A clustered-randomized controlled trial of a self-reflection resilience-strengthening intervention and novel mediators

*[1]Samantha L. Falon, [1]Eyal Karin, [1, 2]Danny Boga, [3]Daniel F. Gucciardi, [1]Barbara Griffin & [1]Monique F. Crane

[1]Department of Psychology, Macquarie University

[2]Australian Army Psychology Corps, Australian Defence Force

[3] School of Physiotherapy and Exercise Science, Curtin University

* Correspondence to:
Samantha Falon
Department of Psychology, 4 First Walk,
Macquarie University, North Ryde, NSW, 2109, Australia
E: samantha.falon@students.mq.edu.au
P: +61 434 934 288

Author Note: This research was supported by an Australian Government Research Training Program Scholarship. Previously, parts of this research have been presented to the Australian Defence Force and at the Australasian Military Medicine Association (AMMA) 2018 Annual Conference, the Centre for Elite Performance Expertise and Training (CEPET) 2018 Annual Conference, and the Society for Industrial and Organizational Psychology (SIOP) 2020 Annual Conference.

Falon, S. L., Karin, E., Boga, D., Gucciardi, D. F., Griffin, B., & Crane, M. F. (2020). A clustered-randomized controlled trial of a self-reflection resilience-strengthening intervention and novel mediators. *Journal of Occupational Health Psychology.* https://doi.org/10.1037/ocp0000268

Abstract

This study tests the efficacy of a unique resilience-strengthening intervention using a clustered-randomized controlled trial. It was hypothesized that the training, which encourages adaptive self-reflection on stressor events and the effectiveness of coping strategies and resources, would exert a positive effect on mental health outcomes via increased reflection and decreased brooding. The trial was conducted during a significant stressor period with a final sample of 204 second-class Officer Cadets from the Royal Military College, Australia. Platoons of Cadets were randomly allocated to either Self-Reflection Resilience Training (n = 96) or an exposure-matched active control group that received training as usual (i.e., cognitive-behavioral skill development training) and communication skills seminars (n = 108). Compared to the active control group, Self-Reflection Resilience Training was more effective at preventing the onset of depression symptoms and promoting stable levels of perceived stress during a period of increased exposure to training stressors, consistent with a resilient trajectory. The Self-Reflection group unexpectedly demonstrated higher anxiety symptoms than the Control group at immediate follow-up, but these symptoms returned to baseline levels at longer-term follow-up. In contrast, the Control group experienced increasing anxiety symptoms between immediate and longer-term follow-up. Mediation analyses supported an indirect effect of Self-Reflection Resilience Training on all three outcome measures via brooding, but not via reflection. This study provides support for the capacity of a practical, sustainable, and scalable intervention based on self-reflection to strengthen resilience in the military training setting.

Keywords: mental health; military; stressors; rumination; brooding

The capacity for resilience is widely regarded to be important in contemporary workplaces (e.g., Ceschi et al., 2017; Shatté et al., 2017). Although resilience-strengthening interventions are being disseminated rapidly as a result of this interest, significant limitations in the efficacy of currently available interventions remain (Leppin et al., 2014; Robertson et al., 2015; Vanhove et al., 2016). Resilience as an outcome is characterized by "the maintenance or quick recovery of mental health during and after exposure to significant stressors" (p. 786; Kalisch et al., 2017). In this way, resilience is observed in the context of some level of risk, whereby individuals either 'bounce back' and rapidly return to normal levels of functioning following a stressful event (Bonanno, 2005; Masten, 2007; Rutter, 2006) or experience 'robust' resilience (i.e., no decline in functioning; Fletcher & Sarkar, 2016).

An individual's capacity for resilience is widely regarded to be teachable and amenable to change (Chmitorz et al., 2018). Models of resilience and empirical work have highlighted the malleable capacities, resources, and coping strategies required to facilitate resilience during times of increased risk (e.g., Bonanno et al., 2007; Bonanno et al., 2011). Traditionally, interventions aimed at developing resilience have focused on supporting the development of cognitive and behavioral coping skills that serve to increase the likelihood of a resilient outcome (Robertson et al., 2015). Largely, these interventions have their basis in cognitive-behavioral therapeutic strategies (e.g., Cohn & Pakenham, 2008; Millear et al., 2008; Sood et al., 2011; Waite & Richardson, 2004). However, recent meta-analyses (Joyce et al., 2018; Vanhove et al., 2016) and a systematic review (Robertson et al., 2015) conclude that these interventions collectively demonstrate small to moderate and inconsistent effects on mental health outcomes. When positive outcomes are initially observed, they often fail to be maintained over time (Vanhove et al., 2016). The challenge of achieving observable effects is especially clear in high-risk occupational environments such as the military (e.g., Adler et al., 2009; Castro et al., 2012; Cohn & Pakenham, 2008; Cornum et al., 2011; Harms et al., 2013).

Self-Reflection and the Strengthening of Resilience: Theoretical Underpinnings

Capacities for resilience are the biological, social, psychological, and environmental characteristics that optimise the likelihood of experiencing positive mental health following a psychological risk (Kalisch et al., 2015). Yet, little is known about how such capacities for resilience emerge and are strengthened. Cumulating research has highlighted the potential for resilience to emerge from exposure to moderately stressful events and adversity (Crane & Searle 2016; Moore et al., 2018; Seery et al., 2010; Seery et al., 2013). However, exposure to stressors alone is insufficient for the ongoing development of resilience (Angel, 2016; Sarkar & Fletcher, 2017).

Previous theoretical work proposed that introspective processes are involved in converting major life events into resilience-strengthening experiences (Richardson, 2002), yet the nature of these introspective processes were undefined making the translation into application difficult. Crane, Searle, Kangas, et al. (2019) extended this initial theoretical work via the Systematic Self-Reflection model of resilience strengthening, which proposed five adaptive self-reflective practices entailing specific reflection on: (1) one's emotional, physical, behavioral, and cognitive responses to triggering events, (2) values and value-based goals in relation to the situation, (3) strategies applied to address the situation, (4) strategy effectiveness in relation to values and value-based goals, and (5) constructive adaptations of strategies in order to improve one's coping approach, including a search for new strategies using available resources (e.g., peers, psychologists). These five practices are thought to scaffold the stressor experience in a way that develops: (1) self-*awareness*, referring to insight into one's emotional, cognitive, physical, and behavioral reactions to the stressor, their relationship to each other, the events that trigger those reactions, and one's values in the context of the stressor, (2) self-*evaluation*, referring to an understanding of the effectiveness of the coping strategies, resources, and beliefs applied in different stressor or adversity contexts, and whether these

approaches move a person toward or away from their personally endorsed values, and (3) self-*development*, referring to a consideration of how one's coping approach or access to coping resources may be improved and adapted. The self-developmental style of questions prompt individuals to identify what could be done differently in the future, describe how they might search for new solutions, or describe additional resources that could be engaged to support coping. The process of iterative structured self-reflection during and after exposure to moderate everyday stressors and the emergence of greater self-understanding in the context of stressors is proposed to drive adaptations in these capacities that promote resilient outcomes (Crane, Searle, Kangas, et al., 2019).

This multifaceted approach to self-reflection applied in the resilience training context is based on insights from both psychology and education. Within the field of psychology, self-reflection has been described by several researchers as an inherently cognitive process that prompts individuals to evaluate past experience, adaptively consider how external feedback relates to important outcomes, and implement behavioral changes that drive learning (Anseel et al., 2015; Ellis et al., 2014; Yang et al., 2018). Crane, Searle, Kangas, et al. (2019) draw on work highlighting the use of self-reflection to broaden behavioural options and improve workplace performance (Ellis et al., 2014). These authors also draw from the clinical literature, and in particular approaches aligned with Acceptance and Commitment Therapy (ACT). For example, self-as-context in ACT describes the pure awareness of the observing self whereby the individual is willing to observe the self with dispassionate curiosity (Harris, 2009). Moreover, the identification of, and behavioural consistency with, personally endorsed values is also central to both the self-reflective approach and ACT. According to the Systematic Self-Reflection model (Crane, Searle, Kangas, et al., 2019), values are used as a benchmark for evaluating coping outcomes and as a source of feedback. Feedback regarding the usefulness of coping strategies is achieved by monitoring value adherence in the context of stressors.

Self-reflection is a key topic of interest in the adult education literature (Cranton, 2006; Tusting & Barton, 2003), and is thought to facilitate learning by allowing the individual to explore subject matter first-hand, integrate new insights into their existing knowledge base, learn from mistakes, and understand the real-world applications of what they have learned (Kolb, 1984). Transformational learning theory (Mezirow, 1991; 1997) describes that the purpose of self-reflection is to challenge one's current assumptions about the world, revise assumptions that lack coherence, and integrate new information into the rest of one's worldview. Other studies have drawn attention to the capacity for interventions based on self-reflection to enhance complex decision-making (Yang et al., 2018) and negotiation performance (Kray et al., 2009). Yet, little research to date has investigated the implications of self-reflection for the development of resilience.

The Self-Reflective Approach to Resilience Training: Application and Modifications

Preliminary work has demonstrated the promise of the self-reflective approach to resilience training in the military context. A clustered-randomized controlled trial with a sample of 226 Officer Cadets compared Self-Reflection Resilience Training (SRT), which is based on the aforementioned self-reflective approach, to the currently implemented Cognitive-Behavioral Skill Development Training (CBST; Crane, Boga, Karin, et al., 2019). Findings demonstrated that following exposure to significant training demands, SRT resulted in moderate to large reductions in symptoms of depression and anxiety at longer-term follow-up (3 months post-intervention). In order to encourage the aforementioned multifaceted approach to self-reflection, the training was comprised of a single 40-minute lecture-style session and five weekly guided journaling exercises. In the current study, we have made four key modifications to the original training. First, the introductory brief provided to Cadets was simplified into six key learning objectives that were articulated at the outset of training: (1) how resilience relates to performance, (2) that resilience is malleable and can be developed

over time, (3) that stressors experienced as part of training can function as an opportunity to strengthen resilience, (4) that everyone has the capacity for resilience, and greater capacity can be developed, (5) how identifying leadership values and valued goals can help us to determine the usefulness of our coping strategies and resources, and (6) why self-reflection is important and how to achieve it meaningfully.

Second, the design of the values activity completed during the initial brief was modified. In the first trial, Cadets were asked to identify their values and value-based goals pertaining to leadership under pressure using a writing activity whereby they were required to consider the characteristics of leaders they admired. In contrast, during the current trial, we provided Cadets with a list of 32 values (e.g., integrity) and asked Cadets to rate these values from 'very important' to 'not at all important' on a six-point Likert scale. Cadets also had the option to specify their own values. The values that Cadets identified as being important or very important were referred to during the self-reflection journaling questions as a way of evaluating coping strategies and resources (e.g., "Consider the characteristics that you want to demonstrate during the process of achieving your desired outcomes, reflecting on the values you identified").

Third, the self-reflective journal questions received by Cadets were also modified after the initial trial to emphasize the learning and self-development intentions of the self-reflective process and to increase implementation of new strategies. For example, the question "*why* do you think your strategy was effective" which was intended to prompt the identification of mechanisms for efficacy, was changed to "*what* aspects of your strategy were effective" in order to more simply ask participants to list the effective elements of their strategy (i.e., 'talking to my girlfriend helped me to see my problems from a different perspective'). Finally, all self-reflective activities in the previous trial required Cadets to reflect on their own stressor reactions and coping approaches. In this version, we added an additional reflective activity that

encouraged the observation of others. Reflecting on how others bounce back from adversity may be an avenue through which individuals can continue to develop their personal resilience capacities through vicarious learning (Bandura, 1986; Smith et al., 2010). Sample questions included in each week of the reflective activities are provided as supplementary material (see Table S1). In light of these modifications, a key objective of the current trial was to ensure that SRT continues to be effective at strengthening resilience.

Distinguishing SRT from Existing Resilience Training Approaches

The focus of SRT is to cultivate insights into one's coping approach via the process of self-reflection. These insights are then proposed to drive personalized adaptations in coping application and resource use that facilitate the ongoing development of resilience. In this way, exactly which coping strategies and resources are used and how they are applied in context is based on individual choice. This feature distinguishes the SRT approach from other resilience training interventions that focus on the cultivation of cognitive and behavioral skill-development approaches to coping that are broadly considered to be adaptive, such as controlled breathing, positive self-talk, support seeking, mindful meditation practices (e.g., Cohn & Pakenham, 2008; Liossis et al., 2009; McCraty & Atkinson, 2012; Millear et al., 2008). In contrast, the SRT approach encourages participants to seek, formulate, and implement new coping approaches on the basis of identified personal coping 'gaps' that are pinpointed via the self-reflection process. Participants may then reflect on the efficacy of these coping approaches in the following self-reflection activity.

This self-directed refinement of an individual's coping repertoire espoused by the SRT program is motivated by three key observations across resilience research. First, research investigating the demonstration of resilience typically identifies that a majority of individuals experience resilient outcomes (e.g., Bonanno et al., 2006). Thus, individuals already engender an important set of personal resilience capacities that could be highlighted via the self-

reflection process. Second, accumulating research is demonstrating the importance of possessing a diverse range of coping strategies that are flexible and fit for the demand (Aldao et al., 2015; Bonanno & Burton, 2013). The ability to continuously reflect on and adjust one's strategies is likely to encourage the flexible deployment of strategies in dynamic stressor contexts. Finally, the SRT approach inherently conveys that stressors are opportunities for growth, rather than a threat to wellbeing, thereby capitalizing on research emphasizing the relationship between the tendency to appraise potentially aversive events as non-threatening or as having the potential to culminate in positive and resilient outcomes (e.g., Kalisch et al., 2015).

Despite the initial promise of previous findings for the SRT approach, there is an important need for replication in prevention science (Valentine et al., 2011) against rigorous control groups. In the study conducted by Crane, Boga, Karin, et al. (2019), the SRT intervention was 75 minutes longer in duration than the active control program. In the absence of an active control that is equivalent in length to the intervention, it is not possible to rule out whether non-specific treatment effects (such as greater exposure to a psychologist or trainer) are responsible for observed changes in mental health outcomes, rather than the content of the SRT intervention (Chmitorz et al., 2018). Thus, a further aim of this trial is to evaluate the SRT program against a more rigorously designed control condition.

Mediators of the Effect of Self-Reflection Resilience Training on Mental Health Outcomes

In the current research, we also explore two possible mechanisms that we hypothesize may underpin the effectiveness of the SRT program. Although the Systematic Self-Reflection model of resilience strengthening proposes several possible intervening mechanisms as individuals adjust their coping repertoire (see: Crane, Searle, Kangas, et al., 2019), we focus on a mechanism that is proximal to the action of the intervention: *self-reflection.* We thus seek

to test whether the SRT intervention enhances resilience via engagement in adaptive self-reflection. However, as a form of self-focused attention, there is also a potential dark side to engaging in self-reflection in that self-reflective activities may engender greater brooding. Although the distinctions between self-reflection and brooding have been under-researched, the available evidence suggests that these terms are distinct constructs as opposed to opposite ends of a single continuum (Trapnell & Campbell, 1999; Treynor et al., 2003). Nevertheless, they share the characteristic that attention is directed toward the self and one's thoughts, actions and affect.

Initially, both forms of self-focused attention were considered to result in harmful outcomes for wellbeing. Trapnell and Campbell (1999) coined the term "self-absorption paradox" in order to explain the reasons for these harmful outcomes, suggesting that "frequent inspection of one's feelings and thoughts seems to improve the accuracy of self-knowledge, but at the cost of psychological wellbeing" (Trapnell & Campbell, 1999, p. 299). However, recent evidence suggests that self-reflection does not necessarily result in negative mental health outcomes, with Treynor et al. (2003) demonstrating that although brooding predicted an increase in depression symptoms over a one-year period, self-reflection conversely predicted a *decline* in depression symptoms over the same time period. Other empirical work has also linked self-reflection to a number of valued outcomes, including increased perspective taking and empathetic concern (Joireman et al., 2002), increased personal growth (Harrington & Loffredo, 2010), increased vigor and decreased exhaustion (Casper et al., 2019), decreased depression symptoms (Takano & Tanno, 2009), and decreased stress (Samaie & Farahani, 2011).

Authors have subsequently proposed that adaptive self-reflection is characterised by key differences in the motivational orientation, subject, structure and controllability of a person's self-reflection. Specifically, adaptive forms of self-reflection are proposed to be

solution-focused, constructive, driven by curiosity, time-bound and characterized by a genuine intent to learn from the experience (Takano & Tanno, 2009; Treynor et al., 2003; Schoofs et al., 2010). In this way, self-reflection is adaptive when the purpose of one's self-reflection is to solve problems, plan for the future, or develop oneself in a constructive manner (Segerstrom et al., 2003). These distinctions in the nature of self-reflection are considered to account for why self-reflection results in these seemingly contradictory outcomes. Drawing on these adaptive elements of self-reflection, Crane, Searle, Kangas, et al. (2019) proposed that, rather than being focused on the reasons for depressed mood as a sub-type of rumination (Miranda & Nolen-Hoeksema, 2007), adaptive self-reflection on everyday stressors is focused on developing awareness and understanding of one's approach to stressors in context for the purpose of the development of resilient capacities. Adaptive self-reflection of this type is proposed to develop resilient capacities through the purposeful dedication of one's attention to triggers and reactions to stressors, what can be learnt from stressor exposure, the efficacy of coping strategies and resources used to manage the stressor, and new capacities to be developed and applied during future exposure to stressors. The SRT program encourages participants to structure self-reflection in this adaptive manner through the use of structured journaling activities.

In contrast to this adaptive and solution-focused characterization of self-reflection, brooding is regarded to be a maladaptive cognitive reaction to stressful events (Takano & Tanno, 2009; Treynor et al., 2003; Schoofs et al., 2010). Brooding is thought to entail repetitive and uncontrollable thoughts about the causes, negative consequences and symptoms of current negative affect (Nolen-Hoeksema, 1991; Schoofs et al., 2010), a self-referential, narrowly-focused fixation on perceived threats, losses, shortcomings or injustices to the self (Takano & Tanno, 2009; Van Vugt & Van der Velde, 2018), repetitive thoughts about interferences associated with stressful life events (Alloy et al., 2000), continued processing of a past event

or interaction (Clark & Wells, 1995), and a preoccupation with the discrepancy between one's actual and desired status (Treynor et al., 2003; Wells & Matthews, 1996). The adverse implications of brooding are well known, with several studies drawing attention to its relationship with decreased perspective taking and empathetic concern (Joireman et al., 2002), decreased psychological wellbeing and life satisfaction (Harrington & Loffredo, 2010), decreased vigor and increased exhaustion (Casper et al., 2019), impaired sleep (Demsky et al., 2018), increased depression symptoms (Takano & Tanno, 2009), and increased stress (Samaie & Farahani, 2011).

Through the use of structured journaling activities, we propose that the SRT program encourages the adaptive self-reflective aspects of self-focused attention and simultaneously reduces harmful brooding by reshaping the motivation for, subject of, and structure of self-focused attention into more adaptive forms of self-reflection characterized by a focus on the development of resilience and a curiosity about one's own coping approach. This focus on positive self-development and future thinking has been previously associated with lower levels of brooding (O'Connor & Williams, 2014). In the present study, the structure of the journaling activity means that self-reflection is time-bound, focused on addressing past gaps in coping, and focused on living consistently with one's values in the future. Previous research has demonstrated that self-focused attention directed at correcting past mistakes and achievement produces significantly greater performance improvements than state-focused or task-irrelevant rumination (Ciarocco et al., 2010).

Other related research has demonstrated that the positive effects of expressive writing interventions are mediated by reductions in brooding (Gortner et al., 2006). Gortner and colleagues (2006) hypothesized that this outcome was attributable to the capacity for expressive writing to restructure the negative, brooding cognitions that typically accompany depression into solution-focused, constructive patterns of thinking. However, their positive

outcomes were not mediated by the reflection subscale. Perhaps this was because the expressive writing intervention, unlike SRT, did not explicitly promote adaptive self-reflection on stressors and coping approaches. Collectively, these distinguishing aspects of the SRT process are thought to enhance adaptive self-reflection, but at the same time promote reductions in brooding.

The Present Study

This clustered-randomized controlled trial evaluated the effectiveness of SRT for a sample of Officer Cadets during a period of increasing stressor demands. Our operationalization of resilience in this context was consistent with our earlier definition involving the measurement of mental health outcomes in the context of psychological risk. The depression and anxiety measures used assess several symptom criteria including emotional, cognitive, somatic, and behavioral dimensions of depression and anxiety (Kroenke et al., 2001; Spitzer et al., 2006). These outcomes were compared between a group who received the SRT program and an active control group. It was anticipated that the SRT group would experience a significantly greater decline in depression symptoms (H1), anxiety symptoms (H2), and perceived stress (H3) between immediate follow-up and longer-term follow-up testing (3 months post-intervention), compared to the Control group. In addition, decreases in brooding (H4) and increases in reflection (H5) were expected to mediate the relationship between SRT and all three outcome measures.

Method

Research Context

We conducted the SRT intervention during routine training at the Australian Royal Military College.[1] Officer training consists of three phases, each six months long, with several

[1] The study was approved by the Department of Defence and Veterans' Affairs Human Research Ethics Committee (Approval Number: 015-17), reviewed by the Macquarie University Human Research Ethics Committee (Approval Number: 5201800005), and pre-registered with the Australian New Zealand Clinical Trials Registry (Trial ID: ACTRN12617001377325).

weeks break between phases. Cadets may return home or take periods of leisure time during these breaks. Officer training involves a combination of lectures, physical training, field exercises, and ongoing assessments of fitness, academic ability, leadership skills, operational skills, and character. As part of these activities, Cadets have the opportunity to demonstrate their capacity to lead their peers as part of field exercises and day-to-day training. The current study was conducted during the second phase (second six months) of training that is considered to be an especially stressful and formative period for many Cadets due to their involvement in several field-based activities, dynamic simulations of ground-warfare, and other arduous training activities. Training demands are minimal during the initial few days of second class, but increase rapidly in intensity over the subsequent three to four months and reach a peak during an arduous field exercise involving reduced food and sleep, after which they receive a short rest period. Our intention was to capture the rise and fall of training demands over the course of the second phase of training. To achieve this goal, the evaluation surveys were timed such that: (1) the baseline survey was administered in the first few days of this second phase of training where demands are relatively low (i.e. prior to the delivery of the majority of lectures, field exercises, and assessments); (2) the immediate follow-up survey was administered during a more intensive combination of lectures, training, and assessments, but prior to the most demanding field exercises; and (3) the longer-term follow-up survey was administered during a short rest period immediately following the completion of the arduous field exercise.

Participants

Participants were 218 second-class Cadets recruited from the Royal Military College, Australia. There were no exclusion criteria for participation in the study. Of the total sample, 14 participants were excluded from the final analysis either because they changed platoons during the study period and therefore changed study conditions ($n = 4$), or they failed to provide

identifying information and their data could not be matched across time points (*n* = 10). The sample at baseline therefore consisted of 204 Cadets (86% males, mean age = 23.18 years, *SD* = 3.91, age range = 18-41 years). Participants had served in the Army for an average of 2.40 years (*SD* = 2.06, range = 0 to 17 years). Figure 1 provides information about the pattern of participant attrition from the study.

Cadets typically commence their training at the Royal Military College through one of three entry pathways: (1) they have completed an undergraduate degree at the Australian Defence Force Academy and are required to complete training at the Royal Military College to become an Officer, (2) they are a civilian who has been selected for their potential to become an Officer, or (3) they are a general enlistee (Australian or International) who has been selected to become an Officer. Of the total group, 47.06% were completing a degree at the Australian Defence Force Academy (*n* = 96) and 52.94% entered through other pathways (*n* = 108); 8.8% of Cadets (*n* = 18) had served in international military organizations prior to commencing training at the Royal Military College.

Design and Procedure

The design of the study was a clustered-randomized controlled trial whereby randomization occurred by platoons. Although randomization was conducted using a clustering variable (i.e., platoons), the unit of analysis was the *individual* over time between intervention groups. Self-reflection Resilience Training is an individual-level intervention involving personal self-reflection that is hypothesised to influence individual mental health outcomes.

Prior to the collection of baseline data, eight platoons were randomly assigned by the lead researcher into two groups. Four platoons were assigned to the SRT group and the other four platoons were assigned to the Control group. The groups were separated into two lecture theatres. The researchers were blind to the baseline functioning of platoon members and had no prior knowledge of platoon functioning on allocation. This approach to group allocation is

common practice during military evaluations as it avoids the cross-contamination of training content and is most likely to ensure continuity in intervention group membership (e.g., Cacioppo et al., 2015). All platoons were deliberately similar in their composition of males and females and were not distinct by function (e.g., infantry). However, differences can emerge as a function of platoon sub-culture and platoon membership needs to, at least initially, be considered in the analysis. Attempts were made to blind participants to the experimental condition by informing them that the purpose of the research was to compare two different resilience development programs and that the researchers held no expectations regarding which program would be more beneficial.

A baseline survey was administered prior to the delivery of any training content. The immediate follow-up survey was administered after the final self-reflection exercise and communication skills seminars were completed in week five, and the longer-term follow-up survey was administered approximately three months later. All surveys were completed under supervision using pen and paper. Data were linked over time using the participant's employee number to allow longitudinal change to be examined at the individual level. This is a number that is unique to each employee at the Australian Defence Force, but cannot identify the individual to non-Defence employees. The CONSORT flowchart summarizing this procedure is provided in Figure 1.

Intervention Groups

Cadets in the SRT group (n = 96) undertook a 40-minute lecture-style session, proceeded by a weekly guided self-reflection journaling exercise (15 minutes each) for five weeks (five in total). During these journaling sessions, the training instructor was present to answer any questions about the self-reflection exercises, but otherwise provided Cadets time to write. In contrast, Cadets in the Control group (n = 108) group completed "training as usual"; that is, a 40-minute lecture-style session of the currently applied cognitive-behavioral skills

training (CBST) utilized by the Australian Defence Force. CBST includes content on specific techniques such as controlled breathing, progressive muscle relaxation, grounding, thought appraisal, positive self-talk, and support-seeking (see Cohn et al., 2011 for a detailed description). In order to control for the length of contact time with a psychologist and for other non-specific treatment effects (e.g., attention), Cadets in the Control group also completed weekly communication skills seminars (15 minutes each) for five weeks (five in total) in lieu of the five self-reflection exercises completed by those in the SRT group.

The communication skills seminars were developed by author DB for the present study and included lessons on: (1) the communication process, (2) perspectives on communication, (3) elements of communication, (4) communication styles, and (5) basic listening skills. In contrast to the five guided journaling exercises completed by participants in the SRT group, these five communication skills seminars did not involve the use of a workbook or the completion of individual activities. As such, instructors for the Control group spent comparatively more time directly teaching the participants during the communication skills seminars than the instructors who facilitated the guided journaling exercises for the SRT group. The trainers were not blind to the study conditions. However, attempts were made to reduce bias in the quality of delivery. First, the CBST program was delivered by a Royal Military College psychologist experienced in delivering the CBST program with no involvement in the current study. Second, the development of the communication skills training package was overseen by Royal Military College psychologists to ensure that the training was useful and relevant to Cadets. Third, all training was delivered by military psychologists experienced in training large groups of soldiers.

Measures[2]

[2] A number of secondary variables were also measured for future research. Such measures have not been examined in the present paper, but will be analyzed as part of a separate study reflecting distinct research questions. These include measures of coping self-efficacy (Chesney et al., 2006), optimism (Scheier et al., 1994), and dispositional hope (Snyder, 1995).

Depression Symptoms

The 9-item Patient Health Questionnaire (PHQ-9; Kroenke et al., 2001) was included as a measure of current depression symptoms at each time point. Participants indicated how often they had been bothered by particular problems over the last two weeks using a 4-point Likert scale from 0 (Not at all) to 4 (Nearly every day). The PHQ-9 demonstrated satisfactory internal reliability at all three time points ($\alpha_{Baseline} = .84$; $\alpha_{Immediate} = .87$; $\alpha_{Long\text{-}Term} = .87$).

Anxiety Symptoms

The 7-item Generalized Anxiety Disorder Scale (GAD-7; Spitzer et al., 2006) was included as a measure of current anxiety symptoms at each time point. Participants indicated how often they had been bothered by particular problems over the last two weeks on a 4-point Likert scale from 0 (Not at all) to 4 (Nearly every day). The GAD-7 demonstrated satisfactory internal reliability at all three time points ($\alpha_{Baseline} = .92$; $\alpha_{Immediate} = .92$; $\alpha_{Long\text{-}Term} = .93$).

Perceived Stress

The 14-item Perceived Stress Scale (PSS; Cohen et al., 1983) was included as a measure of perceived stress at each time point. Participants indicated how often they had felt or thought a certain way (e.g., "How often have you felt upset because of something that happened unexpectedly?" or "How often have you felt difficulties were piling up so high that you could not overcome them?") during the last two weeks on a 5-point Likert scale from 0 (Never) to 4 (Fairly often). The PSS demonstrated satisfactory internal reliability at all three time points ($\alpha_{Baseline} = .81$; $\alpha_{Immediate} = .80$; $\alpha_{Long\text{-}Term} = .81$).

Mediator Variables

Brooding and reflection were measured at each time point using sub-scales from the Rumination Response Scale (RRS; Treynor et al., 2003), which is one of the most commonly used measures of ruminative thought patterns (e.g., Glass et al., 2019; Goldin et al., 2016). In line with studies conducted by authors such as Gortner and colleagues (2006), the brooding

sub-scale included 5 items pertaining to ruminative thoughts and behaviors (e.g., "I think: 'What am I doing to deserve this?'"), and the reflection sub-scale contained 2 items pertaining to reflective thoughts and behaviors (e.g., "I go away by myself to think about why I feel this way"; "I go someplace alone to think about my feelings"). The original item about "I write down what I am thinking about and analyze it" was removed as it was considered to be confounded with the treatment condition. Participants indicated how often they had thought or carried out each response when they have felt down, sad, or depressed on a 4-point Likert scale from 1 (Almost never) to 4 (Almost always). The brooding sub-scale demonstrated satisfactory internal reliability at all three time points ($\alpha_{Baseline}$ = .82; $\alpha_{Immediate}$ = .82; $\alpha_{Long\text{-}Term}$= .84), however the reflection sub-scale had somewhat lower coefficient alphas, especially at the immediate follow-up survey ($\alpha_{Baseline}$ = .83; $\alpha_{Immediate}$ = .67; $\alpha_{Long\text{-}Term}$ = .76)

Stressor Frequency

Stressor frequency was assessed using a checklist designed for use in the present research. Participants indicated how often 23 stressors occurred in the past week (e.g., a lot of academic work, performance being assessed, physical demands, inefficient communication from superiors or other Cadets regarding training-related tasks) on a 6-point Likert scale from 0 (No days) to 6 (Every day).

Analysis Strategy

Statistical Power

The determination of sample size and statistical power followed the DELTA2 guidelines for clinical trials (Cook et al., 2018). According to the DELTA2 guidelines, previously reported estimates from similar contexts and designs should be used to determine statistical parameters such as change over time, symptom score variance, and within-subject correlation. We used data from a similar self-reflection intervention conducted within the context of the Royal Military College (Crane, Boga, Karin, et al., 2019) to inform estimates

regarding observed change over time, symptom score variance, and within-subject correlation using the Longpower package in R (Donohue & Edland, 2016). The resulting power calculation reported in Crane, Boga, Karin, et al. (2019) indicated that the projected sample of 100 participants within each group was adequately powered to refute minimally detectable symptom difference between the groups that were as small as 20% of the total effect. For example, hypothetically, if the rate of symptom change within one group condition would be 50%, the sample is large enough to determine whether a null difference of 42% could be considered a true null group difference or a null difference that may result from insufficient power (false non-significant results).

Missing Cases Analysis and Replacement

All analyses were conducted using SPSS Version 25. Analysis of the rate of missing cases demonstrated that 34.63% of participants who were present at baseline had missing data at immediate follow-up and/or longer-term follow-up. Backward logistic regression models were conducted in order to determine whether any measured baseline variables were significant predictors of missing data at immediate follow-up or longer-term follow-up (Little, 1995; Karin et al., 2018). Enrolment at the Australian Defence Force Academy significantly predicted missing cases at both immediate and longer-term follow-up time points, and were thus considered to be missing conditionally (conditional MAR assumptions). Replacement estimates for missing cases at immediate follow-up and longer-term follow-up were subsequently replaced through a model-based replacement estimation process that took into account the clustered nature of the data, whereby intervention group, platoon, Australian Defence Force Academy enrolment, time, their two-way interactions with time, and baseline outcome scores were included as predictors of outcome scores during the process of missing data estimation (Karin et al., 2018).

To determine whether missing data replacement changed the general pattern of results, we re-ran identical change analyses using the raw un-replaced data. The results of these analyses, reported in Tables S2 and S3, demonstrates that the pattern of findings obtained through the longitudinal models was consistent both with and without the use of model replaced data. The analyses reported in the main manuscript use a dataset with model-based replacement.

Testing Longitudinal Group Change

Longitudinal changes in outcome measures were estimated and compared between the two groups using a series of generalized estimation equation models (GEE; Liang & Zeger, 1986). Within the context of this study, and within similar studies in the military training context (Crane, Boga, Karin, et al., 2019), authors have identified that symptom scores can rapidly increase and then rapidly decrease between the few sampling moments. To account for this pattern, GEE methodology was employed with the use of a working correlation matrix that is unstructured. This method enables us to account for the within-subject correlation of participants between the different time points in a robust manner that: (1) does not require us to make assumptions about the overall trajectory of participants from baseline; (2) is an alternative to more cumbersome models that require quadric or cubic change over time that can be computationally demanding and may inaccurately describe the change patterns over time; and (3) estimate and test for differences in the experimental group means in a way that can flexibly alternate between binary, skewed, and normal score distributions. This approach allows the exploration of nonlinear change in outcome scores over time, whilst also accounting for within-subject variance between different time points.

Given that the intervention condition allocation was randomized by platoon, a potential source of non-independence of data was platoon membership. The researchers also received informal feedback from lead instructional staff at the Royal Military College regarding sub-

group differences in conduct, morale, and performance relating to Cadets from the Australian Defence Force Academy. Our models therefore included additional covariates that were designed to account for the nesting of individuals within platoons and Australian Defence Force Academy enrolment (akin to level 3 nesting in mixed models). Each GEE model included the effect of condition, time, platoon membership, baseline symptoms, Australian Defence Force Academy enrolment, and the two-way interaction with time for condition, platoon membership, and Australian Defence Force Academy enrolment. Effect sizes were reported using Hedge's *g* for between-groups and within-groups effects (Hedges, 1981).

Although our preference is to use GEEs as specified in our pre-registered protocol, we also provide equivalent analyses using multilevel modelling (MLM) as supplementary materials, given that readers may be most familiar with this statistical model. In order to conduct the MLMs we first explored the intra-class correlation for platoon membership. An initial model with no fixed-effects explored the unique variance explained by the random intercept for platoon. Comparable to Crane, Boga, Karin, et al. (2019), platoon membership explained between 0% and 0.12% of the variance in all three predictors. Thus, minimal variance in these outcomes was associated with platoon membership. Subsequently, two-level MLMs were conducted including the random intercept and slope for individuals within platoons. The results of these analyses are reported in Table S4 and S5. The pattern of findings for the MLMs were similar across time to those derived from the GEE models.

Testing for Mechanisms of Symptom Change

To explore possible mechanisms that mediated the change in symptoms, parallel multiple mediation analyses were conducted using Hayes' PROCESS macro for SPSS (Model 4; Hayes, 2013). Parallel mediation models allow the exploration of possible mechanisms simultaneously and a comparison of their effects in context (Hayes, 2013). Three analyses examined whether brooding and reflection explored in parallel mediated the relationship

between intervention condition and depression symptoms, anxiety symptoms, and perceived stress. As demonstrated in Figure 2, parallel mediation was performed using two mediation variables, self-reflection and brooding at longer-term follow-up, in the relationship between intervention condition and the outcome measures at longer-term. In the model, a_n denotes the effect of condition (0 = Control vs. 1 = Intervention) on the mediator variables; b_n denotes the effect of the mediator variables on the outcomes; and *c'* is the direct effect of condition on the outcomes accounting for all other variables in the model. Indirect effects are calculated in the PROCESS macro by multiplying the coefficients relating to the indirect effect (e.g., a_1*b_1; Hayes, 2013). Tests of indirect effects were conducted using 1000 bootstrap samples and 95% bias-corrected confidence intervals. Models also controlled for whether participants were from the Australian Defence Force Academy, baseline outcome variables, and mediator variables at baseline and immediate follow-up. However, platoon membership was not included as its inclusion produced model overfit.

Results

Preliminary Analyses

Preliminary analyses showed no statistically significant differences between the SRT and Control groups at baseline in terms of gender, age, education level, years of service, enrolment at the Australian Defence Force Academy, depression symptoms, anxiety symptoms, or perceived stress scores (see Table 1). A Pearson's chi-square test also found no statistically significant differences between the two groups in terms of Cadets who were lost to follow-up as a result of transferring services, resigning, or being terminated from the Royal Military College. A further chi-square analysis was conducted on all missing participants at immediate and longer-term follow-up. There were no statistically significant differences between groups. Furthermore, correlations between the outcome and mediator variables at each time point for each training condition are reported in Table 2.

A repeated measures ANOVA was used to explore the mean differences in perceived stressor frequency at different time points. The results confirmed a significant effect for time, whereby stressors were reported to be significantly less frequent at baseline (M = 57.80, SE = 1.24) than at immediate follow-up (M = 61.99, SE = 1.13; $p < .001$) and longer-term follow-up (M = 65.60, SE = 1.16; $p < .001$).

Analysis of Primary Outcomes

As specified in the pre-registered protocol, analyses were initially conducted using the PHQ-9 and the GAD-7 as measures of depression symptoms and anxiety symptoms, respectively. These analyses initially did not reveal any statistically significant effects. On examination of the scale items, some items were potentially invalid symptom determinants given that Cadets were subject to periods of reduced food, sleep and personal time in the context of their training; specifically, the items of the PHQ-9 and GAD-7 related to eating (e.g., "Poor appetite or overeating"), sleeping (e.g., "Feeling tired or having little energy") and recreational time (e.g., "Trouble relaxing") lacked construct validity in the military training context. For example, in the case of overeating, Cadets often 'carb-load' prior to an activity where they know there will be restrictions on food as a coping strategy. Moreover, lacking energy is likely to be a consequence of a demanding training environment, rather than symptoms of depression. This is particularly the case at longer-term follow-up, where Cadets had just experienced a very arduous field activity involving these aspects. Thus, analyses were repeated using the PHQ-2 (Kroenke et al., 2003) and the GAD-2 (Kroenke et al., 2007), as these scales specifically focus on the emotional and cognitive aspects of depression and anxiety (PHQ-2: "Little interest or pleasure in doing things", "Feeling down, depressed, or hopeless"; GAD-2: "Feeling anxious or on edge", "Not being able to stop or control worrying"). The PHQ-2 and GAD-2 were likely to be a more sensitive assessment of mental health outcomes in this context and have used in previous military research (e.g., Allison-Aipa et al., 2010;

Corson et al., 2004; Fanning & Pietrzak, 2013; Milliken et al., 2007; Sharma et al., 2017). The PHQ-2 demonstrated satisfactory internal reliability at immediate follow-up and longer-term follow-up, but not at baseline ($\alpha_{Baseline}$ = .63; $\alpha_{Immediate}$ = .76; $\alpha_{Long\text{-}Term}$ = .80). The GAD-2 demonstrated satisfactory internal reliability at all three time points ($\alpha_{Baseline}$ = .82; $\alpha_{Immediate}$ = .85; $\alpha_{Long\text{-}Term}$ = .88).

Depression Symptoms (PHQ-2)

The findings revealed support for H1. Parameter estimates indicated that the rate of increase in depression symptoms between baseline and longer-term follow-up was significantly greater for the Control group than for the SRT group (*Wald* χ^2 = 5.510; p = .019), as was the rate of increase between immediate follow-up and longer-term follow-up (*Wald* χ^2 = 6.084; p = .014). Within-groups pairwise comparisons revealed that the SRT group experienced no statistically significant change in depression symptoms between baseline and immediate follow-up (p = .101), between baseline and longer-term follow-up (p = .178), or between immediate follow-up and longer-term follow-up (p = .884). In contrast, the Control group reported significant *increases* in depression symptoms between baseline (M = 1.95; SE = .07) and immediate follow-up (M = 2.32; SE = .18; p = .049; g = 0.26), between baseline and longer-term follow-up (M = 3.59; SE = .28; p < .001; g = 0.76) and between immediate follow-up (M = 2.32; SE = .18) and longer-term follow-up (M = 3.59; SE = .28; p < .001; g = 0.51). Between-groups pairwise comparisons at individual time points demonstrated that although the mean depression symptoms of the SRT group were not statistically different from the Control group at immediate follow-up (p = .763), the SRT group reported significantly lower depression symptoms (M = 2.46; SE = .25) than the Control group at longer-term follow-up (p = .018; g = -0.41). The estimated mean scores and effect size statistics for depression symptoms by condition are presented in Table 3 and depicted graphically in Figure 3.

Anxiety Symptoms (GAD-2)

The GEE for anxiety symptoms reflected less clear support for H2. Parameter estimates did not identify a difference in the rate of change between baseline and longer-term follow-up (*Wald* χ^2 = 0.523; p = .470). However, the rate of increase in anxiety symptoms between immediate follow-up and longer-term follow-up was significantly greater for the Control group than for the SRT group (*Wald* χ^2 = 8.016; p = .005). The Control group demonstrated no statistically significant mean change in anxiety symptoms between baseline and immediate follow-up, but significant mean increases in symptoms between baseline (M = 1.93; SE = .12) and longer-term follow-up (M = 2.81; SE = .34; p = .032; g = 0.33), and between immediate follow-up (M = 1.67; SE = .23) and longer-term follow-up (M = 2.81; SE = .34; p < .001; g = 0.37). In contrast, the SRT group demonstrated a significant mean *increase* in anxiety symptoms between baseline (M = 1.91; SE = .12) and immediate follow-up (M = 2.65; SE = .26; p = .021; g = 0.36), but no statistically significant change between baseline and longer-term follow-up, nor between immediate follow-up and longer-term follow-up. Further, between group pairwise comparisons demonstrated that, contrary to H2, the SRT group (M = 2.65; SE = .26) reported significantly higher anxiety symptoms than the Control group (M = 1.67; SE = .23) at immediate follow-up (p = .030; g = 0.39), however the two groups were not statistically different at longer-term follow-up. The estimated mean scores and effect size statistics for anxiety symptoms by condition are presented in Table 3 and depicted graphically in Figure 4.

Post-Hoc Analyses

Given that the use of the PHQ-2 and GAD-2 was inconsistent with the scales used in the original trial reported by Crane, Boga, Karin, et al. (2019), we re-ran the GEE analyses with the data collected by Crane, Boga, Karin, et al. (2019) using the PHQ-2 and GAD-2 to permit a comparison between trials. This analysis will help us to determine whether any differences between trials are attributable to the use of different scales or to the modifications

in the intervention program. The results of these analyses, reported in Tables S6 and S7 and depicted graphically in Figures S1 and S2, demonstrated that the findings for the PHQ-2 and GAD-2 for the previous trial were more comparable to the findings for the complete scales, than the findings of the current trial. This indicates that the differences in the pattern of findings between trials is unlikely to be attributable to the scale used.

Perceived Stress (PSS)

As with the GEE for depression symptoms, the findings revealed some support for H3. The rate of increase in perceived stress between baseline and longer-term follow-up was significantly greater for the Control group than for the SRT group (*Wald* χ^2 = 9.329; p = .002). However, there was no statistically significant difference in the rate of change for each group between immediate follow-up and longer-term follow-up (*Wald* χ^2 = 2.612; p = .106). Within-group pairwise comparisons demonstrated that the SRT group experienced no statistically significant mean changes in perceived stress between baseline and immediate follow-up (p = .320) or longer-term follow-up (p = .106), nor between immediate follow-up and longer-term follow-up (p = .534). In contrast, the Control group reported significant increases in perceived stress between baseline (M = 22.08; SE = .15) and immediate follow-up (M = 25.66; SE = .81; p < .001; g = 0.58), between baseline (M = 22.08; SE = .15) and longer-term follow-up (M = 28.44; SE = .94; p < .001; g = 0.95), and between immediate follow-up (M = 25.66; SE = .81) and longer-term follow-up (M = 28.44; SE = .94; p = .001; g = 0.34). Between-group pairwise comparisons at each time point demonstrated that the mean perceived stress scores for the SRT group were not significantly different to the Control group at immediate follow-up (p = .092), yet the SRT group demonstrated significantly lower perceived stress (M = 23.69; SE = .87) than the Control group (M = 28.84; SE = .94) at longer-term follow-up (p = .001; g = -0.55). The estimated mean scores and effect size statistics for perceived stress by condition are presented in Table 3 and depicted graphically in Figure 5.

Parallel Multiple Mediation Analyses

When brooding and reflection at immediate follow-up were used as mediators, no significant indirect effects were identified. However, given that many of the significant changes were emerging at longer-term follow-up, we repeated the mediation analyses using the longer-term follow-up version of the mediating variables, controlling for the corresponding measure at other time points. The results of these analyses are presented in Table 4. For all three outcomes there was a significant indirect effect via brooding at longer-term follow-up. There were no statistically significant indirect effects via self-reflection. Membership in the SRT group predicted lower brooding at longer-term follow-up, and brooding was positively associated with greater depression symptoms, anxiety symptoms, and perceived stress. Interestingly, the SRT group was significantly related to greater self-reflection at longer-term follow-up, but self-reflection was not statistically related to any of the outcome variables.

Discussion

In this study, the findings from the longitudinal analysis of change over time indicated that Cadets in the SRT group demonstrated no increases in depression symptoms and perceived stress over time, even as training stressors reportedly increased. In contrast, the Control group demonstrated increases in depression symptoms and perceived stress, particularly between immediate and longer-term follow-up. At longer-term follow-up, average depression symptoms and perceived stress over time was significantly lower for the SRT group than for the Control group. Although these results are not entirely consistent with the anticipated declines in outcomes for the SRT group across time (H1, H2 and H3), the findings demonstrate that those in the SRT group maintained stable mental health in the context of increasing occupational demands that may have resulted in worsening mental health for the Control group.

As noted previously, this pattern of findings is somewhat different to that obtained in the previous trial (Crane, Boga, Karin, et al., 2019), which was more consistent with

demonstrations of ‘bouncing back’ resilience, whereby an initial increase in symptoms was followed by a symptom decline. In contrast, the pattern of findings observed for the SRT group in the present study is consistent with a pattern of ‘robust’ resilience (Fletcher & Sarkar, 2016), whereby symptoms of depression or perceived stress remain at baseline levels irrespective of increasing training demands.

The results for anxiety symptoms were less consistent with anticipated patterns (H2). Specifically, the SRT group demonstrated an initial *increase* in anxiety symptoms between baseline and immediate follow-up, followed by no statistically significant change between immediate and longer-term follow-up. In contrast, the Control group demonstrated no change in anxiety symptoms between baseline and immediate follow-up, followed by a significant increase in anxiety symptoms between immediate and longer-term follow-up. Consequently, at longer-term follow-up there were no significant between-group differences in anxiety symptoms. There are several possible reasons for this initial increase in anxiety symptoms for those in the SRT group. In order to be effective, reflectors must first draw their attention to some of their more confronting thoughts, emotions and experiences, followed by an intensive process of self-observation and self-analysis (Şimşek, 2013; Takano & Tanno, 2009; Yip, 2006). Dealing with inner conflicts, discrepancies or personal weaknesses during this process may generate initial feelings of anxiety, particularly for individuals operating in a highly demanding psychosocial environment such as Officer training (Atkins & Murphy, 1993; Boyd & Fales, 1983; Schön, 1991; Takano & Tanno, 2009; Yip, 2006). For example, the discrepancy between actions and leadership values may trigger anxious thoughts in this context. Yet, developing or highlighting discrepancies is a key behaviour change technique (e.g., identifying a discrepancy between behavioural performance and a set goal; Michie et al., 2011). Although discrepancy may promote initial uncomfortable feelings, it is a necessary stimulus for change. In this context, in order to reduce felt distress in relation to value discrepancies, Cadets are

encouraged to view their values as aspirational and falling short of personal values is normalized during the SRT program.

A further potential explanation for the observed 'spike' in the SRT group at immediate follow-up might be related to the complex relationship between self-reflection and emotional expression. Previous work has shown an increase in negative emotion initially after writing about traumatic experiences, but that improvements in affect are observed over the course of writing sessions (Murray & Segal, 1994). Other research exploring symptom reports over the course of writing has demonstrated similar initial increases in negative affect immediately after writing about stressful events (Petrie et al., 1995). These researchers describe that initial increases in negative affect are potentially the product of emotional expression, rather than emotional suppression, which may ultimately be positive for both longer-term wellbeing and immune function.

The indirect effects of brooding on all three outcome measures and reflection on anxiety symptoms provide some insight into the mechanisms that potentially underpin the effects of SRT on mental health outcomes. Consistent with H4, there was an indirect relationship between the SRT intervention and all outcomes at longer-term follow-up via brooding. The SRT intervention was associated with lower brooding at longer-term follow-up, which was in turn positively associated with all three outcome measures at longer-term follow-up. This finding is consistent with Gortner et al.'s (2006) evaluation of an expressive writing intervention and the proposal that the self-reflection activity may encourage participants to spend available "thinking time" in adaptive forms of self-reflection rather than in maladaptive brooding.

Contrary to H5, self-reflection did not mediate the relationship between SRT and the outcome measures. However, although the SRT intervention was associated with greater reflection at longer-term follow-up, reflection was not associated with reductions in the

outcome measures. This provides some support for the proposal that the SRT program enhances self-reflection. However, self-reflection did not appear to have a direct relationship with mental health outcomes in this study. There are two salient reasons for the failure to identify a relationship between self-reflection and mental health outcomes. Methodologically, the measure used in this study is a broad measure of reflective thinking that does not capture the motivational orientation, subject, and purpose of reflection that is unique to the adaptive self-reflection proposed in the Systematic Self-Reflection model (Crane, Searle, Kangas, et al., 2019). It is possible that the results of this study may have shown a significant effect as hypothesized if the measure of self-reflection used had contained questions that focused specifically on the motivational orientation, subject, and purpose of a person's self-reflection (e.g., whether the reflections are solution-focused, constructive, future-focused, and centred on growth and learning). Conceptually, the role of self-reflection is to stimulate promising adaptations in resilient capacities (e.g., coping strategies, coping resources). It is via these adaptations in resilient capacities that self-reflection may promote downstream resilience. Thus, a more direct positive outcome of this self-reflective process is perhaps the more flexible use of coping strategies or increased breadth of coping resources. Similar explanations may also account for the mixed findings observed between self-reflection and mental health outcomes in other studies, whereby some studies find that reflective thinking is positively correlated with desirable mental health and wellbeing outcomes (Harrington & Loffredo, 2010; Joireman et al., 2002; Samaie & Farahani, 2011; Takano & Tanno, 2009), while others (e.g., Grant et al., 2002) demonstrate a negative relationship.

In the previous trial of SRT (Crane, Boga, Karin, et al., 2019), Cadets in the SRT group experienced what could be characterized as "bouncing back" resilience, where there was an initial increase in symptoms during a period of high demand, followed by a reduction in the experience of symptoms. In the current trial, the pattern of results over time for depression and

perceived stress were more consistent with a pattern of "robust" resilience, given that there was no significant increase in symptoms across time compared to the Control group, which tended to demonstrate a greater increase in symptoms between immediate and longer-term follow-up.

These differences in the findings for the SRT group between trials could be explained by the previously described changes made to the intervention program, which were intended to improve the training. However, the pattern of findings was also inconsistent for the Control group between trials, suggesting that a mechanism was at play affecting both groups more holistically. One possibility was the change in measurement of outcomes across trials. Rather than measuring the behavioral and somatic dimensions of depression and anxiety, the ultra-brief scales used in this study focus on the measurement of emotional and cognitive aspects of depression and anxiety corresponding to the core diagnostic criteria for depression and generalized anxiety (Kroenke et al., 2010). Thus, these ultra-brief scales may function as a measure of the possible presence of anxiety or depression, rather than severity. In terms of the implications for our findings, limiting the breadth of symptoms measured to emotional and cognitive dimensions may mean that we can only interpret the SRT program as modifying these aspects of symptoms. However, when the data from the previous trial (Crane, Boga, Karin, et al., 2019) was re-analysed using the PHQ-2 and the GAD-2, the pattern of results was similar to those found using the PHQ-8 and GAD-7. This suggests that the change in scales cannot account for the differences observed between studies and indicates that the ultra-brief scales may provide a comparable estimation of depression and anxiety outcomes, as has been observed in previous research (e.g., Kroenke et al., 2010).

Consideration also needs to be given across trials of interventions to the role of organizational system-level factors that influence the effect of training on individual mental health outcomes. It is notable that other research has demonstrated inconsistent effects of resilience training on mental health outcomes for trials of similar programs (e.g., Liossis et al.,

2009; Millear et al., 2008). Like other forms of organizational training, resilience interventions do not occur in a vacuum and their efficacy can be affected by other factors that influence pre- and post- training conditions and therefore the transfer of training to the workplace setting (Kozlowski & Salas, 1997). For example, perceived leader/supervisor endorsement of training practice is a factor affecting training engagement and skill transfer (Colquitt et al., 2000; Sitzmann et al., 2008). Given the routine changes in leadership and Cadet training instructors at the Royal Military College, these factors may have a notable influence of the pattern of training outcomes across trials. However, the nature of the organizational system is rarely considered in the context of resilience interventions. Moving forward, resilience research should consider manipulating, or at least measuring, some of the organizational conditions pre and post-training that may influence training outcomes.

A further factor worth considering is training engagement of participants. This is of particular concern in the military, where personnel are typically required to complete mandatory resilience training (e.g., Cohn & Pakenham, 2008). Participants may not perceive a personal need for training or demonstrate cynicism, which can impair training efficacy (Hurtz & Williams, 2009). A recent study demonstrated the important implications of training engagement in a trial of SRT in older adults (Crane et al., 2020). In this trial, self-reported training engagement moderated the effect of SRT on perceived resilience, particularly for those in the employee group who were encouraged by managers to attend training. Individual levels of engagement in training is often not considered in the context of resilience training, but may have important implications for training efficacy and should be addressed in future research.

Applied Implications

There are several applied implications associated with the present research that have potential value across diverse occupational settings. First, the need for a scalable and effective resilience training program is an issue shared by many organizations. The findings from the

present trial and Crane, Boga, Karin, et al. (2019) provide support for the efficacy of a scalable approach to resilience training that still allows for a higher degree of training personalization. Interventions delivered to large groups may limit training efficacy, given the increased risk of depersonalization (Robertson et al., 2015). The SRT program addresses this risk by encouraging participants to reflect on their personal coping strategies and develop an individualized set of resilience capacities that align with their demand context, strengths, values and goals.

Second, a potential unexplored advantage of this approach is that SRT may encourage participants to reframe stressors as opportunities for growth. The SRT program seeks to achieve this by asking participants to consider what can be learned from stressors encountered and explaining how they can strategically leverage their strengths and resources during exposure to future stressors. Training focused on the development of cognitive-behavioral skills aimed at reducing the negative effects of stress may imply that stress and adversity are inherently problematic, inadvertently promoting a "stress-is-debilitating", rather than a "stress-as-enhancing" mindset (Crum et al., 2013). The potential for stress-as-enhancing mindset to act as an explanatory mechanism for Self-Reflection Resilience Training needs to be investigated in future research. Other psychological processes that serve as mechanisms or moderators, such as reappraisal, cognitive control, coping self-efficacy, optimism, and dispositional hope should also be explored.

Third, the present study demonstrates the importance of a rigorous control group design when evaluating training programs. Most published trials of resilience interventions do not compare intervention outcomes to another group (e.g., Burton et al., 2010; Pipe et al., 2012), used a passive control such as a waitlist control group (e.g., McCraty & Atkinson, 2012; Sood et al., 2011), or compared their intervention to an active control group that was not comparable in terms of treatment length (e.g., Crane, Boga, Karin, et al., 2019; Rose et al., 2013), although

there are exceptions (e.g., Adler et al., 2009; Cohn & Pakenham, 2008). The current outcomes were obtained using an active control group with comparable levels of intervention length, allowing us to infer with greater confidence that it is the content of SRT itself, as opposed to simply a greater quantity of exposure to a psychologist or trainer that is responsible for observed changes in mental health outcomes (Chmitorz et al., 2018).

Finally, this research draws attention to the importance of considering the effect of group-based features (e.g., team dynamics) on the efficacy of training. Platoon membership and enrolment at the Australian Defence Force Academy were important covariates in the analysis of mean change over time. These observations indicate that SRT may be effective at strengthening resilience at the individual-level, yet is not intended to overcome systematic features of culture attributable to group-level characteristics that may have negative effects of wellbeing. Research in sport psychology (Alliger et al., 2015; Chapman et al., 2018; Morgan et al., 2017) and the military (Cacioppo et al., 2011) draws attention to team-level characteristics such as group structure, social capital, and collective efficacy that contribute to team-level resilience. These findings raise important considerations about the training context highlighted by the organization training theory literature (Salas & Cannon-Bowers, 2001), but that are rarely considered in the context of individual resilience training. Future research could be conducted to understand the influence of ecological factors (e.g., antecedent training conditions, team dynamics) on the effectiveness of resilience training.

Limitations and Future Directions

The present study has limitations that should be taken into account during future work. First, methodological considerations may explain why we did not observe the anticipated mediating role of self-reflection. In particular, as noted above, the measure of self-reflection may not have captured the adaptive characteristics of self-reflection that are thought to promote the development of resilience (Crane, Searle, Kangas, et al., 2019). This potential is

underscored by the significant moderate to strong *positive* correlation between brooding and reflection at all time points in the present study. Other work has demonstrated that the outcomes of repetitive thought are distinct, depending on the type of thoughts under investigation (Segerstrom et al., 2003). The non-specific way that self-reflection was measured in this study fails to specify the target of, and motivation for, a person's self-reflection. This is potentially problematic given that several authors suggest that these elements of self-reflection are critical to the outcomes of the reflective process (e.g., Takano & Tanno, 2009; Treynor et al., 2003). Future research is required to develop a measure of adaptive self-reflection that captures the important nuances in motivation (e.g., personal growth) and content (e.g., task-focused) of self-reflection that are likely to have implications for outcomes.

Second, this study was limited by a reliance on subjective outcome measures based on self-report data. The results could be strengthened through the use of objective measures (e.g., cortisol; academic grades). Examining change in these objective outcomes in future investigations may circumvent some of the issues associated with self-report data, including social desirability and recall bias (Beehr, 2019; Rogler et al., 2001).

Third, findings were not statistically significant when analyses were conducted using the full versions of the outcome scales (PHQ-9 and GAD-7). However, the scale items themselves may have limited validity in this context because the Cadets may experience these outcomes in a way that is unrelated to mood and cannot be changed by either the CBST or SRT training. To enhance the ecological validity of measurement, the analyses were re-conducted using ultra-brief versions of the outcome scales (PHQ-2 and GAD-2). Although this issue was not observed during previous trials of SRT (Crane, Boga, Karin, et al., 2019), it should be noted that the current trial is the first of its kind to directly compare SRT to an exposure-matched active control group. This more rigorous study design may have accentuated the need for a more sensitive assessment of depression and anxiety symptoms in this context. Future work is

needed to examine these measures in a context where the constraints (e.g., diet, sleep, recreation) of the current training setting do not exist.

The longer-term follow-up survey occurred after the peak of training demands during a short rest period, but not at the completion of second-class training. This is one possible temporal frame selected to capture both diverse training demands and also the rise and fall of demands. However, as indicated in the analysis of perceived stressor frequency, the Cadets may not have been experiencing the full benefit of the rest period having recently encountered a very demanding field operation exercise and the potential for further training demands. Although the training demands were not entirely alleviated at the final time point, the SRT training still demonstrates potential to enable benefits in the context of risk in comparison to the Control group. Having noted this, future research should endeavour to conduct assessments at the completion of Officer training.

It is the intention of the SRT program that participants continue to engage in systematic self-reflection on stressors after the training has ceased. This may not be in a written form, but as a cognitive process involved in converting major life events into resilience-strengthening experiences to drive the iterative development of resilient capacities as stressor events are encountered (Crane & Boga, 2017; Crane, Searle, Kangas, et al., 2019). However, at present we do not know whether the practice of adaptive self-reflection continued after the formal training ceased. Further research could attempt to address this gap by measuring the continued use of daily self-reflection either as part of follow-up diary studies or the retrospective measurement of the extent to which adaptive reflection has been used post-training.

A further consideration is the generalizability of effects to other populations and workplaces. Cadets at the Royal Military College are generally young, selected for their Officer potential, and highly motivated. These characteristics may possibly make them more amenable to the self-reflection process than those in the general population (Beehr, 2019). Moreover, the

content of SRT is ostensibly metacognitive in nature and requires participants to reflect on their coping strategies and resources in a written format, which may render the program less accessible to individuals with poorer literacy skills or less exposure to academic environments that promote critical thinking.[3] Research in support of this idea indicates that there is a positive association between metacognition and intellectual ability (Van der Stel & Veenman, 2008), and that educational institutions such as high schools and universities play a critical role in the development of metacognitive competencies (Dimmitt & McCormick, 2012) that support reflective processes. Further, both the current trial and its predecessor (Crane, Boga, Karin, et al., 2019) have been conducted with a military population experiencing highly specific and intense stressors that, according to Britt and colleagues (2016) and Vanhove and colleagues (2016), may not be directly comparable to the more cumulative, lower-intensity stressors typically encountered in civilian workplaces. More research is required to investigate whether comparable effects of the SRT program occur in other populations and workplaces.

Opportunities exist to explore how self-reflection can be imparted beyond traditional classroom-based learning. This flexibility is likely to be important for individuals working in action teams such as the military, paramedics, and other first responders. Future work could explore how these small group and one-to-one opportunities for experiential learning could be adapted to reinforce and consolidate the content currently imparted during weekly workbook exercises. Self-reflective practices for strengthening resilience can be enhanced through opportunities for peer group discussion (Riley-Doucet & Wilson, 1997) and direct contact between instructors and participants, as these allow greater opportunities for monitoring and feedback (Kraiger, 2003; Vanhove et al., 2016). Future trials could compare SRT delivered in isolation to SRT accompanied by small group instruction or individualized coaching in order

[3] Officer Cadets in the present study undertake an extensive selection process prior to their entry into the ADF and must meet minimum cognitive aptitude requirements above the general population mean.

to determine whether these additional opportunities for experiential learning increase the effectiveness of the program.

A further area of future research may also involve identifying the possible moderators and theoretical covariates that may moderate training effects, including baseline levels of coping flexibility, coping efficacy and coping resources. With a dedicated study and larger sample size, we can test more complex and nuanced relationships that are currently not the main focus of the present trial. Finally, it will important to further qualify in future research whether or not the self-reflection journals in this study genuinely encourage adaptive outcomes such as self-development and growth, rather than maladaptive outcomes such as brooding. This may be extended via qualitative analyses of the self-reflection journals in order to understand the insights, emerging from self-reflective activity, that strengthen the capacity for resilience.

Concluding Remarks

The findings of this trial offer a valuable insight into the capacity for a resilience training intervention teaching self-reflection on recent stressor events to enhance mental health outcomes in a military setting. The present study extends previous research by testing a modified version of a previously trialled training program using a similar self-reflective approach (Crane, Boga, Karin, et al., 2019) and clarifying the role of two theoretically derived mediators. The findings demonstrate the promise of this practical, sustainable, and scalable approach to resilience training.

References

Adler, A. B., Bliese, P. D., McGurk, D., Hoge, C. W., & Castro, C. A. (2009). Battlemind debriefing and Battlemind training as early interventions with soldiers returning from Iraq: Randomization by platoon. *Journal of Consulting and Clinical Psychology*, *77*(5), 928–940. https://doi.org/10.1037/a0016877

Aldao, A., Sheppes, G., & Gross, J. J. (2015). Emotion regulation flexibility. *Cognitive Therapy and Research*, *39*(3), 263–278. https://doi.org/10.1007/s10608-014-9662-4

Alliger, G. M., Cerasoli, C. P., Tannenbaum, S. I., & Vessey, W. B. (2015). Team resilience: How teams flourish under pressure. *Organizational Dynamics*, *44*(3), 176–184. https://doi.org/10.1016/j.orgdyn.2015.05.003

Allison-Aipa, T. S., Ritter, C., Sikes, P., & Ball, S. (2010). The impact of deployment on the psychological health status, level of alcohol consumption, and use of psychological health resources of postdeployed US Army Reserve soldiers. *Military Medicine*, *175*(9), 630–637. https://doi.org/10.7205/MILMED-D-09-00212

Alloy, L. B., Abramson, L. Y., Hogan, M. E., Whitehouse, W. G., Rose, D. T., Robinson, M. S., Kim, R. S., & Lapkin, J. B. (2000). The Temple-Wisconsin Cognitive Vulnerability to Depression Project: Lifetime history of Axis I psychopathology in individuals at high and low cognitive risk for depression. *Journal of Abnormal Psychology*, *109*(3), 403–418. https://doi.org/10.1037/0021-843X.109.3.403

Angel, C. M. (2016). Resilience, post-traumatic stress, and posttraumatic growth: Veterans' and active duty military members' coping trajectories following traumatic event exposure. *Nurse Education Today*, *47*, 57–60. https://doi.org/10.1016/j.nedt.2016.04.001

Anseel, F., S. Beatty, A., Shen, W., Lievens, F., & R. Sackett, P. (2015). How are we doing after 30 years? A meta-analytic review of the antecedents and outcomes of feedback-

seeking behavior. *Journal of Management*, *41*(1), 318–348. https://doi.org/10.1177/0149206313484521

Atkins, S., & Murphy, K. (1993). Reflection: A review of the literature. *Journal of Advanced Nursing*, *18*(8), 1188–1192. https://doi.org/10.1046/j.1365-2648.1993.18081188.x

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Beehr, T. A. (2019). Interventions in occupational health psychology. *Journal of Occupational Health Psychology*, *24*(1), 1–3. https://doi.org/10.1037/ocp0000140

Bonanno, G. A. (2005). Resilience in the face of potential trauma. *Current Directions in Psychological Science*, *14*(3), 135–138. https://doi.org/10.1111/j.0963-7214.2005.00347.x

Bonanno, G. A., & Burton, C. L. (2013). Regulatory flexibility: An individual differences perspective on coping and emotion regulation. *Perspectives on Psychological Science*, *8*(6), 591–612. https://doi.org/10.1177/1745691613504116

Bonanno, G. A., Galea, S., Bucciarelli, A., & Vlahov, D. (2006). Psychological resilience after disaster: New York City in the aftermath of the September 11th terrorist attack. *Psychological Science*, *17*(3), 181–186. https://doi.org/10.1111/j.1467-9280.2006.01682.x

Bonanno, G. A., Galea, S., Bucciarelli, A., & Vlahov, D. (2007). What predicts psychological resilience after disaster? The role of demographics, resources, and life stress. *Journal of Consulting and Clinical Psychology*, *75*(5), 671–682. https://doi.org/10.1037/0022-006X.75.5.671

Bonanno, G. A., Westphal, M., & Mancini, A. D. (2011). Resilience to loss and potential trauma. *Annual Review of Clinical Psychology*, *7*(1), 511–535. https://doi.org/10.1146/annurev-clinpsy-032210-104526

Boyd, E. M., & Fales, A. W. (1983). Reflective learning: Key to learning from experience. *Journal of Humanistic Psychology*, *23*(2), 99–117. https://doi.org/10.1177/0022167883232011

Britt, T., Shen, W., Sinclair, R., R. Grossman, M., & Klieger, D. (2016). How much do we really know about employee resilience? *Industrial and Organizational Psychology*, *9*(2), 378–404. https://doi.org/10.1017/iop.2015.107

Burton, N. W., Pakenham, K. I., & Brown, W. J. (2010). Feasibility and effectiveness of psychosocial resilience training: A pilot study of the READY program. *Psychology, Health & Medicine*, *15*(3), 266–277. https://doi.org/10.1080/13548501003758710

Cacioppo, J. T., Reis, H. T., & Zautra, A. J. (2011). Social resilience: The value of social fitness with an application to the military. *American Psychologist*, *66*(1), 43–51. https://doi.org/10.1037/a0021419

Cacioppo, J., Adler, A., Lester, P., McGurk, D., Thomas, J., Chen, H. Y., & Cacioppo, S. (2015). Building social resilience in soldiers: A double dissociative randomized controlled study. *Journal of Personality and Social Psychology*, *109*(1), 90–105. https://doi.org/10.1037/pspi0000022

Casper, A., Tremmel, S., & Sonnentag, S. (2019). Patterns of positive and negative work reflection during leisure time: A latent profile analysis. *Journal of Occupational Health Psychology*, *24*(5), 527–542. https://doi.org/10.1037/ocp0000142

Castro, C. A., Adler, A. B., McGurk, D., & Bliese, P. D. (2012). Mental health training with soldiers four months after returning from Iraq: Randomization by platoon. *Journal of Traumatic Stress*, *25*(4), 376–383. https://doi.org/10.1002/jts.21721

Ceschi, A., Fraccaroli, F., Costantini, A., & Sartori, R. (2017). Turning bad into good: How resilience resources protect organizations from demanding work environments. *Journal*

of Workplace Behavioral Health, *32*(4), 267–289. https://doi.org/10.1080/15555240.2017.1398659

Chapman, M. T., Lines, R. L. J., Crane, M., Ducker, K. J., Ntoumanis, N., Peeling, P., Parker, S. K., Quested, E., Temby, P., Thøgersen-Ntoumani, C., & Gucciardi, D. F. (2018). Team resilience: A scoping review of conceptual and empirical work. *Work & Stress*, 1–25. https://doi.org/10.1080/02678373.2018.1529064

Chesney, M. A., Neilands, T. B., Chambers, D. B., Taylor, J. M., & Folkman, S. (2006). A validity and reliability study of the coping self-efficacy scale. *British Journal of Health Psychology*, *11*(3), 421–437. https://doi.org/10.1348/135910705X53155

Chmitorz, A., Kunzler, A., Helmreich, I., Tüscher, O., Kalisch, R., Kubiak, T., Wessa, M., & Lieb, K. (2018). Intervention studies to foster resilience: A systematic review and proposal for a resilience framework in future intervention studies. *Clinical Psychology Review*, *59*, 78–100. https://doi.org/10.1016/j.cpr.2017.11.002

Ciarocco, N., Vohs, K., & Baumeister, R. (2010). Some good news about rumination: Task-focused thinking after failure facilitates performance improvement. *Journal of Social and Clinical Psychology*, *29*(10), 1057–1073. https://doi.org/10.1521/jscp.2010.29.10.1057

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69–93). Guilford Press.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396. https://doi.org/10.2307/2136404

Cohn, A., Crane, M. F., & Hodson, S. E. (2011). Psychological resilience training in the Australian Defence Force. In K. M. Gow & M. J. Celinski (Eds.), *Wayfinding through life's challenges: Coping and survival*. Nova Science Publishers.

Cohn, A., & Pakenham, K. (2008). Efficacy of a cognitive-behavioral program to improve psychological adjustment among soldiers in recruit training. *Military Medicine*, *173*(12), 1151–1157. https://doi.org/10.7205/MILMED.173.12.1151

Colquitt, J. A., LePine, J. A., & Noe, R. A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*, *85*(5), 678–707. https://doi.org/10.1037/0021-9010.85.5.678

Cook, J. A., Julious, S. A., Sones, W., Hampson, L. V, Hewitt, C., Berlin, J. A., Ashby, D., Emsley, R., Fergusson, D. A., Walters, S. J., Wilson, E. C. F., MacLennan, G., Stallard, N., Rothwell, J. C., Bland, M., Brown, L., Ramsay, C. R., Cook, A., Armstrong, D., … Vale, L. D. (2018). DELTA2 guidance on choosing the target difference and undertaking and reporting the sample size calculation for a randomised controlled trial. *BMJ*, *363*, k3750. https://doi.org/10.1136/bmj.k3750

Cornum, R., Matthews, M. D., & Seligman, M. E. P. (2011). Comprehensive Soldier Fitness: Building resilience in a challenging institutional context. *American Psychologist*, *66*(1), 4–9. https://doi.org/10.1037/a0021420

Corson, K., Gerrity, M. S., & Dobscha, S. K. (2004). Screening for depression and suicidality in a VA primary care setting: 2 items are better than 1 item. *American Journal of Managed Care*, *10*(11), 839–845.

Crane, M. F., Searle, B. J., Kangas, M., & Nwiran, Y. (2019). How resilience is strengthened by exposure to stressors: The systematic self-reflection model of resilience strengthening. *Anxiety, Stress and Coping*, *32*(1), 1–17. https://doi.org/10.1080/10615806.2018.1506640

Crane, M., & Boga, D. (2017). A commentary: Rethinking approaches to resilience and mental health training. *Journal of Military and Veterans Health*, *25*(1), 30–33.

Crane, M. F., Boga, D., Karin, E., Gucciardi, D. F., Rapport, F., Callen, J., & Sinclair, L. (2019). Strengthening resilience in military officer cadets: A group-randomized controlled trial of coping and emotion regulatory self-reflection training. *Journal of Consulting and Clinical Psychology*, *87*(2), 125–140. https://doi.org/10.1037/ccp0000356

Crane, M. F., Kho, M., Kangas, M., Griffin, B., Karin, E., Earl, J. K., & Harris, C. B. (2020). Strengthening resilience in over 50's: A nested clustered-randomized controlled trial of adaptive systematic self-reflection. *Anxiety, Stress, & Coping*, 1–19. https://doi.org/10.1080/10615806.2020.1768375

Crane, M. F., & Searle, B. J. (2016). Building resilience through exposure to stressors: The effects of challenges versus hindrances. *Journal of Occupational Health Psychology*, *21*(4), 468–479. https://doi.org/10.1037/a0040064

Cranton, P. (2006). Fostering authentic relationships in the transformative classroom. *New Directions for Adult and Continuing Education*, *2006*(109), 5–13. https://doi.org/10.1002/ace.203

Crum, A. J., Salovey, P., & Achor, S. (2013). Rethinking stress: The role of mindsets in determining the stress response. *Journal of Personality and Social Psychology*, *104*(4), 716–733. https://doi.org/10.1037/a0031201

Demsky, C. A., Fritz, C., Hammer, L. B., & Black, A. E. (2019). Workplace incivility and employee sleep: The role of rumination and recovery experiences. *Journal of Occupational Health Psychology*, *24*(2), 228–240. https://doi.org/10.1037/ocp0000116

Dimmitt, C., & McCormick, C. B. (2012). Metacognition in education. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Volume 1: Theories, constructs, and critical issues* (pp. 157–187). American Psychological Association. https://doi.org/10.1037/13273-007

Donohue, M. C., & Edland, S. D. (2016). *Longpower: Power and sample size calculators for longitudinal data* (R package version 1.0-16).

Ellis, S., Carette, B., Anseel, F., & Lievens, F. (2014). Systematic reflection: Implications for learning from failures and successes. *Current Directions in Psychological Science*, *23*(1), 67–72. https://doi.org/10.1177/0963721413504106

Fanning, J. R., & Pietrzak, R. H. (2013). Suicidality among older male veterans in the United States: Results from the National Health and Resilience in Veterans Study. *Journal of Psychiatric Research*, *47*(11), 1766–1775. https://doi.org/10.1016/j.jpsychires.2013.07.015

Fletcher, D., & Sarkar, M. (2016). Mental fortitude training: An evidence-based approach to developing psychological resilience for sustained success. *Journal of Sport Psychology in Action*, *7*(3), 135–157. https://doi.org/10.1080/21520704.2016.1255496

Glass, O., Dreusicke, M., Evans, J., Bechard, E., & Wolever, R. Q. (2019). Expressive writing to improve resilience to trauma: A clinical feasibility trial. *Complementary Therapies in Clinical Practice*, *34*, 240–246. https://doi.org/10.1016/j.ctcp.2018.12.005

Goldin, P., Morrison, A., Jazaieri, H., Brozovich, F., Heimberg, R., & J Gross, J. (2016). Group CBT versus MBSR for Social Anxiety Disorder: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *84*(5), 427–437. https://doi.org/10.1037/ccp0000092

Goldstein, I. L. (1993). *Training in organizations: Needs assessment, development and evaluation* (3rd ed.). Brooks/Cole.

Gortner, E. M., Rude, S., & Pennebaker, J. (2006). Benefits of expressive writing in lowering rumination and depressive symptoms. *Behavior Therapy*, *37*(3), 292–303. https://doi.org/10.1016/j.beth.2006.01.004

Grant, A. M., Franklin, J., & Langford, P. (2002). The Self-Reflection and Insight Scale: A new measure of private self-consciousness. *Social Behavior and Personality: An International Journal*, *30*(8), 821–835. https://doi.org/10.2224/sbp.2002.30.8.821

Harms, P., Herian, M., V. Krasikova, D., Vanhove, A., & Lester, P. (2013). *The Comprehensive Soldier and Family Fitness Program Evaluation. Report #4: Evaluation of resilience training and mental and behavioral health outcomes*. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?referer=https://scholar.google.com.a%0Au/&httpsredir=1&article=1009&context=pdharms

Harrington, R., & Loffredo, D. A. (2010). Insight, rumination, and self-reflection as predictors of well-being. *The Journal of Psychology: Interdisciplinary and Applied*, *145*(1), 39–57. https://doi.org/10.1080/00223980.2010.528072

Harris, R. (2009). *ACT made simple*. New Harbinger Publications.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis*. Guilford Press.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Hurtz, G. M., & Williams, K. J. (2009). Attitudinal and motivational antecedents of participation in voluntary employee development activities. *Journal of Applied Psychology*, *94*(3), 635–653. https://doi.org/10.1037/a0014580

Joireman, J. A., Parrott III, L., & Hammersla, J. (2002). Empathy and the self-absorption paradox: Support for the distinction between self-rumination and self-reflection. *Self and Identity*, *1*(1), 53–65. https://doi.org/10.1080/152988602317232803

Joyce, S., Shand, F., Tighe, J., Laurent, S. J., Bryant, R. A., & Harvey, S. B. (2018). Road to resilience: A systematic review and meta-analysis of resilience training programmes and interventions. *BMJ Open*, *8*(6), e017858. https://doi.org/10.1136/bmjopen-2017-017858

Kalisch, R., Baker, D., Basten, U., Boks, M., Bonanno, G., Brummelman, E., Chmitorz, A., Fernàndez, G., J. Fiebach, C., Galatzer-Levy, I., Geuze, E., Groppa, S., Helmreich, I., Hendler, T., J. Hermans, E., Jovanovic, T., Kubiak, T., Lieb, K., Lutz, B., & Kleim, B. (2017). The resilience framework as a strategy to combat stress-related disorders. *Nature Human Behaviour*, *1*(11), 784–790. https://doi.org/10.1038/s41562-017-0200-8

Kalisch, R., Müller, M. B., & Tüscher, O. (2015). A conceptual framework for the neurobiological study of resilience. *Behavioral and Brain Sciences*, *38*, e92. https://doi.org/10.1017/S0140525X1400082X

Karin, E., Dear, B. F., Heller, G. Z., Crane, M. F., & Titov, N. (2018). "Wish you were here": Examining characteristics, outcomes, and statistical solutions for missing cases in web-based psychotherapeutic trials. *JMIR Mental Health*, *5*(2), e22. https://doi.org/10.2196/mental.8363

Kozlowski, S. W. J., & Salas, E. (1997). A multilevel organizational systems approach for the implementation and transfer of training. In J. K. Ford (Ed.), *Improving training effectiveness in work organizations* (pp. 247–290). Taylor & Francis.

Kraiger, K. (2003). Perspectives on training and development. In I. B. Weiner (Ed.), *Handbook of Psychology*. https://doi.org/10.1002/0471264385.wei1208

Kray, L., D. Galinsky, A., & Markman, K. (2009). Counterfactual structure and learning from experience in negotiations. *Journal of Experimental Social Psychology*, *45*(4), 979–982. https://doi.org/10.1016/j.jesp.2009.03.008

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, *16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care*, *41*(11), 1284–1292. https://doi.org/10.1097/01.MLR.0000093487.78664.3C

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2010). The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: A systematic review. *General Hospital Psychiatry*, *32*(4), 345–359. https://doi.org/10.1016/j.genhosppsych.2010.03.006

Kroenke, K., Spitzer, R. L., Williams, J. B. W., Monahan, P. O., & Lowe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, *146*(5), 317–325. https://doi.org/10.7326/003-4819-146-5-200703060-00004

Leppin, A. L., Bora, P. R., Tilburt, J. C., Gionfriddo, M. R., Zeballos-Palacios, C., Dulohery, M. M., Sood, A., Erwin, P. J., Brito, J. P., Boehmer, K. R., & Montori, V. M. (2014). The efficacy of resiliency training programs: A systematic review and meta-analysis of randomized trials. *PLoS ONE*, *9*(10), e111420. https://doi.org/10.1371/journal.pone.0111420

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22. https://doi.org/10.1093/biomet/73.1.13

Liossis, P. L., Shochet, I., Millear, P., & Biggs, H. (2009). The Promoting Adult Resilience (PAR) program: The effectiveness of the second, shorter pilot of a workplace prevention program. *Behaviour Change*, *26*(2), 97–112. https://doi.org/10.1375/bech.26.2.97

Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, *90*(431), 1112–1121. https://doi.org/10.1080/01621459.1995.10476615

Masten, A. S. (2007). Resilience in developing systems: Progress and promise as the fourth wave rises. *Development and Psychopathology*, *19*(3), 921–930. https://doi.org/10.1017/S0954579407000442

McCraty, R., & Atkinson, M. (2012). Resilience training program reduces physiological and psychological stress in police officers. *Global Advances in Health and Medicine*, *1*(5), 44–66. https://doi.org/10.7453/gahmj.2012.1.5.013

Mezirow, J. (1997). Transformative learning: Theory to practice. *New Directions for Adult and Continuing Education*, *1997*(74), 5–12. https://doi.org/10.1002/ace.7401

Mezirow, J. (1991). *Transformative dimensions of adult learning*. Jossey Bass.

Michie, S., Ashford, S., Sniehotta, F. F., Dombrowski, S. U., Bishop, A., & French, D. P. (2011). A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: The CALO-RE taxonomy. *Psychology & Health*, *26*(11), 1479–1498. https://doi.org/10.1080/08870446.2010.540664

Millear, P., Liossis, P., Shochet, I., Biggs, H., & Donald, M. (2008). Being on PAR: Outcomes of a pilot trial to improve mental health and wellbeing in the workplace with the Promoting Adult Resilience (PAR) program. *Behaviour Change*, *25*(4), 215–228. https://doi.org/10.1375/bech.25.4.215

Milliken, C. S., Auchterlonie, J. L., & Hoge, C. W. (2007). Longitudinal assessment of mental health problems among active and reserve component soldiers returning from the Iraq war. *JAMA: The Journal of the American Medical Association*, *298*(18), 2141–2148. https://doi.org/10.1001/jama.298.18.2141

Miranda, R., & Nolen-Hoeksema, S. (2007). Brooding and reflection: Rumination predicts suicidal ideation at 1-year follow-up in a community sample. *Behaviour Research and Therapy*, *45*(12), 3088–3095. https://doi.org/10.1016/j.brat.2007.07.015

Moore, L. J., Young, T., Freeman, P., & Sarkar, M. (2018). Adverse life events, cardiovascular responses, and sports performance under pressure. *Scandinavian Journal of Medicine & Science in Sports*, *28*(1), 340–347. https://doi.org/10.1111/sms.12928

Morgan, P. B. C., Fletcher, D., & Sarkar, M. (2017). Recent developments in team resilience research in elite sport. *Current Opinion in Psychology*, *16*, 159–164. https://doi.org/10.1016/j.copsyc.2017.05.013

Murray, E. J., & Segal, D. L. (1994). Emotional processing in vocal and written expression of feelings about traumatic experiences. *Journal of Traumatic Stress*, *7*(3), 391–405. https://doi.org/10.1007/BF02102784

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology*, *100*(4), 569–582. https://doi.org/10.1037/0021-843X.100.4.569

O'Connor, R. C., & Williams, J. M. G. (2014). The relationship between positive future thinking, brooding, defeat and entrapment. *Personality and Individual Differences*, *70*, 29–34. https://doi.org/10.1016/j.paid.2014.06.016

Petrie, K. J., Booth, R. J., Pennebaker, J. W., Davison, K. P., & Thomas, M. G. (1995). Disclosure of trauma and immune response to a hepatitis B vaccination program. *Journal of Consulting and Clinical Psychology*, *63*(5), 787–792. https://doi.org/10.1037//0022-006x.63.5.787

Pipe, T. B., Buchda, V. L., Launder, S., Hudak, B., Hulvey, L., Karns, K. E., & Pendergast, D. (2012). Building personal and professional resources of resilience and agility in the healthcare workplace. *Stress and Health*, *28*(1), 11–22. https://doi.org/10.1002/smi.1396

Richardson, G. E. (2002). The metatheory of resilience and resiliency. *Journal of Clinical Psychology*, *58*(3), 307–321. https://doi.org/10.1002/jclp.10020

Riley-Douchet, C., & Wilson, S. (1997). A three-step method of self-reflection using reflective journal writing. *Journal of Advanced Nursing*, *25*(5), 964–968. https://doi.org/10.1046/j.1365-2648.1997.1997025964.x

Robertson, I. T., Cooper, C. L., Sarkar, M., & Curran, T. (2015). Resilience training in the workplace from 2003 to 2014: A systematic review. *Journal of Occupational and Organizational Psychology*, *88*(3), 533–562. https://doi.org/10.1111/joop.12120

Rogler, L. H., Mroczek, D. K., Fellows, M., & Loftus, S. T. (2001). The neglect of response bias in mental health research. *Journal of Nervous and Mental Disease*, *189*(3), 182–187. https://doi.org/10.1097/00005053-200103000-00007

Rose, R. D., Buckey, J. C., Zbozinek, T. D., Motivala, S. J., Glenn, D. E., Cartreine, J. A., & Craske, M. G. (2013). A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. *Behaviour Research and Therapy*, *51*(2), 106–112. https://doi.org/10.1016/j.brat.2012.11.003

Rutter, M. (2006). Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Sciences*, *1094*(1), 1–12. https://doi.org/10.1196/annals.1376.002

Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, *52*(1), 471–499. https://doi.org/10.1146/annurev.psych.52.1.471

Samaie, G., & Farahani, H. (2011). Self-compassion as a moderator of the relationship between rumination, self-reflection and stress. *Procedia - Social and Behavioral Sciences*, *30*, 978–982. https://doi.org/10.1016/j.sbspro.2011.10.190

Sarkar, M., & Fletcher, D. (2017). Adversity-related experiences are essential for Olympic success: Additional evidence and considerations. *Progress in Brain Research*, *232*, 159–165. https://doi.org/10.1016/bs.pbr.2016.11.009

Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): A reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, *67*(6), 1063–1078. https://doi.org/10.1037//0022-3514.67.6.1063

Schön, D. A. (1991). *The reflective turn: Case studies in and on educational practice*. Teachers College Press.

Schoofs, H., Hermans, D., & Raes, F. (2010). Brooding and reflection as subtypes of rumination: Evidence from confirmatory factor analysis in nonclinical samples using the Dutch Ruminative Response Scale. *Journal of Psychopathology and Behavioral Assessment*, *32*(4), 609–617. https://doi.org/10.1007/s10862-010-9182-9

Seery, M. D., Holman, E. A., & Silver, R. C. (2010). Whatever does not kill us: Cumulative lifetime adversity, vulnerability, and resilience. *Journal of Personality and Social Psychology*, *99*(6), 1025–1041. https://doi.org/10.1037/a0021344

Seery, M. D., Leo, R. J., Lupien, S. P., Kondrak, C. L., & Almonte, J. L. (2013). An upside to adversity? Moderate cumulative lifetime adversity is associated with resilient responses in the face of controlled stressors. *Psychological Science*, *24*(7), 1181–1189. https://doi.org/10.1177/0956797612469210

Segerstrom, S. C., Stanton, A. L., Alden, L. E., & Shortridge, B. E. (2003). A multidimensional structure for repetitive thought: What's on your mind, and how, and how much? *Journal of Personality and Social Psychology*, *85*(5), 909–921. https://doi.org/10.1037/0022-3514.85.5.909

Sharma, V., Marin, D. B., Koenig, H. K., Feder, A., Iacoviello, B. M., Southwick, S. M., & Pietrzak, R. H. (2017). Religion, spirituality, and mental health of US military veterans: Results from the National Health and Resilience in Veterans Study. *Journal of Affective Disorders*, *217*, 197–204. https://doi.org/10.1016/j.jad.2017.03.071

Shatté, A., Perlman, A., Smith, B., & Lynch, W. D. (2017). The positive effect of resilience on stress and business outcomes in difficult work environments. *Journal of Occupational and Environmental Medicine*, *59*(2), 135–140. https://doi.org/10.1097/JOM.0000000000000914

Şimşek, Ö. F. (2013). Self-absorption paradox is not a paradox: Illuminating the dark side of self-reflection. *International Journal of Psychology*, *48*(6), 1109–1121. https://doi.org/10.1080/00207594.2013.778414

Sitzmann, T., Brown, K. G., Casper, W. J., Ely, K., & Zimmerman, R. D. (2008). A review and meta-analysis of the nomological network of trainee reactions. *Journal of Applied Psychology*, *93*(2), 280–295. https://doi.org/10.1037/0021-9010.93.2.280

Smith, B. W., Tooley, E. M., Christopher, P. J., & Kay, V. S. (2010). Resilience as the ability to bounce back from stress: A neglected personal resource? *The Journal of Positive Psychology*, *5*(3), 166–176. https://doi.org/10.1080/17439760.2010.482186

Snyder, C. R. (1995). Conceptualizing, measuring, and nurturing hope. *Journal of Counseling & Development*, *73*(3), 355–360. https://doi.org/10.1002/j.1556-6676.1995.tb01764.x

Sood, A., Prasad, K., Schroeder, D., & Varkey, P. (2011). Stress management and resilience training among Department of Medicine faculty: A pilot randomized clinical trial. *Journal of General Internal Medicine*, *26*(8), 858–861. https://doi.org/10.1007/s11606-011-1640-x

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Takano, K., & Tanno, Y. (2009). Self-rumination, self-reflection, and depression: Self-rumination counteracts the adaptive effect of self-reflection. *Behaviour Research and Therapy*, *47*(3), 260–264. https://doi.org/10.1016/j.brat.2008.12.008

Trapnell, P. D., & Campbell, J. D. (1999). Private self-consciousness and the five-factor model of personality: Distinguishing rumination from reflection. *Journal of Personality and Social Psychology*, *76*(2), 284–304. https://doi.org/10.1037/0022-3514.76.2.284

Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination reconsidered: A psychometric analysis. *Cognitive Therapy and Research*, *27*(3), 247–259. https://doi.org/10.1023/A:1023910315561

Tusting, K., & Barton, D. (2003). *Models of adult learning: A literature review*. National Research and Development Centre for Adult Literacy and Numeracy.

Valentine, J., Biglan, A., Boruch, R., Castro, F., Collins, L., Flay, B., Kellam, S., Moscicki, E., & Schinke, S. (2011). Replication in prevention science. *Prevention Science*, *12*(2), 103–117. https://doi.org/10.1007//s11121-011-0217-6

Van der Stel, M., & Veenman, M. V. J. (2008). Relation between intellectual ability and metacognitive skillfulness as predictors of learning performance of young students performing tasks in different domains. *Learning and Individual Differences*, *18*(1), 128–134. https://doi.org/10.1016/j.lindif.2007.08.003

Van Vugt, M. K., & Van der Velde, M. (2018). How does rumination impact cognition? A first mechanistic model. *Topics in Cognitive Science*, *10*(1), 175–191. https://doi.org/10.1111/tops.12318

Vanhove, A. J., Herian, M. N., Perez, A. L. U., Harms, P. D., & Lester, P. B. (2016). Can resilience be developed at work? A meta-analytic review of resilience-building programme effectiveness. *Journal of Occupational and Organizational Psychology*, *89*(2), 278–307. https://doi.org/10.1111/joop.12123

Waite, P. J., & Richardson, G. E. (2004). Determining the efficacy of resiliency training in the work site. *Journal of Allied Health*, *33*(3), 178–183.

Wells, A., & Matthews, G. (1996). Modelling cognition in emotional disorder: The S-REF model. *Behaviour Research and Therapy*, *34*(11–12), 881–888. https://doi.org/10.1016/S0005-7967(96)00050-2

Yang, M. M., Zhang, Y., & Yang, F. (2018). How a reflection intervention improves the effect of learning goals on performance outcomes in a complex decision-making task. *Journal of Business and Psychology*, *33*(5), 579–593. https://doi.org/10.1007/s10869-017-9510-0

Yip, K. (2006). Self-reflection in reflective practice: A note of caution. *British Journal of Social Work*, *36*(5), 777–788. https://doi.org/10.1093/bjsw/bch323

Table 1: Demographic characteristics by training condition.

	Self-Reflection Resilience Training (SRT) Group		Control Group			
Variable Name	***n***	**% or Mean**	***n***	**% or Mean**	***t* or χ^2**	***p*-values**
Gender						
Male	82	85.42	94	87.04	0.11	.74
Female	14	14.58	14	12.96		
Age (in years)	96	*M* = 23.18	108	*M* = 23.18	0.01	.99
Education level						
Completed Year 11 (High School)	0	n/a	1	0.93	1.90	.93
Completed Year 12 (High School)	14	14.74	21	19.63		
Some trade qualification	2	2.11	2	1.87		
completed	12	12.63	12	11.21		
Some university completed	6	6.32	7	6.54		
Completed trade qualification	5	5.26	6	5.61		
Currently enrolled in university	56	58.95	58	54.21		
Completed university degree						
Years of Service	92	*M* = 2.42	102	*M* = 2.39	-0.13	.90
Enrolment at Australian Defence						
Force Academy	48	50.00	48	44.44	0.63	.43
Yes	48	50.00	60	55.56		
No						
Depression (Baseline)						
PHQ-2	96	*M* = 2.06	108	*M* = 1.94	-0.57	.57
Anxiety (Baseline)						
GAD-2	96	*M* = 2.02	108	*M* = 1.83	-0.73	.47
Baseline Perceived Stress (Baseline)	96	*M* = 22.22	108	*M* = 21.96	-0.25	.80

Table 2: Correlations at each time point for each training condition. Correlations above the diagonal refer to the Self-Reflection Resilience Training (SRT) group, whereas correlations below the diagonal refer to the Control group. * $p < .05$; ** $p < .01$ (two-tailed)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. PHQ-2 at Baseline	1	.700**	.311**	.693**	.392**	.188	.643**	.532**	.395**	.572**	.513**	.232*	.387**	.309**	.187
2. PHQ-2 at Immediate Follow-Up	.624**	1	.374**	.641**	.650**	.373**	.484**	.640**	.510**	.470**	.582**	.376**	.351**	.422**	.258*
3. PHQ-2 at Longer-Term Follow-Up	.156	.272**	1	.282**	.264**	.727**	.181	.338**	.558**	.299**	.292**	.619**	.205*	.141	.481**
4. GAD-2 at Baseline	.636**	.483**	.205*	1	.431**	.280**	.734**	.613**	.455**	.701**	.579**	.390**	.466**	.291**	.408**
5. GAD-2 at Immediate Follow-Up	.354**	.610**	.147	.406**	1	.373**	.288**	.504**	.301**	.349**	.517**	.296**	.224*	.361**	.229*
6. GAD-2 at Longer-Term Follow-Up	.086	.251**	.643**	.233*	.281**	1	.178	.333**	.589**	.305**	.304**	.609**	.131	.156	.455**
7. Perceived Stress at Baseline	.613**	.452**	.167	.638**	.387**	.181	1	.590**	.542**	.727**	.503**	.352**	.525**	.294**	.370**
8. Perceived stress at Immediate Follow-Up	.591**	.641**	.345**	.550**	.517**	.302**	.766**	1	.643**	.591**	.660**	.488**	.451**	.389**	.428**
9. Perceived stress at Longer-Term Follow-Up	.301**	.297**	.604**	.433**	.243*	.547**	.614**	.605**	1	.573**	.588**	.696**	.375**	.397**	.587**
10. Brooding at Baseline	.580**	.491**	.114	.658**	.354**	.141	.671**	.546**	.438**	1	.851**	.672**	.653**	.465**	.563**
11. Brooding at Immediate Follow-Up	.606**	.655**	.176	.530**	.594**	.169	.654**	.670**	.395**	.730**	1	.597**	.534**	.617**	.504**
12. Brooding at Longer-Term Follow-Up	.324**	.480**	.586**	.424**	.348**	.568**	.409**	.507**	.631**	.503**	.499**	1	.481**	.319**	.682**
13. Reflection at Baseline	.323**	.176	.101	.432**	.129	.100	.450**	.235*	.380**	.671**	.397**	.297**	1	.587**	.601**
14. Reflection at Immediate Follow-Up	.318**	.214*	.181	.420**	.250**	.216*	.499**	.358**	.424**	.533**	.552**	.351**	.720**	1	.539**
15. Reflection at Longer-Term Follow-Up	.205*	.203*	.387**	.400**	.023	.458**	.351**	.283**	.439**	.448**	.247*	.610**	.534**	.449**	1

Table 3: Adjusted means, standard errors and average effect size of within-group and between-group change in depression symptoms, anxiety symptoms and perceived stress for the Control group and the Self-Reflection Resilience Training (SRT) group.

	Means (Standard Errors) for each Time Point			**Within-Group Effect Sizes calculated using Hedge's *g* [95% Confidence Interval]**			**Between-Group Effect Sizes calculated using Hedge's *g* [95% Confidence Interval]**	
Primary Outcomes	**Baseline**	**Immediate follow-up**	**Longer-term follow-up**	**Baseline to Immediate follow-up**	**Baseline to Longer-term follow-up**	**Immediate to Longer-term follow-up**	**Immediate follow-up (Control vs. SRT)**	**Longer-term follow-up (Control vs. SRT)**
Depression symptoms (PHQ-2)								
Control	1.95 (0.07)	2.32 (0.18)	3.59 (0.28)	0.26 [-0.02, 0.53]	0.76 [0.48, 1.04]	0.51 [0.23, 0.79]	0.05 [-0.23, 0.33]	-0.41 [-0.7, -0.13]
SRT	2.09 (0.08)	2.42 (0.20)	2.46 (0.25)	0.22 [-0.06, 0.51]	0.20 [-0.09, 0.49]	0.02 [-0.27, 0.31]		
Anxiety symptoms (GAD-2)								
Control	1.93 (0.12)	1.67 (0.23)	2.81 (0.34)	-0.14 [-0.41, 0.13]	0.33 [0.05, 0.60]	0.37 [0.10, 0.65]	0.39 [0.11, 0.67]	-0.16 [-0.44, 0.12]
SRT	1.91 (0.12)	2.65 (0.26)	2.26 (0.32)	0.36 [0.07, 0.66]	0.15 [-0.14, 0.44]	-0.13 [-0.42, 0.15]		
Perceived stress								
Control	22.08 (0.15)	25.66 (0.81)	28.84 (0.94)	0.58 [0.30, 0.86]	0.95 [0.66, 1.24]	0.34 [0.07, 0.62]	-0.29 [-0.58, -0.01]	-0.55 [-0.84, -0.26]
SRT	22.20 (0.15)	23.12 (0.88)	23.69 (0.87)	0.15 [-0.14, 0.44]	0.24 [-0.05, 0.53]	0.07 [-0.22, 0.35]		

GEE models for quantifying change over time specified $Yij = \beta_0 + \beta_1 \text{Time}_j + \beta_2 \text{Platoon}_k + \beta_2 \text{Platoon}_k * \beta_1 \text{Time}_j + \beta_3 \text{ADFA} + \beta_3 \text{ADFA} * \beta_1 \text{Time}_j + \beta_4 \text{Baseline symptoms} + \epsilon ij$; with $m_i \times m_i$ working correlation matrix for each Yij, $\text{Var}(Yij) = \Phi v(\mu ij)$ y; where Φ is a scale parameter and $v(\cdot)$ is a normal distribution; $\epsilon ij \sim N(0, \sigma 2)$; i is the individual cadet, clustered within j time points (Pre, post and follow-up), k platoons (eight in total); Australian defence force academy (ADFA) enrolment (Yes/No).

Table 4: Parallel multiple mediation model, with brooding and reflection at longer-term follow-up mediating the relationship between intervention condition and outcomes at longer-term follow-up. * $p < .05$

	Follow-up depression symptoms		**Follow-up anxiety symptoms**		**Follow-up perceived stress**	
	b	95% Confidence Interval	*b*	95% Confidence Interval	*b*	95% Confidence Interval
Total effect (X→Y)	-0.44*	-0.86, -0.01	-0.22*	0.25, -0.13	-2.34*	-3.95, -0.73
Total indirect effects	-0.30*	-0.57, -0.05	-0.34*	-0.60, -0.09	-0.93*	-1.85, -0.04
X→M1 (*a1*)	-1.02*	-1.83, -.22	-1.09*	-1.89, -0.28	-1.08*	-1.88, -0.27
X→M2 (*a2*)	0.34*	0.01, 0.09	0.32*	0.08, 0.56	0.33*	0.09, 0.57
M1→Y (*b1*)	0.30*	0.24, 0.37	0.31*	0.23, 0.38	0.99*	0.72, 1.25
M2→Y (*b2*)	0.02	-0.20, 0.24	-0.03	-0.28, 0.21	0.40	-0.48, 1.28
X→M1→Y (indirect effect)	-0.31*	-0.57, -0.07	-0.33*	-0.59, -0.10	-1.06*	-1.88, -0.28
X→M2→Y (indirect effect)	<0.01	-0.07, 0.09	-0.01	-0.09, 0.07	-0.13	-0.12, 0.44

Note: M1 = Brooding; M2 = Reflection; X = Condition; Y = Outcome variable; values represent unstandardized coefficients.

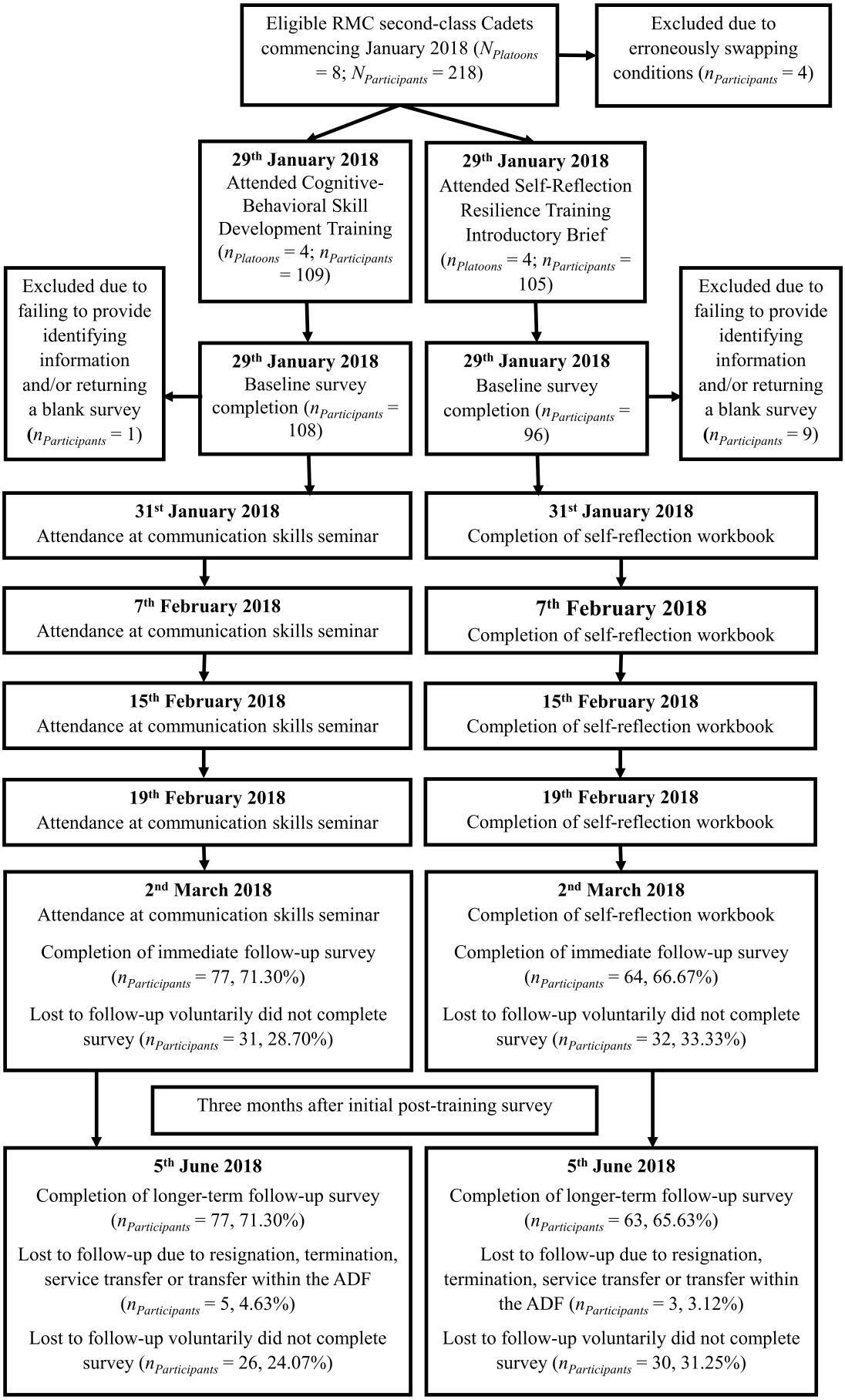

Figure 1: CONSORT flow diagram for intervention delivery and survey administration.

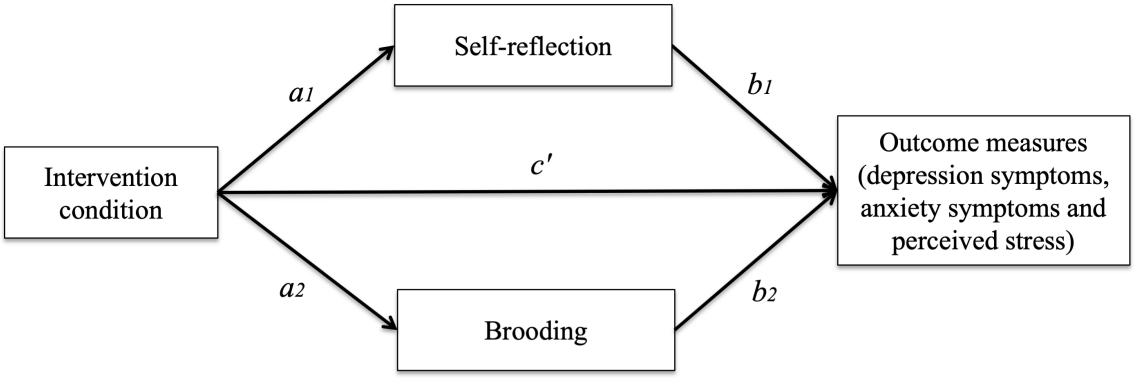

Figure 2. Hypothesised mediation model with self-reflection and brooding each mediating the relationship between intervention condition and outcome measures (depression symptoms, anxiety symptoms and perceived stress).

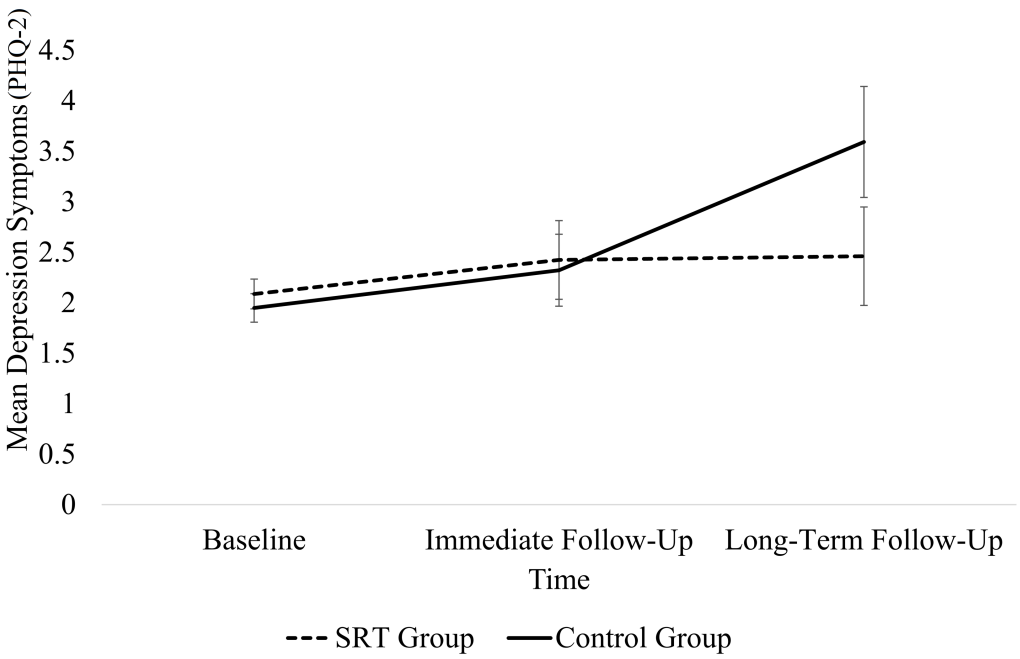

Figure 3. Mean depression symptoms for the SRT and Control groups at baseline, immediate follow-up and longer-term follow-up.

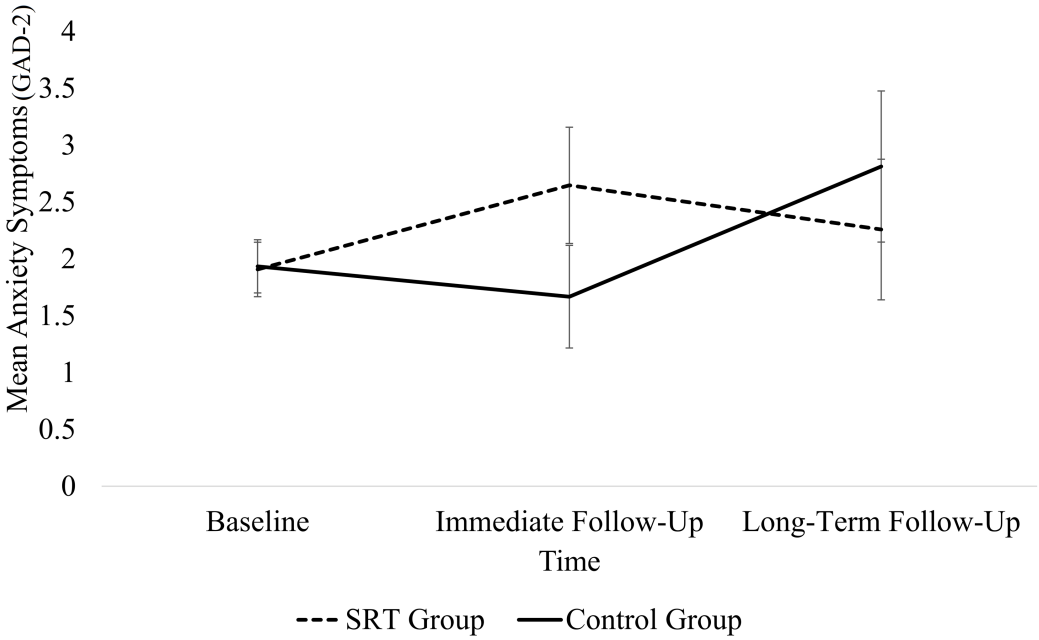

Figure 4. Mean anxiety symptoms for the SRT and Control groups at baseline, immediate follow-up and longer-term follow-up.

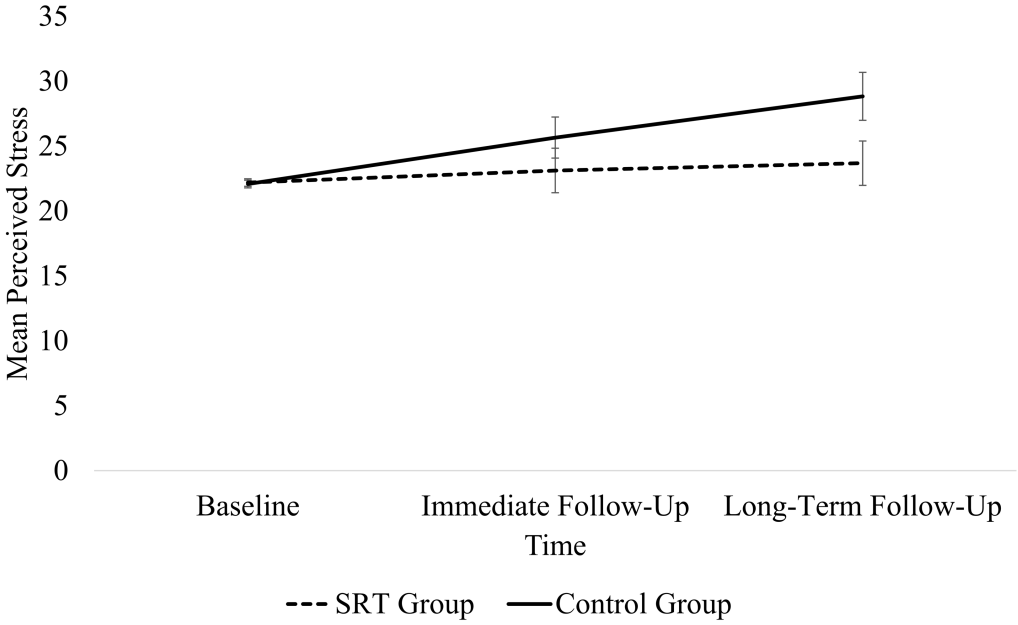

Figure 5. Mean perceived stress scores for the SRT and Control groups at baseline, immediate follow-up and longer-term follow-up.

Supplementary Materials

Table S1: Sample constructs and questions included in the self-reflection workbook exercises.

Reflection construct	Reflective question
1. Awareness of one's emotional, physical, behavioral and cognitive coping responses to stressor events	• Describe one of the more difficult events that you have experienced during the past week. In your response, consider: a) What you were thinking, b) How you behaved, c) How you felt physically and emotionally, and d) How your emotions, thoughts and physical feelings influenced you and your performance
2. Awareness of values and value-based goals in relation to the situation	• Describe your desired outcomes during this situation. a) Consider the task outcomes, and b) The characteristics that you want to demonstrate during the process of achieving those outcomes (reflecting on the values you have previously identified)
3. Awareness of the strategies applied to deal with the situation	• What strategies did you use to reduce your levels of stress in order to maximize your performance in response to this event? Provide as much detail as you can.
4. Evaluation of coping strategy effectiveness in relation to values and goals	• How effective were the strategies that you applied (Q3) for achieving your desired outcomes (Q2)? Rated on a 7-point scale: 1 = Not at all, 4 = Somewhat, 7 = Completely
5. Adaptation of strategies to promote improvements in future coping	• What did you learn about yourself or what useful knowledge, skills or abilities did you develop from this experience?

Table S2: Parameter estimates for the GEE analyses examining longitudinal change in depression symptoms, anxiety symptoms and perceived stress, using only the raw data without model-based replacement.

	Depression Symptoms (PHQ-2)		**Anxiety Symptoms (GAD-2)**		**Perceived Stress**	
	Wald χ^2	***p***	***Wald*** χ^2	***p***	***Wald*** χ^2	***p***
Baseline / Longer-term follow-up	4.070	.044	1.253	.263	5.530	.019
Immediate follow-up / Longer-term follow-up	4.746	.029	7.393	.007	2.471	.116

Table S3: Adjusted means, standard errors, and the effect sizes for within-group and between-group change in depression symptoms, anxiety symptoms and perceived stress for the Control group and the Self-Reflection Resilience Training (SRT) group, calculated during the GEE using only the raw data without model-based replacement.

	Means (Standard Errors) for each Time Point			**Within-Group Effect Sizes calculated using Hedge's *g* [95% Confidence Interval]**			**Between-Group Effect Sizes calculated using Hedge's *g* [95% Confidence Interval]**	
Primary Outcomes	**Baseline**	**Immediate follow-up**	**Longer-term follow-up**	**Baseline to Immediate follow-up**	**Baseline to Longer-term follow-up**	**Immediate to Longer-term follow-up**	**Immediate follow-up (Control vs. SRT)**	**Longer-term follow-up (Control vs. SRT)**
Depression Symptoms (PHQ-2)								
Control	1.92 (.06)	2.29 (.24)	3.67 (.36)	0.21 [-0.09, 0.52]	0.71 [0.39, 1.02]	0.47 [0.16, 0.77]	0.05 [-0.27, 0.37]	-0.38 [-0.71, -0.06]
SRT	2.04 (.07)	2.40 (.26)	2.40 (.36)	0.22 [-0.12, 0.55]	0.16 [-0.18, 0.5]	0 [-0.34, 0.34]		
Anxiety Symptoms (GAD-2)								
Control	1.90 (.09)	1.80 (.32)	3.02 (.45)	-0.04 [-0.35, 0.26]	0.37 [0.06, 0.67]	0.33 [0.02, 0.63]	0.28 [-0.04, 0.6]	-0.26 [-0.58, 0.07]
SRT	1.88 (.09)	2.67 (.37)	1.91 (.52)	0.34 [0.01, 0.68]	0.01 [-0.32, 0.35]	-0.2 [-0.53, 0.14]		
Perceived Stress								
Control	21.92 (.18)	25.55 (1.04)	28.42 (1.29)	0.52 [0.21, 0.82]	0.74 [0.43, 1.06]	0.26 [-0.05, 0.56]	-0.25 [-0.57, 0.07]	-0.48 [-0.81, -0.16]
SRT	22.06 (.18)	23.15 (1.12)	22.66 (1.36)	0.16 [-0.18, 0.49]	0.07 [-0.26, 0.41]	-0.05 [-0.38, 0.29]		

Table S4: Parameter estimates for the MLM analyses examining longitudinal change in in depression symptoms, anxiety symptoms and perceived stress.

	Depression Symptoms (PHQ-2)		**Anxiety Symptoms (GAD-2)**		**Perceived Stress**	
	t	***p***	***t***	***p***	***t***	***p***
Baseline / Longer-term follow-up	-1.950	.052	-0.592	.555	-1.780	.076
Immediate follow-up / Longer-term follow-up	-2.807	.005	-3.093	.002	-1.104	.271

Table S5: Adjusted means, standard errors, and the effect sizes for within-group and between-group change in depression symptoms, anxiety symptoms and perceived stress for the Control group and the Self-Reflection Resilience Training (SRT) group.

	Means (Standard Errors) for each Time Point			**Within-Group Effect Sizes calculated using Hedge's *g* [95% Confidence Interval]**			**Between-Group Effect Sizes calculated using Hedge's *g* [95% Confidence Interval]**	
Primary Outcomes	**Baseline**	**Immediate follow-up**	**Longer-term follow-up**	**Baseline to Immediate follow-up**	**Baseline to Longer-term follow-up**	**Immediate to Longer-term follow-up**	**Immediate follow-up (Control vs. SRT)**	**Longer-term follow-up (Control vs. SRT)**
Depression symptoms (PHQ-2)								
Control	2.45 (.60)	2.04 (.66)	2.76 (.72)	-0.06 [-0.33, 0.21]	0.04 [-0.23, 0.32]	0.1 [-0.17, 0.37]	0.09 [-0.19, 0.37]	-0.11 [-0.39, 0.17]
SRT	2.58 (.62)	2.68 (.67)	1.94 (.74)	0.02 [-0.27, 0.3]	-0.09 [-0.38, 0.19]	-0.11 [-0.39, 0.18]		
Anxiety symptoms (GAD-2)								
Control	1.84 (.72)	0.93 (.59)	2.40 (.69)	-0.13 [-0.4, 0.14]	0.08 [-0.2, 0.35]	0.22 [-0.06, 0.49]	0.17 [-0.12, 0.45]	-0.08 [-0.36, 0.2]
SRT	1.64 (.74)	1.94 (.60)	1.86 (.71)	0.05 [-0.24, 0.33]	0.03 [-0.26, 0.32]	-0.01 [-0.3, 0.28]		
Perceived stress								
Control	22.99 (2.86)	25.27 (2.94)	26.12 (3.12)	0.07 [-0.2, 0.35]	0.1 [-0.17, 0.37]	0.03 [-0.25, 0.3]	-0.05 [-0.33, 0.23]	-0.11 [-0.39, 0.17]
SRT	22.67 (2.93)	23.77 (2.99)	22.49 (3.21)	0.04 [-0.25, 0.33]	-0.01 [-0.29, 0.28]	-0.04 [-0.33, 0.25]		

* Note: The absolute mean values are different from the GEE given that the baseline symptoms are now part of the random-effects model, rather than the fixed-effects model as is the case for GEE.

Table S6: Parameter estimates for the GEE analyses examining longitudinal change in depression symptoms (PHQ-2) and anxiety symptoms (GAD-2), calculated using the dataset reported by Crane, Boga, Karin, et al. (2019).

	Depression Symptoms (PHQ-2)		**Anxiety Symptoms (GAD-2)**	
	Wald χ^2	***p***	***Wald*** χ^2	***p***
Baseline / Longer-term follow-up	0.007	.932	4.375	.036
Immediate follow-up / Longer-term follow-up	2.957	.086	9.022	.003

Table S7: Adjusted means, standard error, and the effect sizes for within-group and between-group change in depression symptoms (PHQ-2) and anxiety symptoms (GAD-2) for the Control group and the Self-Reflection Resilience Training (SRT) group, calculated during the GEE using the dataset reported by Crane, Boga, Karin, et al. (2019).

	Means (Standard Errors) for each Time Point			**Within-Group Effect Sizes calculated using Hedge's *g* [95% Confidence Interval]**			**Between-Group Effect Sizes calculated using Hedge's *g* [95% Confidence Interval]**	
Primary Outcomes	**Baseline**	**Immediate follow-up**	**Longer-term follow-up**	**Baseline to Immediate follow-up**	**Baseline to Longer-term follow-up**	**Immediate to Longer-term follow-up**	**Immediate follow-up (Control vs. SRT)**	**Longer-term follow-up (Control vs. SRT)**
Depression Symptoms (PHQ-2)								
Control	1.38 (.12)	2.18 (.16)	2.30 (.16)	0.58 [0.28, 0.88]	0.66 [0.36, 0.97]	0.08 [-0.22, 0.37]	-0.11 [-0.39, 0.16]	-0.37 [-0.65, -0.1]
SRT	1.04 (.11)	2.02 (.12)	1.76 (.12)	0.76 [0.49, 1.02]	0.56 [0.3, 0.82]	-0.19 [-0.45, 0.06]		
Anxiety Symptoms (GAD-2)								
Control	1.42 (.45)	2.00 (.17)	1.83 (.18)	0.17 [-0.12, 0.47]	0.12 [-0.17, 0.42]	-0.1 [-0.39, 0.19]	-0.12 [-0.4, 0.15]	-0.5 [-0.78, -0.22]
SRT	1.17 (.12)	1.82 (.12)	1.06 (.12)	0.48 [0.22, 0.74]	-0.08 [-0.34, 0.17]	-0.56 [-0.83, -0.3]		

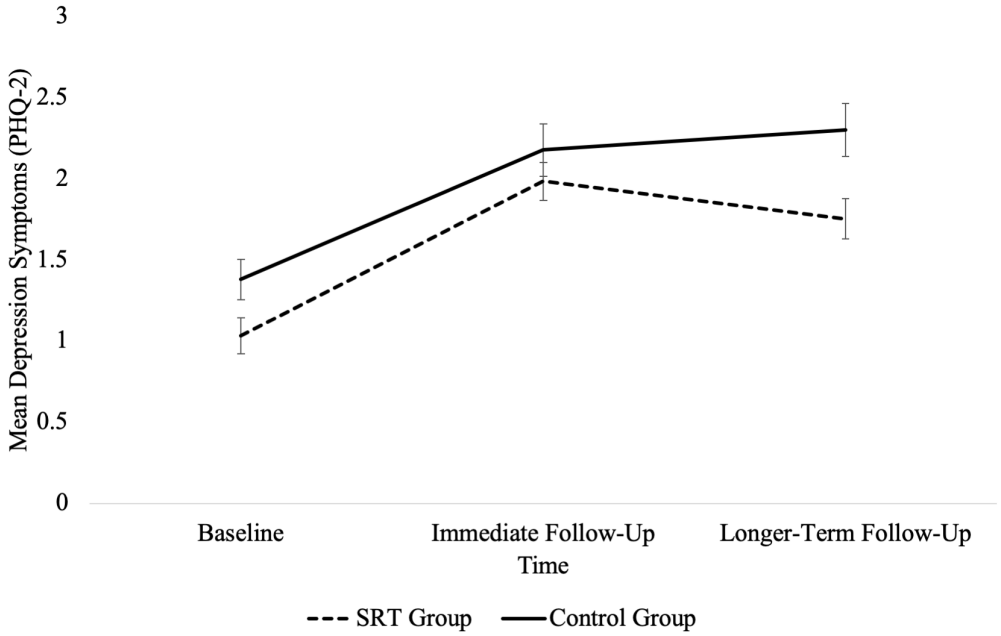

Figure S1. Mean depression symptoms (PHQ-2) for the SRT and Control groups at baseline, immediate follow-up and longer-term follow-up, calculated during the GEE using the dataset reported by Crane, Boga, Karin, et al. (2019).

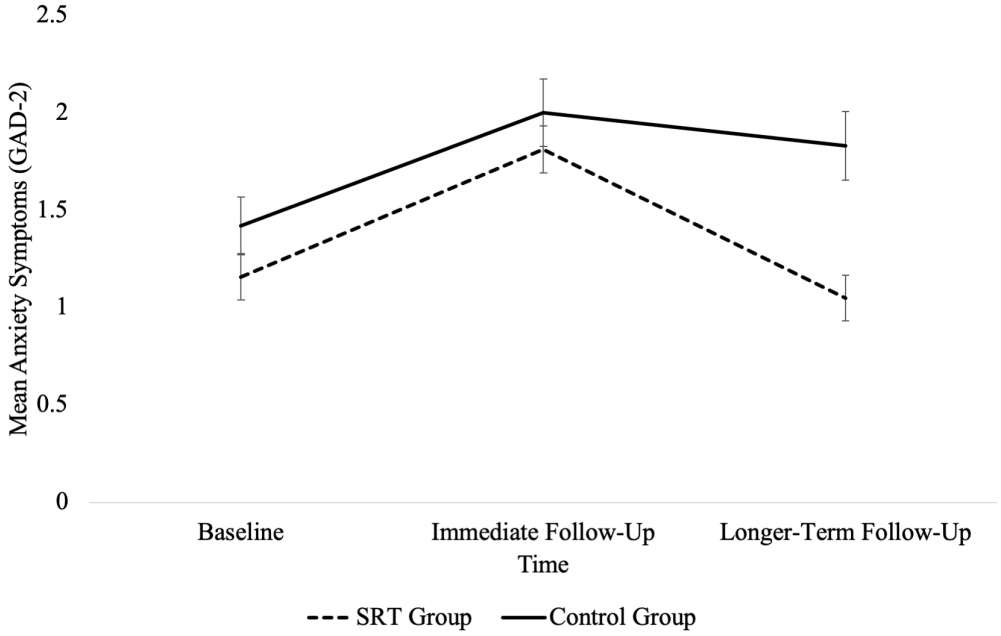

Figure S2. Mean anxiety symptoms (GAD-2) for the SRT and Control groups at baseline, immediate follow-up and longer-term follow-up, calculated during the GEE using the dataset reported by Crane, Boga, Karin, et al. (2019).