide range of treatment options, most interventions for teenagers are based on approaches which were originally developed for adults ([Crisp, 1997; Rutter et al., 2008](#)). However, adolescents' needs differ substantially from adults' needs ([Rutter & Taylor, 2005](#)). Teenagers who have mental health issues are often hesitant to seek help and can be difficult to engage in traditional treatment modalities ([Rickwood, Deane, & Wilson, 2007](#)). Therapeutic approaches for teenagers should be developed with an understanding of adolescent developmental needs in order to address mental health problems in a way that decreases stigma and promotes growth in fundamental domains of competency and performance, responsibility, judgement, social orientation, motivation, and identity ([Crisp, 2014](#)). Teenagers can benefit from modulated exposure to risk, stress, and precipitants of mental health problems, thereby increasing resiliency and developing protective strategies against future problems ([Crisp, 1997; Rutter et al., 2008](#)). Adventure therapy is one such option to address at-risk teenagers' mental health problems ([Schell, Cotton, & Luxmoore, 2012; Tucker, Javorski, Tracy, & Beale, 2012](#)).

1.1. Adventure therapy

Adventure therapy uses experiential learning activities in outdoor environments for assessment and intervention at an individual and group level, in order to effect psychological and/or behavioural therapeutic change ([Gass, Gillis, & Russell, 2012; Norton, Carpenter, & Pryor, 2015](#)). Adventure therapy utilises an eclectic therapeutic approach, drawing on aspects of cognitive-behavioural, systemic, existential, psychodynamic, and occupational therapy ([Association for Experiential Education, 2012](#)). Adventure therapy may be used as a form of brief intervention or embedded in a broader case management approach. Adventure therapy empowers participants by providing fun and engaging activities that involve real obstacles which, although often appearing to be impossible to overcome, are attainable. Activities are sequenced for success in order to provide participants with a sense of self-efficacy and mastery. Adventure therapy activities can include problem-solving activities, ropes challenge courses, outdoor adventure activities (such as rock climbing, abseiling, rafting, caving, and bushwalking), and extended overnight expeditions involving backpacking, canoeing and rafting, skitouring and/or snow camping. Meanings derived from participating in adventure therapy programs are intended to be incorporated back into the participant's individual and social world.

Although a relatively small field, research to date has shown adventure therapy to be effective in treating a range of behavioural and mental health problems. A recent meta-analysis of 197 studies of adventure therapy program outcomes ([Bowen & Neill, 2013a](#)) found that adventure therapy is moderately effective in facilitating positive short-term change in psychological, behavioural, emotional, and interpersonal domains and that these changes appear to be maintained in the longer term. The overall short-term

adventure therapy standardised mean effect size (ES) was moderately positive and statistically significant ($g = 0.47, p < 0.05$), and larger than for alternative treatment (0.14) and no treatment (0.08) comparison groups. In this meta-analysis, adventure therapy participants also reported small, positive, and not statistically significant change during the lead-up period (0.09) and maintenance of the moderately positive short-term improvements during the follow-up period (0.03). Of the eight major outcome categories, clinical outcomes (e.g., anxiety, depression, emotional stability, locus of control, resilience) had the highest ES (0.50). The only significant moderator was participant age. An ES of approximately 0.5 is recommended as a benchmark for adventure therapy programs, ESs between 0.3 and 0.5 are more typical of programs for 9–17 year olds and ESs between 0.5 and 0.7 are more typical of participants aged 18 years and over (see [Bowen & Neill, 2013b](#) for specific outcomes by age). Despite these promising results, there is a lack of well-established adventure therapy models that are supported by rigorous evaluations of their effectiveness ([Newes, 2000](#)).

1.2. Wilderness adventure therapy

Wilderness Adventure Therapy[®] (WAT) is a clinical psychology treatment model developed by Simon Crisp ([Crisp, 1997, 1998; Crisp, Noblet, & Hinch, 2004](#)). The WAT model is based on principles of best practice service design and includes a systemic framework and theoretical paradigm, client psychological assessment, intake processes and treatment planning, group composition, psychological safety procedures, therapeutic group procedures, monitoring of client outcomes, therapist skill training, management of ethical issues, and research evaluation ([Crisp, 1997, 1998](#)). The WAT model emphasises development of social-emotional competencies and coping skills through group-based adventure experiences that are facilitated by a psychologist. The WAT model has undergone three phases of development ([Crisp & Hinch, 2004](#)):

1. 1992–2000: WAT was designed as a multi-systemic, group therapy intervention to treat adolescents with severe psychological, behavioural and psychiatric problems ([Crisp & O'Donnell, 1998](#)). It was developed in 1992 as a component of a clinical adolescent day-program at the Austin Health, Child and Adolescent Mental Health Service, Melbourne, Victoria, Australia (CAMHS; [Crisp, O'Donnell, Kingston, Poot, & Thomas, 2000](#)). Between 1992 and 1996, 115 adolescents completed the program ([Kingston, Poot, & Thomas, 1997](#)).
2. 2000–2001: A WAT program was provided as a “stand-alone” out-patient treatment at the Barwon Health Adolescent Mental Health Service, Geelong, Victoria, Australia ([Crisp, 2003](#)) targeting a broad range of adolescent out-patients and their parents who required more intensive out-patient treatment than traditional approaches ([Crisp et al., 2004](#)). Six programs during 2000 and 2001 were conducted, with 39 adolescents completing the WAT treatment.
3. 2001–2003: WAT programs were implemented at the Inner East Community Health Service, Melbourne, Victoria, Australia, to work with local government high schools and community youth and family agencies and to provide early intervention, prevention and treat psychological, behavioural, and family-based problems in adolescents before they required referral to a clinical service ([Crisp et al., 2004](#)). Nine WAT programs were implemented in this phase, with 36 adolescents completing the WAT treatment: two early intervention programs involved Year 9 and 10 students at risk of educational non-completion or school failure, one program was for families with one or more adolescent child who had experienced substantial domestic

violence that significantly affected the adolescent and family's ability to function, one program was for female adolescents who had experienced sexual assault, one program was for female adolescents who had significant body image issues or who were in the early stages of an eating disorder, two programs were conducted for families with adolescents in the early stages of known or suspected substance abuse, and two programs involved families with adolescents who had evidence of significant disruption to parent-adolescent attachment (Crisp et al., 2004). Since 2003, other WAT programs have been conducted in independent and government schools as community and early-intervention programs.

The WAT intervention is a 10-week, manualised, part-time program, which is typically facilitated by three WAT practitioners for six to eight participants. The manual (Crisp et al., 2004) includes protocol about how to implement the program for different target groups, including early intervention at-risk clients, community counselling where clients have already sought help and WAT is part of on-going case management, and clinical treatment that may be part of a multi-pronged approach within a comprehensive range of clinical services.

The WAT program has four components (Crisp et al., 2004). Intake (Week 1) includes screening, assessment, engagement, orientation, and negotiation of client goals. Treatment (Weeks 2–9) involves seven day-based adventure activities (e.g., bushwalking, abseiling, cross country skiing, and white water rafting), a two-day overnight training expedition, and a five-day expedition. Parents, teachers, and support workers also participate in up to eight weekly indoor adventurous problem-solving activities incorporated within group therapy sessions. Termination (Week 10) includes a review of goals and unresolved needs/issues, identification of post-treatment goals and strategies, and enlisting of psycho-social supports. Follow-up includes liaison with other agencies, group reunion, and school or placement outreach follow-up. A sample program structure is included in Table 1.

There have been several previous evaluations of WAT programs (Crisp, 2003; Crisp & Auger, 1998; Crisp & Hinch, 2004; Crisp & Noblet, 2001; Crisp et al., 2004; Crisp & O'Donnell, 1998; Kingston et al., 1997). An evaluation of 101 clients during Phase 1 found statistically significant improvements in self-reported difficulties with social, attentional, and attitudinal problems, self-esteem, and task leadership skills (Kingston et al., 1997). There were non-significant improvements in sociability, appearance, family, school, emotions, parents, and crisis problems, achievement motivation, emotional control, social competence, time management, withdrawal, anxiety/depression, use of some coping strategies (e.g., self-blame, social action, ignoring, keeping problems secret, and relaxation), social supports, physical recreation, and hard work (Kingston et al., 1997). In addition, there were non-significant increases in not coping and somatic problems.

A case study of a 15-year-old girl (Susan) with psychological and social problems, who completed WAT during Phase 1, was provided by Crisp and Auger (1998). Upon program completion, Susan showed increased self-confidence, personal insight, assertiveness, sociality, and clarity of future plans. At the two-week follow-up, Susan had returned to mainstream schooling with perfect attendance and psychometric evaluation showed an increase in coping skills, a reduction in reported emotional problems, and a slight increase in self-esteem. At the six-month follow-up, Susan was continuing to attend school full-time and was involved in a range of extra-curricular activities. Crisp and O'Donnell (1998) presented three other case vignettes of participants who completed WAT programs. Reported outcomes included an increased sense of hope, perseverance, problem-solving, self-sufficiency, willingness to access support, and returning to school.

Table 1
Example WAT program structure.

Week	Activity
–4	Appraisal of referrals
–3	Contacts parents
–2	Parent and student information session Student interviews
–1	Teacher information session
1	Student introduction group Teacher and parent introduction session Day activity: Peak ascent
2	Student group Student, teacher and parent session Day activity: Cross-country skiing
3	Student group Overnight expedition
4	Student group Teacher and parent session Day activity: Rock climbing and abseiling
5	Student group Parent session Day activity: Caving
6	Student group Parent session Day activity: Initiative course
7	Student group Parent session Day activity: Preparation for major expedition
8	Major Expedition: Coastal wilderness hike
9	Student group Parent, student and teacher day: Rafting
10	Student group Teacher and parent session Graduation
11+	Handover to Teachers/Review Reunion/Follow-up

An evaluation of WAT during Phase 1 ($n = 75$) found that post-treatment levels of symptoms were reduced but the differences were not statistically significant. However, at a five-year follow-up mental health symptoms and the frequency of non-productive coping had significantly reduced.

Outcomes for 36 WAT Phase 2 participants were evaluated by Crisp and Hinch (2004). Overall results indicated a statistically and clinically significant reduction in mental illness symptoms (from clinical levels to non-clinical levels following the WAT intervention) which remained at three-month and two-year follow-ups. Furthermore, there were statistically significant improvements in general and social self-esteem following treatment, with non-significant improvements in life-threatening attitudes, productive and non-productive coping, social competence, school functioning, psychological resilience, and family relationships.

Crisp and Noblet (2001) evaluated a Phase 3 WAT program for eight clients who were at risk of school failure and experiencing mental health problems. Results indicated clinically significant improvements in mental health, behaviour problems, and life-threatening attitudes (i.e., future high-risk behaviour). Non-significant improvements were evident for expectation of receiving help and increased connectedness to, and trust in, peers, communication, and overall functioning.

Program evaluation results for 36 clients from six WAT Phase 3 programs who were identified as showing risk factors including school failure, poor body image and eating problems, substance abuse, being victims of sexual abuse or assault, and/or family dysfunction indicated a statistically and clinically significant reduction in mental illness symptoms which remained at follow-up (3 months later; Crisp et al., 2004). Furthermore, there were statistically significant improvements in general and social self-esteem following treatment, and school self-esteem showed a delayed statistically significant improvement at follow-up. There

were also non-statistically significant improvements in life-threatening attitudes, productive and non-productive coping, social competence and school functioning, psychological resilience, and family relationships.

Therapeutic programs for youth that utilise innovative and non-traditional approaches often do so in isolation and with limited knowledge about how to maximise their effect. A critical task for adventure therapy program developers, and for advancing the adventure therapy movement as a whole, is the use of high quality program design along with research and evaluation. Although WAT Phase 1–3 programs have been evaluated, the data should be independently analysed by non-program staff, expressed as effect sizes and benchmarked against adventure therapy meta-analytic findings, and formally peer-reviewed. Due to the varying target groups and level of symptom severity of WAT Phase 1 and Phase 3 programs, as well as the small samples sizes, it was impractical to include data from these programs. Therefore, this study builds on previous WAT research by evaluating the participant outcomes for WAT Phase 2 programs using standardised mean effect sizes and compares the outcomes to adventure therapy meta-analytic findings. It was hypothesised that the WAT treatment would be associated with statistically significant short-term improvements in psychological and behavioural symptomatology, that these changes would be maintained at a three month follow-up, and that the amount of change would be at least equivalent to comparable adventure therapy benchmarks.

2. Method

2.1. Participants

There were 36 participants (21 females and 15 males) who were out-patients of a regional state-based adolescent mental health service in Victoria, Australia (Barwon Health) who required more intensive out-patient treatment than traditional approaches. Participants were aged between 12 and 18 years ($M=14.6$; $SD=1.6$). Participants completed one of six standardised WAT interventions during 2000 and 2001. The only other treatment option of equal or greater intensity was admission to a psychiatric in-patient ward (Crisp & Hinch, 2004).

The most common primary reason for referral to the program was conduct/behaviour (29%), followed by depression (21%), anxiety (14%), identity/self-esteem (14%), relationship problems (14%), and psychosis (8%). On average, there were six additional presenting problems per participant, including low-esteem (71%), poor relationships with adults (65%), family (65%) and peers (47%), anxiety (41%), family mental illness (41%), and depression (35%).

2.2. Materials

Six self-report questionnaires, consisting of 226 closed-ended items, were completed by participants on up to three occasions: pre-program (Time 1), post-program (Time 2), and at a three-month follow-up (Time 3).

2.2.1. Resilience Questionnaire

The Resilience Questionnaire (RQ) is designed to measure psychological resilience (Crisp, 2001). It consists of 14 items. Four items are rated on a dichotomous scale (True (1 point) or False (2 points); e.g., "I can change things in my life if I really try."). The other 10 items are measured on a seven-point Likert scale. For example, the item "How close, or connected to any person your own age do you feel?" is rated from 1 (Very Close/Connected) to 7 (Not Very Close/Connected). A total score was calculated by adding the scores from the true-false and Likert scale items, ranging from

14 to 78. Higher scores signify lower resilience. For the current study, the internal consistency was strong (Cronbach's $\alpha = 0.83$).

2.2.2. Beck Depression Inventory-II

The Beck Depression Inventory-II (BDI-II; Beck, Steer, & Brown, 1996) assesses severity of depressive symptomatology. There are 21 items measured using a four point ordinal scale (0=I Do Not Feel Sad; 1=I Feel Sad Much of the Time; 2=I Am Sad All the Time; 3=I Am So Sad or Unhappy That I Can't Stand It). Items are summed to form a total score that ranges from 0 to 63, with higher scores indicating greater depressive symptomatology (Minimal=0–13; Mild=14–19; Moderate=20–28; Severe=29–63). The BDI-II has high internal consistency amongst a variety of populations ($\alpha=0.91$ – 0.93) and the one-week test-retest reliability is strong ($r=0.93$; Antony & Barlow, 2010). Convergent validity is supported by significant correlations with other indices of depression, including the Beck Depression Inventory (0.93), the Hamilton Rating Scale for Depression (0.71), and the 7-item Depression subscale of the Depression Anxiety Stress Scales (0.76; Antony & Barlow, 2010).

2.2.3. Youth Self-Report

The Youth Self-Report (YSR; Achenbach, 1991) assesses participants' level of behavioural and emotional functioning. There are 112 items measured on a three-point ordinal scale (0=Not True, 1=Somewhat or Sometimes True, 2=Very True or Often True). The YSR is comprised of eight core syndrome scales: Withdrawn, Somatic Complaints, Anxious/Depressed, Social Problems, Thought Problems, Attention Problems, Aggressive Behaviour, and Delinquent Behaviour. The YSR also includes two higher order scales: Internalising (Withdrawn, Somatic Complaints, Anxious/Depressed) and Externalising (Aggressive Behaviour, Delinquent Behaviour) problems, and a Total score (includes all 8 core syndrome scales; Achenbach & Rescorla, 2001). Raw scores are converted to a *T*-score, with higher scores reflecting poorer behavioural and emotional functioning. *T*-scores 67 or higher for the syndrome scales and 60 or higher for the higher order scales are clinically significant (Achenbach, 1991). The YSR core syndrome scales have reasonable internal consistency amongst a variety of populations ($\alpha=0.71$ – 0.89). One-week test-retest coefficients are generally good, with all scales being above 0.70, except for Withdrawn/Depressed. Seven-month test-retest correlations are generally in the 0.50 range (Achenbach & Rescorla, 2001).

2.2.4. Coopersmith Self Esteem Inventory

The Coopersmith Self Esteem Inventory (CSEI; Coopersmith, 1981) measures self-esteem. The CSEI includes 50 items which represent four subscales: General (26 items; e.g., "I am pretty sure of myself"), Social (8 items; e.g., "I am easy to like"), Home/Parents (8 items; e.g., "My family understands me"), and School/Academic (8 items; e.g., "I am proud of my school work"). For each item, the dichotomous response scale is "Unlike Me" (0 points) or "Like Me" (1 point). The sum of all items provides an estimate of global self-esteem, ranging from 0 to 50. Higher scores signify greater self-esteem. The CSEI has reasonable internal consistency amongst a variety of populations ($KR-20=0.80$ – 0.92); Coopersmith, 1981).

2.2.5. CORE Family Functioning Questionnaire

The CORE Family Functioning Questionnaire (Author Unknown, 2001) is designed to assess family functioning. The CORE FFQ consists of five items measured on a five-point Likert scale ranging from 1 (Very Well) to 5 (Poor). A sample item is "How well has your family communicated together over the past month?". Total scores can range from 5 to 25, with higher scores reflecting poorer family functioning. For the current study, the internal consistency was 0.92.

2.2.6. Life Attitudes Schedule – Short Form

The Life Attitudes Schedule – Short Form (LAS-SF; Rohde, Lewinsohn, Seeley, & Langhinrichsen-Rohling, 1996) measures adolescents' suicide-proneness. The dichotomous response options for the 24 items (e.g., "I avoid unnecessary risks") were False (0) or True (1). The 12 positively worded items were reverse-scored and summed with the 12 negatively worded items to create a total score that can range from 0 to 24, with higher scores indicating greater risk of suicidality. The LAS-SF has reasonable internal consistency ($KR-20=0.80$), and the 30-day test-retest reliability was moderate (0.73; Rohde et al., 2003).

2.3. Procedure

The questionnaires were administered in a standardised manner on a one-to-one basis prior to the first program session (Time 1), following the final program session (Time 2), and three months after the completion of the program (Time 3). Additional assistance and/or verbal administration were provided when required (e.g., poor attention or literacy skills). On average, it took 45–60 min for participants to complete the six questionnaires on each occasion.

2.4. Analysis

Thirty six participants provided pre-program (Time 1; T1), post-program (Time 2; T2), and follow-up (Time 3; T3) data. However, 24% of the data values were missing due to participant non-completion of some of the survey questions, with lower completion rates for measures collected in the latter half of the test battery. Thus, multiple imputation was used to replace the missing data. Multiple imputation uses a regression-based procedure to generate multiple copies of the data set, each of which contains different estimates of the missing values (Enders, 2010). Based on Graham's (2012) recommendation, the Markov-Chain Monte Carlo (MCMC) method in SPSS 22 was used to generate 40 imputed data sets, with 50 iterations of MCMC between each imputation. Methodologists regard multiple imputation as a "state of the art" missing data technique because it improves the accuracy and the power of the analyses relative to other missing data handling methods (Schafer & Graham, 2002).

Scale means were reflected, where necessary, so that increased scores over time signified improvement. Short-term (T1–T2) and follow-up (T2–T3) changes were investigated using descriptive statistics and standardised mean ESs (Hedges' g). A commonly referred to rule of thumb for interpreting standardised mean ESs is

0.20 (small), 0.50 (medium), and 0.80 (large; Cohen, 1988). Ninety percent confidence intervals (CIs) for ESs are also reported. If an ES CI excludes zero, then the ES is statistically significant (Ellis, 2010). However, ES CIs should be interpreted with caution as power for a two-tailed test was estimated to be 47%, based on the overall average ES (0.26), a sample size of 36, and a 90% confidence level. ESs are also expressed as estimated percentages of participants who improved, using the Binomial Effect Size Display (BESD; Randolph & Edmondson, 2005) and Cohen's (1988) U_3 . Using the BESD, an effect size of 0.2 is equivalent to a 10% increase in the outcome of interest, whilst an effect size of 0.4 is equivalent to a 20% increase in the outcome of interest. Using Cohen's U_3 , an effect size of 0.2 is equivalent to 58% of participants who received treatment being better off than an equivalent group who did not participate, whilst an effect size of 0.4 means that 66% of participants who received treatment are likely to be better off than an equivalent group who did not participate.

Participants were an out-patient group of adolescents with diverse reasons for referral, thus it was important to consider the outcomes in clinical terms, that is, the presence or absence of pathology (Kazdin, 1999). Additional analyses (Glasser, 2014) were conducted for the two questionnaires (BDI-II and YSR) in which clinical cut-offs were available. Short-term and follow-up changes for participants who reported clinical and non-clinical levels of depression and behavioural and emotional dysfunction pre-program were investigated using descriptive statistics and standardised mean ESs. Short-term and follow-up descriptive statistics and ESs were also calculated for the most severe symptom for each participant, as indicated by the most elevated subscale on the YSR for each participant before treatment.

3. Results

3.1. Overall

Descriptive statistics for participant responses to each of the six questionnaires at T1–T3 are presented in Tables 2–4, along with standardised mean ESs with CIs for short-term changes (T1–T2) and during the follow-up period (T2–T3). Overall and questionnaire total results are shown in Table 2. Subscale results for the CSEI (self-esteem) are shown in Table 3 and the YSR (behavioural and emotional functioning) subscale results for participants with clinical and non-clinical range pre-program scores are shown in Table 4. Tables 2 and 3 also show short-term ES adventure therapy benchmarks for 10–17 years olds from Bowen and Neill's (2013b) meta-analysis. Statistically significant short-term ESs and meta-

Table 2

Descriptive statistics for T1–T3 questionnaire totals and overall, with effect sizes, 90% confidence intervals for WAT participants, and comparative benchmarks ($N=36$).

Constructs	T1 M (SD)	T2 M (SD)	T3 M (SD)	ST _{ES}	ST _{CI}	FU _{ES}	FU _{CI}	ST BM _{ES}	ST BM _{CI}	BM Category
Behavioural & Emotional Functioning	36.51 (7.92)	39.21 (7.09)	39.44 (7.14)	.36	-.03: .75	.03	-.36: .42	.46*	.37: .54	Clinical
Depression	42.33 (15.26)	48.63 (12.04)	48.61 (11.30)	.46*	.07: .85	.00	-.39: .39	.46*	.37: .54	Clinical
Family Functioning	13.25 (4.19)	13.75 (4.01)	11.48 (3.85)	.12	-.27: .51	-.58*	-.97: -.18	.31*	.14: .48	Family Development
Resilience	33.53 (12.09)	38.94 (9.80)	36.00 (9.51)	.49*	.10: .89	-.30	-.69: .09	.46*	.37: .54	Clinical
Self-Esteem	48.58 (20.71)	53.51 (17.76)	54.47 (15.50)	.26	-.13: .64	.06	-.33: .45	.41*	.34: .47	Self-Concept
Suicidal-Proneness	15.85 (3.95)	15.57 (5.28)	17.52 (3.77)	-.06	-.45: .33	.43*	.03: .82	.46*	.37: .54	Clinical
Overall				.26*	.12: .41	-.06	-.28: .16	.44*	.38: .50	Total

Note: ST = Short-Term; FU = Follow-Up; ES = Effect Size (Hedges' g); CI = Confidence Interval; BM = Benchmark (10–17 year olds; obtained from Bowen and Neill, 2013b). * indicates ES is statistically significant. An increase in mean over time signifies improvement.

Table 3
Descriptive statistics for T1, T2, and T3 self-esteem (CSEI) subscales, with effect sizes, 90% confidence intervals for WAT participants, and comparative benchmarks (N = 36).

Self-Esteem Subscales	T1 M (SD)	T2 M (SD)	T3 M (SD)	ST _{ES}	ST _{CI}	FU _{ES}	FU _{CI}	ST BM _{ES}	ST BM _{CI}	BM Used
General	12.61 (6.07)	13.80 (5.14)	15.70 (4.47)	.21	-.18: .60	.39*	.00: .79	.41*	.34: .47	Self-Concept
Home/Parents	3.50 (2.34)	4.29 (2.27)	3.79 (2.24)	.34	-.05: .73	-.22	-.61: .17	.31*	.14: .48	Family Development
School/Academic	3.78 (1.95)	3.46 (1.66)	3.59 (1.52)	-.18	-.57: .21	.08	-.31: .47	.44*	.34: .54	Academic Perception
Social	4.42 (6.07)	5.18 (1.93)	5.13(1.75)	.40*	.01: .79	-.03	-.41: .36	.41*	.32: .49	Social Development
Total	48.58(20.71)	53.51(17.76)	54.47(15.50)	.26	-.13: .64	.06	-.33: .45	.41*	.34: .47	Self-Concept

* indicates ES is statistically significant.

Table 4
Descriptive statistics for T1, T2, and T3 behavioural and emotional (YSR) subscales, with effect sizes, and 90% confidence intervals for WAT participants in the clinical and non-clinical range before treatment.

Behavioural & Emotional Subscales	N	T1 M (SD)	T2 M (SD)	T3 M (SD)	ST _{ES}	ST _{CI}	FU _{ES}	FU _{CI}
Clinical Range Participants								
Internalising	26	31.69 [#] (6.37)	37.46 (7.64)	38.34 (7.45)	0.82*	0.35: 1.30	.12	-0.34: 0.57
Somatic Complaints	11	23.96 [#] (9.45)	35.51 (7.46)	34.81 (10.76)	1.36*	0.58: 2.13	-.08	-0.78: 0.63
Anxious/Depressed	23	25.00 [#] (8.62)	32.86 (10.13)	35.25 (9.88)	0.84*	0.33: 1.34	.24	-0.25: 0.73
Withdrawn	7	29.31 [#] (3.02)	34.82 (12.87)	37.87 (7.22)	0.59	-0.31: 1.49	.29	-0.59: 1.18
Externalising	21	30.13 [#] (7.60)	33.41 [#] (8.17)	34.44 (8.88)	0.42	-0.10: 0.93	.12	-0.39: 0.63
Social Problems	18	28.10 [#] (5.09)	34.09 (11.03)	36.22 (7.61)	0.70*	0.13: 1.26	.22	-0.33: 0.77
Attention Problems	16	22.50 [#] (9.92)	29.95 [#] (14.92)	32.29 [#] (8.90)	0.59	-0.01: 1.18	.19	-0.39: 0.77
Thought Problems	8	27.37 [#] (8.89)	40.19 (5.57)	34.06 (14.42)	1.73*	0.76: 2.69	-.56	-1.40: 0.28
Aggressive Behaviour	10	24.64 [#] (7.94)	31.51 [#] (8.77)	29.77 [#] (10.27)	0.82*	0.06: 1.59	-.18	-0.92: 0.55
Delinquent Behaviour	15	24.86 [#] (7.22)	28.00 [#] (9.06)	32.79 [#] (10.79)	0.38	-0.22: 0.99	.48	-0.13: 1.09
Total	26	32.18 [#] (5.80)	36.54 (6.56)	36.66 (6.74)	0.70*	0.23: 1.17	.02	-0.44: 0.47
Non-Clinical Range Participants								
Internalising	10	45.81 (3.21)	44.32 (4.87)	44.42 (4.00)	-0.36	-1.10: 0.38	.02	-0.71: 0.76
Somatic Complaints	25	42.78 (5.17)	43.66 (6.01)	42.21 (5.73)	0.16	-0.31: 0.62	-.25	-0.71: 0.22
Anxious/Depressed	13	45.07 (4.35)	44.10 (6.33)	44.01 (5.86)	-0.18	-0.83: 0.47	-.01	-0.66: 0.63
Withdrawn	29	41.68 (5.60)	42.35 (5.64)	43.57 (4.93)	0.12	-0.31: 0.55	.23	-0.20: 0.66
Externalising	15	46.03 (2.64)	44.68 (5.41)	44.00 (5.13)	-0.32	-0.92: 0.29	-.13	-0.73: 0.47
Social Problems	18	41.52 (6.12)	41.87 (8.07)	40.22 (8.76)	0.05	-0.50: 0.60	-.20	-0.75: 0.35
Attention Problems	20	42.74 (4.57)	42.30 (6.36)	41.81 (6.96)	-0.08	-0.60: 0.44	-.07	-0.59: 0.45
Thought Problems	28	46.03 (5.13)	45.04 (5.93)	44.65 (6.01)	-0.18	-0.62: 0.26	-.07	-0.51: 0.37
Aggressive Behaviour	26	43.68 (5.46)	43.07 (7.02)	42.74 (6.38)	-0.10	-0.55: 0.36	-.05	-0.51: 0.41
Delinquent Behaviour	21	42.73 (4.57)	42.53 (6.36)	41.36 (6.96)	-0.04	-0.54: 0.47	-.18	-0.68: 0.33
Total	10	45.17 (2.91)	44.55 (4.67)	45.01 (3.97)	-0.16	-0.90: 0.58	.11	-0.63: 0.84

Note: # indicates mean in the clinical range (reflected T-score clinical range = 0– 33). * indicates ES is statistically significant.

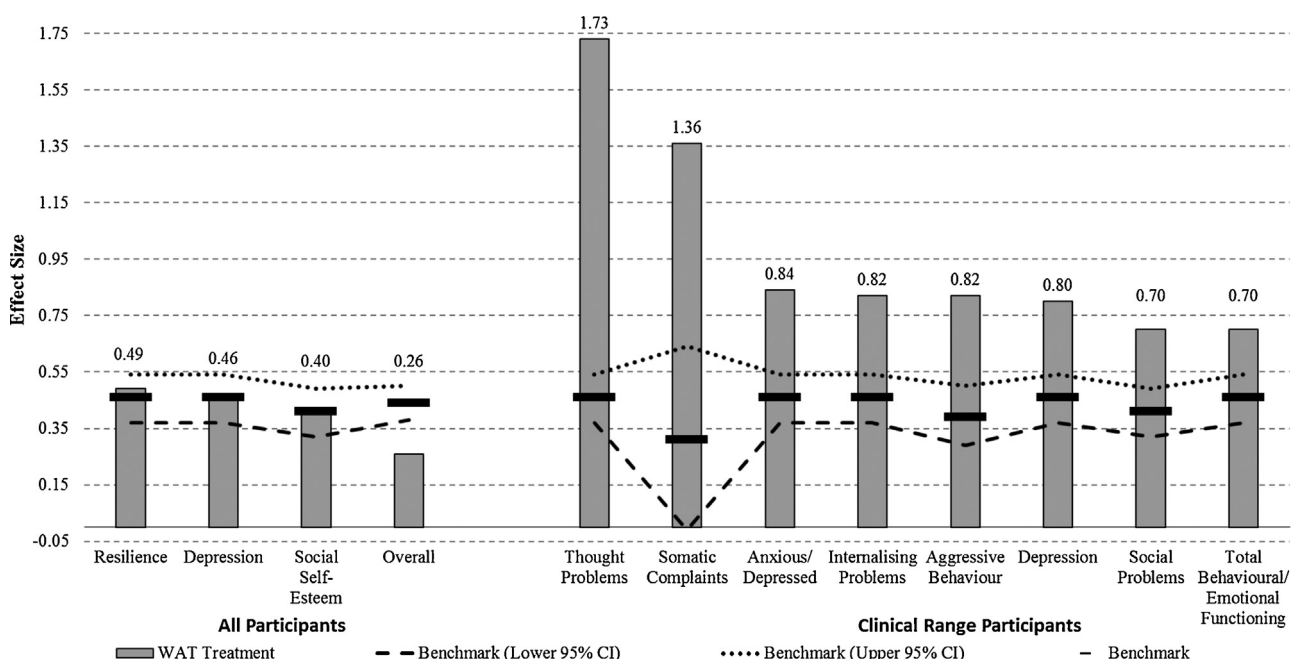


Fig. 1. Statistically significant short-term effect sizes and meta-analytic benchmarks for all participants and clinical range participants.

analytic benchmarks for all participants and clinical range participants are presented in Fig. 1.

The overall short-term ES for WAT program participants was small, positive, and statistically significant (0.26), representing a 13% overall improvement and is akin to 60% of participants reporting improvements at the end of the program. The total short-term ESs were statistically significant for two of the questionnaires ($p < 0.10$; RQ (resilience) $ES = 0.49$; BDI-II (depression) $ES = 0.46$) and non-significant for the other four questionnaires ($p > 0.10$; YSR (behavioural and emotional functioning) $ES = 0.36$; CSEI (self-esteem) $ES = 0.20$; CORE FFQ (family functioning) $ES = 0.12$; LAS-SF (Suicidal-proneness) $ES = -0.06$; see Table 2).

The overall follow-up period (T2–T3) ES was very small, negative, and not statistically significant (-0.06), representing a 3% loss, and thus indicating overall retention of the short-term gains. During the follow-up period, the total ES was statistically significant for one questionnaire (suicidal-proneness $ES = 0.43$) and not statistically significant for the other five measures (family functioning $ES = -0.58$; resilience $ES = -0.30$, depression $ES = 0.00$, behavioural and emotional functioning $ES = 0.03$, and self-esteem $ES = 0.06$; see Table 2). The follow-up result for suicide-proneness represents a 21% improvement and is akin to 67% of participants reporting improvements in suicidality during the three-month follow-up period.

3.2. Self-esteem

The results for the CSEI (self-esteem) subscales are reported in Table 3. The short-term ESs were statistically significant for one of the subscales (Social $ES = 0.40$) and non-significant for the other three subscales (Home/Parents $ES = 0.34$; General $ES = 0.21$; School/Academic $ES = -0.18$). During the follow-up period, the ES for one of the subscales was statistically significant (General $ES = 0.39$), while the ES for the other three subscales was non-significant (School/Academic $ES = 0.08$; Social $ES = -0.03$; Home/Parents $ES = -0.22$).

3.3. Depression

3.3.1. Clinical participants

Twenty-three participants reported clinical levels of depressive symptoms pre-program, as measured by the BDI-II (T1 $M = 34.03$, $SD = 12.84$; Severe range; see Fig. 2). For these participants, the short-term change was large, positive, and statistically significant ($ES = 0.80$; 37% change; $CI = 0.30$:1.31; T2 $M = 44.19$, $SD = 12.45$; Mild range). Change during the follow-up period was very small, positive, and not statistically significant (0.03; 1% change; $CI = -0.46$:0.51; T3 $M = 44.51$, $SD = 11.31$; Mild range), indicating

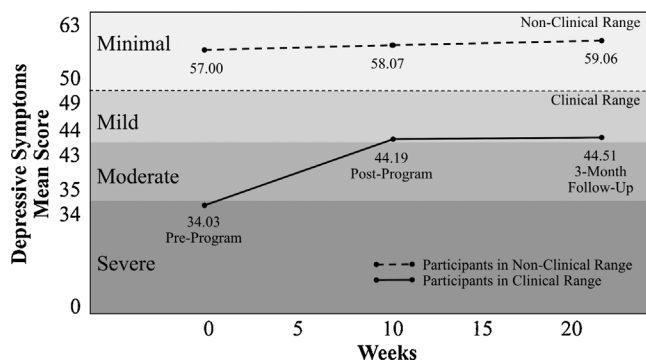


Fig. 2. Pre-program, post-program and 3-month follow-up depressive symptom (BDI-II) means for participants in the clinical ($N = 23$) and non-clinical ($N = 13$) range before treatment.

retention of the short-term gains. The average short-term effect is akin to 79% of clinical range participants reporting improvements in mental health at the end of the program and at follow-up.

3.3.2. Non-clinical participants

Thirteen participants reported non-clinical levels of depressive symptoms pre-program (T1 $M = 57.00$, $SD = 3.99$; Minimal range; see Fig. 2). For these participants, the short-term ES was small, positive, and not statistically significant ($ES = 0.26$; 13% change; $CI = -0.39$:0.91; T2 $M = 58.1$; Minimal range). During the follow-up period for these participants, the ES was small, positive, and not statistically significant ($ES = 0.22$; 11% change; $CI = -0.42$:0.87; T3 $M = 59.1$; Minimal range). Thus, for these non-clinical participants, there was small, positive and non-significant change from pre-program to post-program, and from post-program to follow-up.

3.4. Behavioural and emotional functioning

3.4.1. Clinical participants: total

Twenty six participants reported clinical levels of behavioural and emotional dysfunction pre-program, as measured by the YSR (see Table 4). For these participants, the overall short-term ES was large, positive, and statistically significant (0.70; 33% change). The follow-up period ES was very small, positive, and not statistically significant (0.02; 1% change), indicating retention of the short-term gains. The average overall short-term effect is akin to 76% of clinical range participants reporting improvements in mental health at the end of the program and at follow-up.

3.4.2. Non-clinical participants: total

Ten participants reported non-clinical levels of overall behavioural and emotional dysfunction pre-program (see Table 4). For these participants, the overall short-term ES was small, negative, and not statistically significant (-0.16 ; -8% change). During the follow-up period, the ES was small, positive, and not statistically significant (0.11; 4% change). Thus, for these non-clinical participants, there was minimal change from pre-program to post-program and from post-program to follow-up.

3.4.3. Clinical participants: sub-scales

The short-term ES for the 26 participants who reported clinical levels of Internalising problems was large, positive, and statistically significant (0.82; 38% change) and for the 21 participants who reported clinical levels of Externalising problems, the ES was moderate, positive, and not statistically significant (0.42; 21% change). During the follow-up period, the change in Internalising and Externalising problems was very small, positive, and not statistically significant ($ES = 0.12$; 6% change).

Analyses of changes for participants with clinical-range pre-program scores for each of the YSR subscales found large to very large statistically significant short-term ESs for five subscales (Thought problems, $ES = 1.73$, $n = 8$; Somatic Complaints $ES = 1.36$, $n = 11$; Anxious/Depressed $ES = 0.84$, $n = 23$; Aggressive Behaviour $ES = 0.82$, $n = 10$; Social Problems $ES = 0.70$, $n = 18$) and three of the subscales were non-significant (Attention Problems $ES = 0.59$, $n = 16$; Withdrawn $ES = 0.59$, $n = 7$; Delinquent Behaviour $ES = 0.38$, $n = 15$; see Fig. 1). The YSR subscale ESs during the follow-up period were non-significant and ranged from moderate and positive (Delinquent Behaviour, $ES = 0.48$; 23% change; $n = 16$) to moderate and negative (Thought Problems, $ES = -0.56$; 27% change; $n = 8$). Thus, there was clinically significant change for those within the clinical ranges for five out of eight YSR subscales (internalising problems, social problems, aggressive behaviour, anxiety/depression, somatic complaints and thought problems), which were retained at follow-up.

3.4.4. Non-clinical participants: sub-scales

The short-term ES for participants who reported non-clinical levels of Internalising and/or Externalising problems was moderate, negative, and not statistically significant (-0.36 ; 18% change; $n = 10$ and $ES = -0.32$; 16% change; $n = 15$ respectively). During the follow-up period, the change in Internalising problems was very small, positive, and not statistically significant ($ES = 0.02$; 1% change), whilst the change in Externalising problems was small, negative, and not statistically significant ($ES = -0.13$, a 1% change). The eight core syndrome subscale short-term ESs were non-significant and ranged from small and positive (Somatic Complaints, $ES = 0.16$; 8% change; $n = 25$) to small and negative (Anxious/Depressed and Thought Problems, $ES = -0.18$; 9% change; $n = 13$ and 28 respectively). The eight core syndrome subscale ESs during the follow-up period were non-significant and ranged from small and positive (Withdrawn, $ES = 0.23$; 11% change; $n = 29$) to small and negative (Somatic Complaints, $ES = -0.25$; 12% change; $n = 25$). Thus, for these non-clinical participants, there was minimal change from pre-program to post-program and from post-program to follow-up.

3.4.5. Severest symptoms

Short-term and follow-up ESs were also calculated for the most severe symptom for each participant, as indicated by the most elevated subscale on the YSR for each of the 36 participants before treatment (T1 $M = 24.58$, $SD = 12.34$; $N = 36$; clinical range; see Fig. 3). The short-term ES for the most severe symptoms was moderate, positive, and statistically significant ($ES = 0.57$; 27% change; $CI = 0.18:0.97$; T2 $M = 32.29$, $SD = 14.55$; clinical range). The follow-up ES was very small, negative, and not statistically significant ($ES = -0.02$; 1% change; $CI = -0.40:0.37$; T3 $M = 32.06$, $SD = 14.24$; clinical range), indicating retention of the short-term gains. The average short-term ES of 0.57 is akin to 72% of program participants reporting short-term improvements in the most problematic aspect of their behavioural and emotional functioning.

4. Discussion

This study examined the short- and longer-term changes in WAT participants' psychosocial, behavioural, and psychological functioning. These changes are interpreted here in the context of adventure therapy and psychotherapeutic meta-analytic outcome research. Lessons learned and limitations of the present study are also considered.

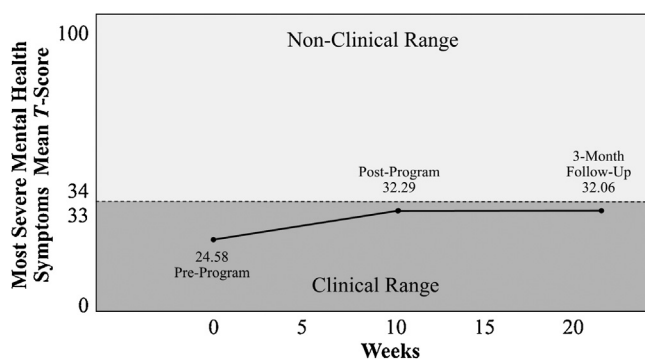


Fig. 3. Pre-program, post-program and follow-up YSR means for the most severe behavioural and emotional dysfunction for participants before treatment ($N = 36$).

4.1. Efficacy of WAT

4.1.1. Overall short-term effect

This study provides mixed evidence about the effectiveness of WAT as a psychotherapeutic treatment. Overall, there was a small, positive, statistically significant short-term impact on the measured outcomes ($ES = 0.26$). This result is similar to meta-analytic results for 10–17 year old adventure therapy participants (Bowen & Neill, 2013b), but is lower than findings from psychotherapy outcome research. There were moderate, statistically significant, short-term improvements in participants' psychological resilience and depression, moderate, non-statistically significant improvements in behavioural and emotional functioning, and small, non-statistically significant improvements in self-esteem and family functioning. There was little to no evidence of short-term improvements in suicidality. All the short-term effects were within the adventure therapy age-based meta-analytic benchmark CIs, except for suicidality which was lower than the comparable benchmark.

4.1.2. Overall follow-up effect

The overall follow-up ES was very small, negative, and not significant ($ES = -0.06$), indicating no substantial change during the three-month follow-up period. This finding is consistent with the very small, positive, non-statistically significant follow-up ES for adventure therapy ($ES = 0.03$) reported by Bowen and Neill (2013a). The follow-up effect was not significant for each of questionnaire totals and subscales, with the exception of family functioning, for which participants reported reduced functioning, and suicide-proneness, which decreased. As family functioning can be a significant factor in the development and maintenance of mental health disorders as well as social and school functioning (Greenberg & Lippold, 2013; Patel, Flisher, Hetrick, & McGorry, 2007), it may be an important consideration in the development of future WAT interventions to use a multi-family group format and/or explicitly aim to teach effective family functioning strategies as an overall treatment goal. As adolescent suicide is a significant public health concern (Sakinofsky et al., 2007), the statistically significant reduction in suicidality at follow-up is noteworthy. The delayed change in suicidality may reflect a process that begins with improvements in resilience and levels of depressive symptoms that take time to generalise to suicide-proneness. This finding may also suggest that reductions in suicidality and self-harm risk may take several months. Adventure therapy programs such as WAT could also consider offering one or more follow-up "booster" session to help generalise and integrate changes.

4.1.3. Effect on self-esteem

The short-term ES for Social self-esteem was moderate, positive, and statistically significant. Self-esteem subscale effects for Social, Home/Parents, and General self-esteem were moderately positive and within the expected age-based benchmark CI, whereas the School/Academic effect was lower than expected. The follow-up effect for each of the self-esteem subscales was not significant, with the exception of General self-esteem, for which participants reported a significant increase. The delayed change in General self-esteem may reflect a process that begins with improvements in social self-esteem.

4.1.4. Effect on depression

The overall short-term effect on depression was small and not significant, however the short-term effect for the 23 participants who reported clinical levels of depressive symptoms pre-program was large, positive, and statistically significant, and this was retained in the longer-term. Consistent with a previous evaluation

of WAT (Crisp et al., 2004), the pre-program mean for these participants was in the Severe range, moved to the Mild range by the end of the program, and remained in the Mild range at the three-month follow-up. There were no statistically significant short- or longer-term changes in depressive symptoms for the 13 participants who were not in the clinical range pre-program. Thus, it seems that clients with depressive symptoms respond well, with clinically meaningful reductions in symptoms. Such magnitude of benefit appears comparable to the most efficacious treatments reported in the literature (Klein et al., 2007; Michael & Crowley, 2002; Singh & Reece, 2014).

4.1.5. Effect on behavioural and emotional functioning

The overall short-term effect for the 26 participants who reported clinical levels of behavioural and emotional dysfunction pre-program was large, positive, and statistically significant, and was retained in the longer-term. Short-term effects for participants who reported clinical levels of symptoms were positive, large, and statistically significant for 6 out of the 10 of the YSR subscales (Social Problems, Aggressive Behaviour, Internalising Problems, Anxious/Depressed, Somatic Complaints, Thought Problems). These short-term changes were retained at follow-up. At the three-month follow-up, 7 out of the 10 subscale means had moved out of the clinical range. This result is important as it shows that participants experienced reduced symptoms that are clinically meaningful. For the 10 participants who reported non-clinical levels of behavioural and emotional dysfunction pre-program, there were no statistically significant short- or longer-term changes. All participants who were in the non-clinical range pre-program remained in the non-clinical range for all YSR subscales at the three-month follow-up.

The short-term WAT program effect on the most severe area of participants' symptoms was moderate, positive, and statistically significant and this change was maintained in the longer-term. This result is important as it shows that in the area of greatest need, clients experienced reduced symptoms. Although the mean *T*-score of the most severe area of participants' symptoms improved from pre-program to the three-month follow-up, it remained in the clinical range.

4.1.6. Summary of effects

Overall, these findings suggest that WAT participants experienced a small, positive, statistically significant overall improvement in psychosocial functioning, with moderate, positive, statistically significant improvements in psychological resilience, depression, and social self-esteem, and non-significant improvements in psychological, emotional, and behavioural functioning. For the most part, the changes appear to have been retained at a three-month follow-up. Although there was no short-term change in suicidality, there was a moderate, statistically significant longer-term reduction in suicidality.

As there were diverse presenting issues, additional analyses focused on participants who were in clinical ranges pre-program. These analyses revealed large, positive, and statistically significant improvements in depressive symptomatology and behavioural and emotional functioning which were retained at a three-month follow-up. In comparison, participants who were in the non-clinical range pre-program, for the most part, experienced relatively little short- or longer-term change.

4.2. Lessons learned

The WAT program is an experiential, adventure activity-based approach to prevention, early intervention, and treatment for adolescents, delivered within a case management framework. Adventure therapy is not a panacea, but it can be useful in a variety

of settings and for a broad spectrum of clients. Overall, there were statistically significant short-term improvements in resilience, social self-esteem and depression, and statistically significant improvements in the follow-up period for suicidality and general self-esteem. These results suggest that WAT can help adolescents to improve dysfunctional beliefs and attitudes, maladaptive behaviours, coping strategies, inadequate problem-solving methods, and to develop greater resilience to overcome risks and avoid negative outcomes in the future. However, the short-term change for academic self-esteem, when compared to the relevant meta-analytic benchmark, was lower than expected. Thus, the WAT program may be less well suited to improving adolescents' academic self-esteem.

Importantly, participants in clinical ranges pre-program, reported clinically and statistically significant improvements in depression, internalising and externalising problems, somatic complaints, anxiety, depression, withdrawal, social problems, thought problems and overall behavioural, and emotional functioning. These results suggest that WAT may be especially well suited to treating participants with clinical psychological symptoms. However, given the notable differences in results between participants in clinical and non-clinical ranges, it may be that WAT is an appropriate treatment modality for participants with clinical symptoms, but less well suited to prevention and early intervention. Alternatively, it may be that the evaluation methodology was not sufficiently targeted at indicators of efficacy with regard to prevention and early intervention.

The statistically significant improvements in suicidality and general self-esteem during the follow-up period may reflect a process that begins with improvements in resilience, alleviation of depressive symptoms, and enhanced social self-esteem but that takes time to generalise to suicide proneness and general self-esteem. Further research could investigate the reasons for, and mechanisms of, these delayed changes, as well as effects experienced by participants at even later post-program timepoints.

Despite the promise of adventure therapy and the WAT model, more in-depth and rigorous program evaluation could be considered. This could take the form of clinical trials of adventure therapy programs tailored to homogenous client groups (e.g., of depressed or conduct disordered adolescents) and use quasi- or fully-experimental designs, including wait-list control groups or cross-over designs with conventional treatments such as CBT. It could also be helpful to investigate which components of adventure therapy programs are most effective and which components could be improved. For example, exit or follow-up interviews with participants could be conducted to systematically document and analyse participants' responses to each of the program components. Additionally, interviews could investigate both intended and unintended effects of programs in order to gain a more holistic understanding of the obtained benefits. Case studies could also be informative in this respect, especially longitudinal case studies to further investigate longer-term changes. It may also be an important consideration for the development of future adventure therapy interventions to investigate the relative benefits of adventure therapy for different client types, presenting problems, and other individual differences.

4.3. Limitations

This study found that WAT programs are reasonably effective, however several limitations should be considered, including the evaluation design, reliance on self-reported data, small sample size, regression to the mean, missing data, and use of non-validated questionnaires.

There was no control or comparison group, therefore conclusions about causality are unable to be made. Use of a comparison and/or wait-list control group could be useful in future evaluations, and would be necessary to demonstrate the effectiveness of WAT interventions. Inclusion of comparison and/or control groups would also help to deal with the potential for regression to the mean (Barnett, van der Pols, & Dobson, 2005), given that participants in this study were pre-selected as being at-risk. Additionally, future such studies should consider including multiple long-term follow-up data collection points.

This study relied exclusively on empirical, self-reported data. Adventure therapy program evaluation studies could be strengthened by triangulating self-reported data with interviews with participants, and ratings and/or interviews with observers such as parents and teachers. Incorporation of other existing data, such as school attendance and behaviour records may also be helpful.

The small sample size ($N=36$) limited the current study's statistical power. Future studies with larger sample sizes would help to validate the findings of this study. Utilising a consistent, minimal set of measurement tools for every participant would allow for easier integration into larger data sets.

In applied, longitudinal research missing data is inevitable. The current study is no exception, with only a small percentage of participants completing all measures at each time point. Thus, multiple imputation was used and can be recommended for use in future adventure therapy research.

Two non-validated questionnaires (RQ (resilience) and CORE FFQ (family functioning)) were used, potentially limiting the reliability and validity of findings. Where possible, existing psychometrically validated assessment tools should be used and the psychometric properties of instrumentation should be reported. A task for future adventure therapy research is the development of purpose-built, multi-dimensional assessment tools using the best available psychometric techniques.

5. Conclusion

The current study contributes to research literature about psychotherapeutic effects of adventure therapy treatment for adolescents by evaluating self-reported outcomes for 36 WAT program adolescent participants. Overall, results indicated small, statistically significant improvements ($ES = 0.26$) which are within the expected range of effects for adventure therapy programs with similar target groups. Large, positive, statistically significant changes were evident for participants in clinical ranges for depression and behavioural and emotional functioning. Importantly, the changes appear to have been retained at a three-month follow-up. These findings indicate that WAT offers a potentially viable alternative treatment modality to more traditional psychotherapeutic approaches for youth at-risk with clinically significant symptoms. More in-depth investigation using triangulated data, qualitative and quantitative data, evaluation of specific program components, comparison or wait-list control groups, and a larger sample would help to better understand the obtained benefits (intended and unintended), as well as what works, how it works, and what could be improved.

References

- Özabacı, N. (2011). Cognitive behavioural therapy for violent behaviour in children and adolescents: A meta-analysis. *Children and Youth Services Review*, 33(10), 1989–1993. <http://dx.doi.org/10.1016/j.childyouth.2011.05.027>.
- Australian Bureau of Statistics. (2007). National survey of mental health and wellbeing: Summary of results (ABS Cat. No. 4326.0). Canberra, ACT: Australian Bureau of Statistics Retrieved from <http://www.ausstats.abs.gov.au/Ausstats/>
- subscriber.nsf/0/6AE6DA447F985FC2CA2574EA00122BD6/\$File/43260_2007.pdf.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- Achenbach, T. M. (1991). *Manual for the youth self report*. Burlington, VT: University of Vermont.
- Antony, M. M., & Barlow, D. H. (Eds.). (2010). *Handbook of assessment and treatment planning for psychological disorders*. 2nd ed. New York: Guilford.
- Association for Experiential Education. (2012). Theory. Retrieved from <http://www.aee.org/tagp-best-p-theory>.
- Author Unknown (2001). *C.O.R.E. family functioning questionnaire*. Kew East, Australia: Neo Psychology Unpublished manuscript.
- Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, 34(1), 215–220. <http://dx.doi.org/10.1093/ije/dyh299>.
- Barrett, P. M., & Ollendick, T. H. (Eds.). (2004). *Handbook of interventions that work with children and adolescents: Prevention and treatment*. Hoboken, NJ: Wiley.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory*, 2nd ed. San Antonio, TX: Harcourt Brace & Company.
- Bennett, D. S., & Gibbons, T. A. (2000). Efficacy of child cognitive-behavioral interventions for antisocial behavior: a meta-analysis. *Child & Family Behavior Therapy*, 22(1), 1–15. http://dx.doi.org/10.1300/J019v22n01_01.
- Bowen, D. J., & Neill, J. T. (2013a). A meta-analysis of adventure therapy outcomes and moderators. *The Open Psychology Journal*, 6, 28–53. <http://dx.doi.org/10.2174/1874350120130802001>.
- Bowen, D. J., & Neill, J. T. (2013b). A meta-analysis of adventure therapy outcomes and moderators: Pre-post adventure therapy age-based benchmarks for outcome categories. Retrieved from <http://www.danielbowen.com.au/meta-analysis>.
- Campbell, E. R. (2010). *The effectiveness of art therapy in reducing symptoms of trauma, anxiety, and stress: a meta-analysis*. Doctor of Psychology. Wheaton, IL: Wheaton College.
- Cohen, J. (1988). *Statistical power analysis for behavioral sciences*, 2nd ed. Hillsdale, NY: Lawrence Erlbaum Associates.
- Coopersmith, S. (1981). *Self esteem inventories*. Palo Alto, CA: Consulting Psychologists.
- Crisp, S. J. R. (2001). Resilience Questionnaire. Unpublished manuscript, Neo Psychology, Kew East, Australia.
- Crisp, S. J. R., & Anger, N. (1998). Wilderness-adventure therapy in adolescent mental health: a case study of a 15-year-old girl with psychological and social problems. *Australian Journal of Outdoor Education*, 2(3), 17–25.
- Crisp, S. J. R., & Hinch, C. (2004). Treatment effectiveness of WAT: A comprehensive evaluation. Kew East, Australia: Neo Psychology. Retrieved from <http://www.neopsychology.com.au/documents/TreatmentEffectivenessofWAT-SummaryFindings.pdf>.
- Crisp, S. J. R., & Noblet, L. (2001). *The adventure challenge experience – ACE program: evaluation report Victoria*. Australia: Rotary Club of North Balwyn.
- Crisp, S. J. R., & O'Donnell, M. (1998). Wilderness adventure therapy in adolescent psychiatry. *Australian Journal of Outdoor Education*, 3(1), 47–57.
- Crisp, S. J. R., O'Donnell, M. J., Kingston, L., Poot, A., & Thomas, N. E. (2000). Innovative multi-modal day-patient treatment for severely disordered at risk adolescents. In N. N. Singh, J. P. Leung, & A. N. Singh (Eds.), *International perspectives on child and adolescent mental health*: (1. pp. 331–345). Kidlington, United Kingdom: Elsevier.
- Crisp, S. J. R., Noblet, L., & Hinch, C. (2004). *Guidelines for the practice of wilderness adventure therapy: findings from the systemic wilderness adventure therapy research and development (SWATRAD) project*. Melbourne, Australia: YouthPsych Consulting.
- Crisp, S. J. R. (1997). International models of best practice in wilderness and adventure therapy: Implications for Australia. Melbourne, Australia: Austin & Repatriation Medical Centre. Retrieved from https://www.churchilltrust.com.au/media/fellows/Crisp_Simon_1996.pdf.
- Crisp, S. J. R. (1998). International models of best practice in wilderness and adventure therapy. In C. Itin (Ed.), *Exploring the boundaries of adventure therapy: international perspectives* (pp. 56–74). Boulder, CO: Association for Experiential Education.
- Crisp, S. J. R. (2003). *Treatment effects of a group based coping skills intervention for high-risk day-patient adolescents with and without a history of suicidal behaviour*. Doctor of Psychology. Melbourne, Australia: La Trobe University.
- Crisp, S. J. R. (2014). *Submission to the Australian human rights commission's examination of intentional self-harm and suicidal behaviour in children*. Unpublished manuscript. Kew East, Australia: Neo Psychology.
- Curtis, N. M., Ronan, K. R., & Bordin, C. M. (2004). Multisystemic treatment: a meta-analysis of outcome studies. *Journal of Family Psychology*, 18(3), 411–419. <http://dx.doi.org/10.1037/0893-3200.18.3.411>.
- Department of Health and Ageing. (2013). National mental health report 2013: Tracking progress of mental health reform in Australia 1993–2011. Canberra, Australia: Commonwealth of Australia. Retrieved from [http://www.health.gov.au/internet/main/publishing.nsf/Content/B090F03865A7FAB9CA257C1B0079E198/\\$File/rep13.pdf](http://www.health.gov.au/internet/main/publishing.nsf/Content/B090F03865A7FAB9CA257C1B0079E198/$File/rep13.pdf).
- Dowd, N. E., Singer, D. G., & Wilson, R. F. (Eds.). (2006). *Handbook of children, culture, and violence*. CA: Sage: Thousand Oaks.
- Ellis, P. D. (2010). *The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results*. Cambridge, UK: Cambridge: University Press.

- Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford.
- Farrington, D. P., & Welsh, B. C. (2003). Family-based prevention of offending: a meta-analysis. *Australian and New Zealand Journal of Criminology*, 36(2), 127–151. <http://dx.doi.org/10.1375/acri.36.2.127>.
- Gass, M. A., Gillis, H. L., & Russell, K. C. (2012). *Adventure therapy: theory, research, and practice*. New York: Routledge.
- Gitman, K. (2009). *The effects of music therapy on children and adolescents with mental or medical illness: a meta-analysis Doctor of Philosophy*. Buffalo, NY: University of New York.
- Glasser, S. P. (Ed.). (2014). *Essentials of clinical research*. 2nd ed. Cham, Switzerland: Springer.
- Gold, C., Voracek, M., & Wigram, T. (2004). Effects of music therapy for children and adolescents with psychopathology: a meta-analysis. *Journal of Child Psychology & Psychiatry*, 45(6), 1054–1063. <http://dx.doi.org/10.1111/j.1469-7610.2004.t01-1-00298.x>.
- Gonzalez, J. E., Nelson, J. R., Gutkin, T. B., Saunders, A., Galloway, A., & Shwery, C. S. (2004). Rational emotive therapy with children and adolescents: a meta-analysis. *Journal of Emotional & Behavioral Disorders*, 12(4), 222–235. <http://dx.doi.org/10.1177/10634266040120040301>.
- Graham, J. W. (2012). *Missing data: analysis and design*. New York: Springer.
- Greenberg, M. T., & Lippold, M. A. (2013). Promoting healthy outcomes among youth with multiple risks: innovative approaches. *Annual Review of Public Health*, 34, 253–270. <http://dx.doi.org/10.1146/annurev-publhealth-031811-124619>.
- Ipsen, J. C., Stein, D. J., Hawkrigde, S., & Hoppe, L. (2010). Pharmacotherapy for anxiety disorders in children and adolescents. *Evidence-Based Child Health*, 5(2), 555–628. <http://dx.doi.org/10.1002/ebch.541>.
- James, A. C., James, G., Cowdrey, F. A., Soler, A., & Choke, A. (2013). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews*, 6, 1–85. <http://dx.doi.org/10.1002/14651858.CD004690.pub3>.
- Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, 67(3), 332–339. <http://dx.doi.org/10.1037/0022-006X.67.3.332>.
- Kingston, L., Poot, A., & Thomas, N. E. (1997). *The brief intervention program under the microscope: a five year review of an adolescent day program*. Melbourne, Australia: Child & Adolescent Mental Health Service, Austin & Repatriation Medical Center.
- Klein, J. B., Jacobs, R. H., & Reinecke, M. A. (2007). Cognitive-behavioral therapy for adolescent depression: a meta-analytic investigation of changes in effect-size estimates. *Journal of the American Academy of Child and Adolescent Psychiatry*, 46(11), 1403–1413. <http://dx.doi.org/10.1097/chi.0b013e3180592aaa>.
- Knorth, E. J., Harder, A. T., Zandberg, T., & Kendrick, A. J. (2008). Under one roof: a review and selective meta-analysis on the outcomes of residential child and youth care. *Children and Youth Services Review*, 30(2), 123–140. <http://dx.doi.org/10.1016/j.childyouth.2007.09.001>.
- Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive-behavioral programs for offenders: a meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, 1(4), 451–476. <http://dx.doi.org/10.1007/s11292-005-3541-7>.
- Michael, K. D., & Crowley, S. L. (2002). How effective are treatments for child and adolescent depression? A meta-analytic review. *Clinical Psychology Review*, 22(2), 247–269. [http://dx.doi.org/10.1016/S0272-7358\(01\)00089-7](http://dx.doi.org/10.1016/S0272-7358(01)00089-7).
- Mrazek, P. J., & Haggerty, R. J. (1994). *Reducing the risks for mental disorders: frontiers for preventive intervention research*. Washington, DC: National Academy Press.
- Newes, S. L. (2000). *Adventure-based therapy: Theory, characteristics, ethics, and research*. Pennsylvania State University. Retrieved from <http://wilderdom.com/SandyNewes.htm>
- Norton, C. L., Carpenter, C., & Pryor, A. (Eds.). ([267_TD\$DIFF]2015). *Adventure therapy around the globe: international perspective and diverse approaches*. Champaign, IL: Common Ground.
- Patel, V., Flisher, A. J., Hetrick, S., & McGorry, P. (2007). Mental health of young people: a global public-health challenge. *The Lancet*, 369, 1302–1313. [http://dx.doi.org/10.1016/S0140-6736\(07\)60368-7](http://dx.doi.org/10.1016/S0140-6736(07)60368-7).
- Perkins, D. F., & Borden, L. M. (2003). Positive behaviors, problem behaviors, and resiliency in adolescence. In I. B. Weiner, R. M. Lerner, M. A. Easterbrooks, & J. Mistry (Eds.), *Handbook of psychology* (pp. 373–394). New York: Wiley.
- Randolph, J. J., & Edmondson, R. S. (2005). Using the binomial effect size display (BESD) to present the magnitude of effect sizes to the evaluation audience. *Practical Assessment Research & Evaluation*, 10(14), 1–7.
- Rickwood, D. J., Deane, F. P., & Wilson, C. J. (2007). When and how do young people seek professional help for mental health problems? *Medical Journal of Australia*, 187(7), 35–39.
- Rohde, P., Lewinsohn, P. M., Seeley, J. R., & Langhinrichsen-Rohling, J. (1996). The Life Attitudes Schedule Short Form: an abbreviated measure of life-enhancing and lifethreatening behaviours in adolescents. *Suicide and Life-Threatening Behaviour*, 26(3), 272–280.
- Rohde, P., Seeley, J. R., Langhinrichsen-Rohling, J., & Rohling, M. L. (2003). The life attitudes schedule—short form: psychometric properties and correlates of adolescent suicide proneness. *Suicide and Life-Threatening Behavior*, 33(3), 249–260. <http://dx.doi.org/10.1521/suli.33.3.249.23216>.
- Rutter, M., & Taylor, E. A. (Eds.). (2005). *Child and adolescent psychiatry*. 4th ed. Malden, MA: Wiley-Blackwell.
- Rutter, M., Bishop, D. V. M., Pine, D. S., Scott, S., Stevenson, J., Taylor, E., & Thapar, A. (Eds.). (2008). *Rutter's child and adolescent psychiatry*. 5th ed. Oxford, UK: Blackwell.
- Sakinofsky, I., Cardish, R. J., Doey, T., Grek, A., Mamo, D. C., McMains, S., et al. (2007). Caring for the suicidal patient: an evidence-based approach. the canadian journal of psychiatry, 52(6. Supplement 1)15–1365.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <http://dx.doi.org/10.1037//1082-989X.7.2.147>.
- Schell, L., Cotton, S., & Luxmoore, M. (2012). Outdoor adventure for young people with a mental illness. *Early Intervention in Psychiatry*, 6(4), 1–8. <http://dx.doi.org/10.1111/j.1751-7893.2011.00326.x>.
- Singh, N., & Reece, J. (2014). Psychotherapy, pharmacotherapy, and their combination for adolescents with major depressive disorder: a meta-analysis. *The Australian Educational and Developmental Psychologist*, 31(1), 47–65. <http://dx.doi.org/10.1017/edp.2013.20>.
- Sukhodolsky, D. G., Kassinove, H., & Gorman, B. S. (2004). Cognitive-behavioral therapy for anger in children and adolescents: a meta-analysis. *Aggression and Violent Behavior*, 9(3), 247–269. <http://dx.doi.org/10.1016/j.avb.2003.08.005>.
- Tucker, A. R., Javorski, S., Tracy, J., & Beale, B. (2012). The use of adventure therapy in community-based mental health: decreases in problem severity among youth clients. *Child and Youth Care Forum*, 1, 25, . <http://dx.doi.org/10.1007/s10566-012-9190-x>.
- Villani, S. (2001). Impact of media on children and adolescents: a 10-year review of the research. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(4), 392–401. <http://dx.doi.org/10.1097/00004583-200104000-00007>.
- Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: a meta-analysis. *Psychological Bulletin*, 132(1), 132–149. <http://dx.doi.org/10.1037/0033-2909.132.1.132>.
- World Health Organization (2001). *The world health report 2001: mental health: new understanding, new hope*. Geneva, Switzerland: World Health Organization.

Daniel J. Bowen is a registered psychologist and is currently completing a PhD in Clinical Psychology in the Centre for Applied Psychology at the University of Canberra. Daniel's research focuses on the effectiveness of adventure therapy programs in Australia. For more information, see <http://www.danielbowen.com.au>.

James T. Neill is an Assistant Professor in the Centre for Applied Psychology at the University of Canberra. James' research examines processes and outcomes of outdoor, experiential, and nature-based programs. For more information, see <http://wilderdom.com>.

Simon J.R. Crisp is a Clinical Child, Adolescent and Family Psychologist and has researched and developed Wilderness Adventure Therapy® (WAT) which provides interventions in public mental health, community counselling, private psychology practice, and school settings. For more information, see <http://www.neopsychology.com.au>.