Word count: 4,532
Tables: 3
Figures: 1

The effectiveness of interventions aimed at improving wellbeing and resilience to stress in first responders: A systematic review

Jennifer Wild[1]
Shama El-Salahi[2]
Michelle Degli Esposti[3]

[1]Department of Experimental Psychology, University of Oxford
[2]The Oxford Centre for Psychological Health, University of Oxford
[3]Department of Social Policy and Intervention, University of Oxford

Acknowledgements: This work was funded by Mind (grant number CQR00510) awarded to Jennifer Wild and supported by the NIHR Oxford Health Biomedical Research Centre. Mind had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Conflicts of interest: None.

Abstract

First responders are routinely exposed to trauma and suffer higher rates of mental and physical ill health compared to the general population. Interventions which could improve resilience to stress may help to protect the health of this high risk population. We systematically reviewed such interventions for first responders to determine which ones work and why. We searched the Cochrane and Campbell Collaboration Library, EMBASE, IBSS, Medline, PILOTS, PubMed, PsycINFO, and SCOPUS from 1 January 1980 to 28 June 2018 for randomised and quasi-randomised controlled studies aiming to improve wellbeing, resilience or stress management for police, ambulance, fire, or search and rescue workers using non-pharmacological interventions. Data were extracted from published reports and obtained from authors. Within- and between-group effect sizes were calculated for mental and physical health outcomes. Risk of bias was assessed using the Cochrane Collaboration's Risk of Bias Tool. The initial search identified 3,816 studies, 13 of which were eligible for analysis (n=634 cases, n=628 controls). Six studies demonstrated intervention-related improvements. However, risk of bias was mostly unclear or high. Within-group intervention effect sizes ranged from -0·82 (95% CI -1·48 – -0·17) to 2·71 (1·99 – 3·42) and between-group intervention effect sizes ranged from -0·73 (-1·25 – -0·21) to 1·47 (0·94 – 2·01), depending on the outcome. Largest effects were seen for interventions that targeted modifiable risk factors for trauma-related psychiatric disorders. Targeting modifiable predictors of trauma-related psychiatric disorders through training may protect the health of first responders who routinely face trauma in their line of work.

Keywords: first responder; post-traumatic stress disorder; major depressive disorder; resilience; review

Introduction

Populations routinely exposed to occupational trauma and ongoing stressors are considered at high risk of poor mental and physical health. Police officers, paramedics, firefighters and search and rescue workers routinely face ongoing stress and experience higher rates of psychiatric and physical disorders than the general population (Reichard & Jackson, 2010; Javidi & Yadollahi, 2012). The most common outcomes first responders develop following exposure to trauma are posttraumatic stress disorder (PTSD) and major depressive disorder (MDD; Kleim & Westphal, 2011), whilst physical health problems, such as sleep disturbances and musculoskeletal problems, are also highly prevalent. Evidence suggests that first responders who develop PTSD or MDD are at an even greater risk of poor physical health (McFarlane et al., 1994; Fjeldheim et al., 2014; Wild et al., 2016). Interventions aimed at improving the mental health of first responders may therefore have beneficial secondary effects on physical health problems. A number of interventions have been developed and evaluated for first responders, which aim to improve wellbeing and resilience to stress. It is unclear which ones work and why.

The approach within medicine to improve resilience to poor health is to target risk factors for disease. Knowledge of risk factors allows preventative interventions to be developed and delivered, improving outcomes, longevity and quality of life. Some predictors of disease are fixed and cannot be modified. However, other predictors, such as behavioural traits, are modifiable with training to reduce the risk of developing life-threatening conditions. Lifestyle modifications targeting hypertension have been shown to reduce the risk of cardiovascular disease and mortality, for example (Samadian et al., 2016). It is similarly possible to reduce the likelihood of psychiatric disorders by identifying and targeting modifiable risk factors for poor mental health (Topper et al., 2017).

Drawing on the extant literature, it would appear that modifiable predictors that have previously predicted PTSD and depression, the most common mental health problems emergency workers develop, fall into five key areas: *personality variables*, such as, neuroticism, trait dissociation, anxiety sensitivity and trait anger; *coping variables*, such as behavioural disengagement, wishful thinking, emotional suppression, rumination, and intentional numbing; *cognitions*, such as

resilience appraisals, attributions, and posttraumatic cognitions; and *social support variables*, such as general social support, and support at work; and *physical inactivity*.

Turning to personality variables, longitudinal twin studies have identified neuroticism as a risk factor for depression (Kendler et al., 2003), and a predictor of PTSD in burns victims (Lawrence & Fauerbach, 2003). Trait dissociation assessed in police new recruits during their academy training predicts PTSD one year later (McCaslin et al., 2008). Anxiety sensitivity assessed post-trauma predicts PTSD in survivors of physical injury (Marshall et al., 2010) and trait anger assessed before deployment predicts PTSD two months afterwards in Dutch soldiers (Lommen et al., 2014). In terms of coping variables, in a cross-sectional study of ambulance workers, Clohessy and Ehlers (1999) found that dissociation, rumination, and suppression in response to intrusive memories predicted PTSD symptom severity as did coping strategies, such as wishful thinking, behavioural disengagement (mental disengagement), and intentional numbing. Nolen-Hoeksema and Morrow (1991) assessed 137 students fourteen days before the Loma Prieta earthquake. Students who ruminated about the earthquake in the ten days that followed were more likely to develop high levels of depressive and stress symptoms seven weeks later.

Turning to cognitions, Wild et al. (2016) assessed 453 newly recruited paramedics during their first week of training. Logistic regressions showed that rumination about memories of stressful events at the start of training uniquely predicted an episode of PTSD. Perceived resilience (appraisals about resilience) uniquely predicted an episode of MDD. Alloy et al. (2006) found that negative cognitive styles predicted first onset and recurrences of major depression in a study of 347 first year students without initial psychiatric disorders. In a study of 967 consecutive patients to an emergency clinic following road traffic accidents, Ehlers et al. (1998) found that negative interpretations of intrusions, rumination, thought suppression, and anger cognitions, enhanced the accuracy of the prediction of chronic PTSD at 1 year follow-up.

With respect to social support, Ozer et al. (2003) reviewed 2,647 studies of PTSD and identified poor social support as a significant predictor of PTSD. Positive social support is a predictor of recovery from depression (Brughal et al., 1990). With respect to physical inactivity, in a review

of the literature, Warburton et al. (2006) identified physical inactivity as a significant predictor of depression. Resilience interventions which target these modifiable predictors may provide a promising approach to protecting the health of at-risk populations.

It is unclear whether interventions aimed at improving wellbeing and resilience to stress are effective at improving the health of first responders and if they are, which ones are most effective. Previous systematic reviews have concluded that such interventions have a small effect on improving resilience and mental health outcomes, and highlight heterogeneity in intervention design, content and outcome measurement, and low methodological quality among studies (Leppin et al., 2014; Macedo et al., 2014; Vanhove et al., 2015). We hypothesise that many resilience and wellbeing interventions fare poorly because they fail to target modifiable predictors of poor mental health. There are currently no reviews that identify or evaluate interventions across first responder populations aimed to improve wellbeing, resilience or stress management and no reviews which quantitatively synthesise the evidence to meaningfully assess and compare the effectiveness of such interventions. This systematic review aims to: (1) identify interventions for first responders, which focus on improving wellbeing, resilience or stress management; and (2) quantitatively synthesise the empirical evidence to determine the effectiveness of such interventions in improving mental and physical health outcomes.

The included interventions differ in length and content, which allows us to investigate the effectiveness of different intervention approaches and garner meaningful conclusions about the optimum content, format and length. This is important because it enables us to make recommendations about how to focus the field in terms of intervention development. We define resilience as a dynamic process that can be developed (e.g., Ong et al., 2009), that will change across the lifespan (e.g., Windle, 2011), and that will buffer against the development of mental and physical health problems in difficult times (e.g., Rutter, 1985; Yi et al., 2008). All interventions included in our review were administered to first responders who were psychologically well and the interventions were provided with the intention to keep them well. We see this as in line with our conceptualisation of resilience as promoting wellbeing over time in the face of stressful work. We define wellbeing as the experience of feeling and functioning well with a sense of satisfaction for life (New Economics Foundation, 2016).

Method

Search strategy and selection criteria

This systematic review followed PRISMA reporting guidelines (Moher et al., 2009). A review protocol was registered in advance on PROSPERO, registration number CRD42016034072. We searched the Cochrane and Campbell Collaboration Library, EMBASE, IBSS, Medline, PILOTS, PubMed, PsycINFO, and SCOPUS from 1 January 1980 to 28 June 2018, in English, using a Boolean search strategy combining keywords related to study type, participants, and interventions (see Appendix A for full search strategy). We also hand searched the reference lists of included studies and relevant reviews identified during the initial searches (Peñalba et al., 2008; Leppin et al., 2014; Macedo et al., 2014; Vanhove et al., 2015), and the publication histories of known resilience researchers. We included studies if they satisfied the following four eligibility criteria:

1. Randomised or quasi-randomised controlled trials (RCT, QCT, respectively), conducted in high income countries (defined by the World Bank at time of publication) and published after 1980 in peer-reviewed journals.
2. Males and/or females aged 18 years or above, currently working within the emergency services (ambulance, fire, first aid responders, police, and search and rescue).
3. Non-pharmacological interventions aiming to improve mental health resilience. Amongst other interventions, these may include cognitive behavioural therapy (CBT)-based interventions (e.g. mental imaging training, stress education and management programs, psycho-education interventions, relaxation), mindfulness courses, supportive therapies (e.g. counselling sessions), psychodynamic therapies, debriefing and exercise therapies.
4. Outcome measures of mental and/or physical health outcomes. Physiological (e.g., heart rate, blood pressure) or performance outcomes were not included.

Data analysis

The second author (SE) removed all duplicates from the initial searches, then screened the titles and abstracts of all papers against the eligibility criteria. Full-text articles of potentially eligible studies were then retrieved. If articles were not found, corresponding authors were contacted to

provide full articles. Two reviewers (MDE and SE) blindly and independently assessed all full-text articles for eligibility, documenting reasons for exclusion. There was high inter-rater agreement (92·9%). Disagreements were resolved by consensus. Reviewers were not masked to the journals or authors of the studies reviewed.

We developed a data extraction sheet (Online Supplementary Material) which included the Cochrane Collaboration's tool for assessing risk of bias at the study level (Higgins et al., 2011). The following data were extracted: 1) study methods (setting, design, arms, control condition, time-points); 2) participant characteristics (sample size, emergency service, gender); 3) intervention characteristics (name, content, deliverers, delivery format, duration and frequency); 4) mental and physical health outcomes (type, measure, informant); and 5) study quality, according to: risk of selection, performance, detection, attrition and reporting bias. Data extraction and quality assessment were checked by the first author at a later time-point.

To carry out a comprehensive and consistent data synthesis across studies, we adopted a two-pronged approach. We calculated effect sizes (Cohen's *d*) and confidence intervals (CIs) for all relevant outcomes (1) within-groups, pre to post-intervention (and where applicable, follow-up); and (2) between-groups (i.e. intervention vs control) at post-intervention (and where applicable, follow-up). We only calculated between-group effect sizes if there were no significant differences at baseline. For calculating within-group effect sizes we assumed statistical independence for pre- and post-intervention and pooled the standard deviation: $Cohen's\ d = \frac{M_{pre} - M_{post}}{SD_{pooled}}$, with $SD_{pooled} = \sqrt{\left(\frac{SD_{pre}^2 + SD_{post}^2}{2}\right)}$ (Van Etten & Taylor, 1998). Where an improvement was characterised by a decrease in outcome scores, we adapted the formula to $Cohen's\ d = \frac{M_{post} - M_{pre}}{SD_{pooled}}$ so that a positive effect size consistently represented a beneficial effect.

The same formula was adapted to calculate between-group effect sizes, where M_{pre} was replaced with the intervention group's post-intervention (or follow-up) mean, and M_{post} the control group's post-intervention mean. Cohen's (1988) suggestion of 0·2 constituting a small effect, 0·5 a moderate effect and 0·8 a large effect, was used for interpreting effect sizes. CIs were calculated using

Hedges and Olkin's (1985, p.86) formula: $CIs = d \pm (1.96 \times SD)$, where $SD = \sqrt{\frac{(N_{pre}+N_{post})}{(N_{pre}\times N_{post})} + \left(\frac{d^2}{2(N_{pre+N_{post}})}\right)}$. Where there was insufficient data reported to calculate Cohen's d, we contacted the corresponding authors. Improvements were considered intervention-specific if there was a significant between-group effect for outcomes at post-intervention or follow-up, and if a within-group effect was not seen in the control group over time. Significant within-group effects were not seen as intervention-specific if no significant between-group effects existed.

Outcomes were grouped into two broad categories: mental health outcomes and physical health outcomes. We further derived sub-categories. Mental health sub-categories included wellbeing (i.e., psychological wellbeing, psychological distress and functioning, and emotional states), resilience, coping, stress, suicidal ideation, mindfulness, quality of life, and subsyndromal mental health symptomatology (i.e., depression, anxiety and PTSD). Physical health sub-categories included sleep, alcohol use, tobacco use, burnout/vital exhaustion (a state of excessive tiredness characterised by fatigue, increased irritability, and demoralisation; Kop et al, 1994), physical activity, healthy eating, and somatic symptoms (including general physical health, physical stress, pain, and stomach and heart complaints).

Results

The review process resulted in the inclusion of 13 studies (See Figure 1 for the PRISMA study selection flowchart; Moher et al., 2009), involving 1,264 participants in total. Six QCTs (Norris et al., 1990; Norvell & Belles, 1993; Ireland et al., 2007; Ângelo & Chambel, 2013; Arnetz et al., 2013; Ramey et al., 2017), four RCTs (Wilson et al., 2001; Tanigoshi et al., 2008; McCraty & Atkinson, 2012; Christopher et al., 2018), and three cluster RCTs (Tuckey & Scott, 2014; Kuehl et al., 2014; Skeffington et al., 2016) were conducted across five countries, with the majority taking place in the United States (n=7). Sample sizes ranged from 34 to 408 and on average 16% of participants were women (range 0% – 42%). Ten studies included samples of police, whilst three included firefighters (Arnetz et al., 2013; Tuckey & Scott, 2014; Skeffington et al., 2016). Fifteen different types of

outcomes were measured across studies; the most commonly reported were wellbeing (n=11), stress (n=8), and mental health symptomatology (n=7). A large variety of scales and statistical analyses were used to assess outcomes, and only one study did not test between-group differences (i.e. intervention vs control) at baseline (Tanigoshi et al., 2008). Table 1 describes the main characteristics of included studies.

The overall risk of bias across studies was unclear to high, particularly for blinding and selective outcome reporting. Four studies demonstrated high selection bias due to poor randomisation of participants or clusters (Norris et al., 1990; Norvell & Belles, 1993; Ireland et al., 2007; Ângelo & Chambel, 2013; Arnetz et al., 2013), whilst attrition bias was moderately high since six studies had high rates of drop-out and non-responders (Norvell & Belles, 1993; Ireland et al., 2007; Tanigoshi et al., 2008; Arnetz et al., 2013; Tuckey & Scott, 2014; Skeffington et al., 2016). See Appendix B for the methodological quality assessment of included studies.

The format and content of interventions varied considerably across studies. Durations ranged from 90 minutes to 16 weeks, with the least frequent occurring as a one-off session (debriefing) and the most frequent occurring three times per week (both physical exercise interventions). Five studies targeted a modifiable risk factor mentioned in the introduction of this review (n=2 behavioural disengagement, n=2 physical inactivity, n=1 emotional suppression). Although intervention content was different in each study, there were five main types: (1) physical exercise interventions; (2) psychological interventions; (3) stress management interventions; (4) self-regulation interventions; and (5) debriefing after a potentially traumatic incident. Intervention details are summarised in Table 2 and the effect sizes and CIs, calculated by the authors of this review, are provided in Table 3.

Physical exercise interventions

Two studies evaluated physical exercise interventions that targeted physical inactivity as a modifiable risk factor. Participants assigned to aerobic exercise and participants assigned to anaerobic exercise demonstrated greater wellbeing (aerobic d=0·83, 0·28 – 1·39; anaerobic d=0·63, 0·06 – 1·20) and quality of life (aerobic d=0·94, 0·38–1·49; anaerobic d=0·72, 0·15 – 1·29) at post-intervention when compared to a control group receiving treatment as usual (TAU; Norris et al.,

1990). The aerobic group also demonstrated greater improvements in stress by post-intervention (*d*=0·72, 0·17 – 1·27). Between-group effect sizes for stress were not calculated for the anaerobic group since they had significantly higher stress at baseline than the aerobic and TAU control groups. There were no significant between-group effects between the aerobic and anaerobic groups suggesting that type of exercise did not affect efficacy. In the second study, circuit weight training was shown to improve psychological functioning (*d*=1·00, 0·24 – 1.76) and physical symptoms (*d*=1·36, 0·57 – 2.15) compared to a waitlist control group at post-intervention, and these improvements were also seen when comparing scores from pre- to post-intervention (Norvell & Belles, 1993. The control group did not improve on any outcome over time. Between-group comparisons with a dropout group were not calculated since the participants who dropped out demonstrated significantly higher baseline scores on all measures. Neither study conducted follow-up assessments.

Psychological interventions

Two studies evaluated psychological interventions. Comparisons between groups showed that Eye Movement Desensitisation and Reprocessing (EMDR), which did not target a modifiable risk factor, was linked to less stress (*d*=0·84, 0·33 – 1·35) at post-intervention compared to a control group receiving a stress management intervention (Wilson et al., 2001). Insufficient data meant that intervention-specific improvements could not be determined for other outcomes measured, such as PTSD symptoms, distress, or coping. In the second study, wellness counselling targeting behavioural disengagement led to significant within-group improvements in wellbeing, but no significant between-group effects (Tanigoshi et al., 2008).

Stress management interventions

Six studies evaluated stress management interventions, two of which targeted modifiable risk factors. An imagery intervention targeted behavioural disengagement and was linked to less vital exhaustion (*d*=0·60, 0·11 – 1·09), better coping (*d*=0·95, 0·45 – 1·45) and better sleep quality (d=0·52, 0·03 – 1·01), when compared to a TAU control group at 18-month follow-up (Arnetz et al., 2013). Imagery was also linked to less vital exhaustion over time (*d*=0·58, 0·08 – 1·08), whilst a TAU

control group showed no significant within-group improvements from pre-intervention to 18-month follow-up. Insufficient data were reported at post-intervention so within- and between-group effect sizes could not be calculated at this time point. A writing intervention targeting emotional suppression led to significant within-group improvements in stress, but did not lead to any intervention-specific effects (Ireland et al., 2007).

The four remaining stress management interventions did not target modifiable risk factors. Two of these studies - a leadership stress management program (Ângelo & Chambel, 2013) and Mental Agility and Psychological Strength (MAPS) training (Skeffington et al., 2016) - could not be linked to intervention-specific improvements. A health and safety program offering psychoeducation on healthy eating, exercise, body weight, stress, sleep and other lifestyle factors, was associated with some significant intervention-specific improvements in health and wellbeing but no improvements in mental health outcomes (Kuehl et al., 2014). Participants receiving this health and safety program had better general health (*d*=0·34, 0·13 – 0·55), less stress (*d*=0·11, 0·16 – 0·58), better sleep quality (*d*=0·69, 0·48 – 0·90), better sleep quantity (*d*=0·65, 0·44 – 0·86), and less tobacco use (*d*=0·35, 0·14 – 0·56) at six-month follow-up compared to a TAU control group. These findings were also found when comparing pre-intervention to six-month follow-up, but not for the control group. However, participants receiving the health and safety program had more depressive symptoms (*d*=-0·36, -0·57 – -0·15), more fatigue (*d*=-0·55, 0·76 – 0·34), and more musculoskeletal pain (*d*=-0·22, -0·43 – -0·01) at six-month follow-up compared to the control group, which contradicts the apparent efficacy of the intervention.

Mindfulness-Based Resilience Training (MBRT) was the only stress management intervention that could be linked to clear intervention-specific improvements, although it did not target a modifiable risk factor. Between-group comparisons revealed that MBRT was linked to fewer sleep difficulties (*d*=0·56, 0·04 – 1·16), less burnout (*d*=0·73, 0·17 – 1·30), and higher nonreactivity to inner experience (*d*=0·63, 0·07 – 1·20), a facet of mindfulness, at post-intervention (Christopher et al., 2018). It did not lead to any intervention-specific improvements in stress, wellbeing, or mental health symptomatology, such as anxiety and depression. Between-group effect sizes were not calculated for resilience and self-compassion since the groups differed significantly on these

outcomes at baseline. There were no significant between or within-group effects at three-month follow-up.

Self-regulation interventions

Two studies evaluated self-regulation interventions that did not target modifiable risk factors (Ramey et al., 2017; McCraty & Atkinson, 2012). No significant between-group effect sizes were found in either study, thus neither intervention could be linked to intervention-specific improvements. All within-group effects were also non-significant except for improvements in distress, which were observed in one of the two studies (McCraty & Atkinson, 2012.)

Debriefing

The final study evaluated a debriefing intervention (Tuckey & Scott, 2014). It did not target a modifiable risk factor and was not linked to any within or between group effects.

Discussion

This is the first systematic review to synthesise randomised and quasi-randomised controlled trial evidence for the effectiveness of interventions aimed to improve wellbeing, resilience or stress management in the emergency service population. Over half the interventions had no significant effects on mental or physical health outcomes. The interventions that were linked to beneficial effects that could be considered intervention-specific were more likely to target modifiable risk factors of poor mental health.

Exercise and imagery interventions showed the most promise as they were linked to the largest between-group effect sizes, and both intervention types targeted modifiable risk factors. The benefits of exercise on mental health seen in this review mirror the substantial body of literature supporting the benefits of exercise on physical and mental health outcomes in the general population (i.e., Warburton et al., 2006). With respect to imagery interventions, the findings mirror the benefits of imagery interventions for anxiety (i.e., Wild et al., 2008) and depression (Brewin et al. 2009). In

contrast, self-regulation and debriefing interventions could not be linked to any intervention-related improvements, and neither targeted modifiable risk factors. Interventions were most likely to improve outcomes such as wellbeing, stress, and sleep problems rather than leading to significant improvements in mental health symptomatology. The latter finding is likely due to the observation that participants demonstrated low levels of symptomatology when they were recruited into the studies with little room to further improve. While targeting wellbeing, stress and sleep problems may reduce the likelihood of developing mental ill health in the long-term, it is also crucial to develop interventions that can alleviate or prevent the development of other major symptoms of psychiatric disorders, such as extreme moods and suicidal thoughts, that are often more debilitating. Evaluations would benefit from plotting and comparing trajectories of outcome following stressor exposure in order to also determine the protective potential of an intervention.

Interventions with a higher number and frequency of sessions tended to fare better. The mean number of sessions was 17.7 for interventions that could be linked to intervention-related improvements, whereas interventions that could not be linked to improvements were delivered with an average of 3.6 sessions. The interventions that were linked to intervention-specific effects mostly took place once a week, whereas the least effective interventions were more likely to take place in a block of a three or four days, or every two weeks. Thus, the frequency of sessions may be an important factor affecting efficacy.

Although the majority of evidence points towards the notion that interventions targeting modifiable risk factors of poor mental health are more effective, there are some contradictory results to consider. Mindfulness-based Resilience Training (MBRT) and Eye Movement Desensitisation and Reprocessing (EMDR) demonstrated significant intervention-specific effects but did not target modifiable risk factors. However, the effects of MBRT were not sustained at follow-up, the sample size was fairly small, and the training did not improve any mental health outcome other than one facet of mindfulness, which would be expected with a mindfulness-based intervention. In the study involving EMDR, insufficient data were provided at follow-up so it is unclear whether the effects seen at post-intervention would have been sustained, the sample size was again fairly small, and participants receiving EMDR had significantly more face-to-face contact with a therapist compared to

the control group receiving typical stress management. These limitations restrict the generalisability of both studies and question the efficacy of the interventions. At the other end of the spectrum, two interventions (writing about strong emotions and wellness counselling) targeting modifiable risk factors were linked to within-group improvements in stress and wellbeing respectively, but could not be linked to intervention-specific effects. The writing intervention only lasted for four days and the wellness counselling took place every two weeks, whereas the more effective interventions tended to take place at least once a week. This difference in session frequency may have reduced the potential efficacy of the interventions. Perhaps allowing time for reflection and consolidation between sessions but ensuring regular contact to improve knowledge and skill retention are important components to consider.

Beyond the limitations of individual studies mentioned so far, several limitations were observed with all the identified evidence. First, there was a sheer lack of randomised and quasi-randomised trials investigating the effectiveness of wellbeing, resilience and stress management interventions for first responders, and particularly for ambulance, fire, and search and rescue personnel, with much of the focus resting on police officers. Second, the quality of evidence was low, although it may not be the case that the execution of the trials themselves were poor, but the reporting of them was often unclear. Third, sample sizes were relatively small, and few trials conducted follow-ups. Fourth, many of the trials compared their interventions to no training, making it difficult to conclude whether the observed effects were linked to the intervention or to non-specific effects, such as contact with a group or a counsellor. Fifth, included studies failed to assess first responders' responses to subsequent trauma exposure and thus, it is unclear whether or not the benefits linked to the training programmes also led to lower rates of mental and physical health disorders.

This review's strengths lie in the rigorous methodology employed. However, there are three main limitations worth considering. First, due to resource restrictions, the initial literature search criteria were limited to high-income countries after 1980, the searches were only conducted in English, and 'grey literature' was not obtained. Second, meta-analyses and corresponding graphical examinations of funnel plots were not conducted due to the limited evidence base and the substantial heterogeneity across included studies. Third, effect size (Cohen's d) calculations for within-group

comparisons require a measure of correlation. Because this was not available, we derived within-group effect sizes by assuming statistical independence between pre- and post-intervention (and where applicable, follow-up) scores. This assumption is inaccurate and may result in over-estimating the true effects. To address this limitation, we adopted a two-pronged approach by calculating pre-intervention to post-intervention (and where applicable, follow-up) effect sizes and group comparisons at post-intervention and follow-up, providing there were no significant group differences at baseline. This allowed more accurate effect sizes and CIs to be estimated.

The results of this review significantly extend those conducted previously, offering a rigorous quantitative approach to evaluating resilience, wellbeing and stress management interventions, and then identifying commonalities of the most effective approaches. A systematic approach to intervention development is needed, first identifying modifiable risk factors in at-risk groups with prospective studies, then developing interventions to modify core risk factors. Such a theory-driven approach may help to unify evidence and offer a promising way forward for developing interventions urgently needed for occupational groups regularly exposed to trauma.

References

**Primary studies included in the systematic review*

Alloy LB, Abramson LY, Whitehouse WG, Hogan ME, Panzarella C, Donn R (2006). Prospective incidence of first onsets and recurrences of depression in individuals at high and low cognitive risk for depression. *Journal of Abnormal Psychology* **115(1)**, 145–156.

***Ângelo RP, Chambel MJ** (2013). An intervention with firefighters to promote psychological occupational health according to the Job Demands-Resources Model. *Revista de Psicología Social* **28(2)**, 197–210.

***Arnetz BB, Arble E, Backman L, Lynch A, Lublin A** (2013). Assessment of a prevention program for work-related stress among urban police officers. *International Archives of Occupational and Environmental Health* **86(1)**, 79–88.

Brewin, C.R., Wheatley, J., Patel, T., Fearon, P., Hackmann, A., Wells, A., Fisher, P., & Myers, S. (2009). Imagery rescripting as a brief stand-alone treatment for depressed patients with intrusive memories. *Behaviour Research and Therapy* 47, 569–576.

Brughal, P.E., Bebbington, B., MacCarthy, E., Sturt, T., Wykes, T., & Potter, J. (1990). Gender, social support and recovery from depressive disorders: a prospective clinical study. *Psychological Medicine*, 20, 147-156.

***Christopher MS, Hunsinger M, Goerling LRJ, Bowen S, Rogers BS, Gross CR, Dapolonia E, Pruessner JC** (2018). Mindfulness-based resilience training to reduce health risk, stress reactivity, and aggression among law enforcement officers: A feasibility and preliminary efficacy trial. *Psychiatry Research* **264**, 104-115.

Clohessy S, Ehlers, A. (1999). PTSD symptoms, response to intrusive memories and coping in ambulance service workers. *British Journal of Clinical Psychology* **38(Pt 3)**, 251–265.

Cohen J (1998). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers: New York.

Ehlers, A., Mayou, R. A., & Bryant, B. (1998). Psychological predictors of chronic posttraumatic

stress disorder after motor vehicle accidents. *Journal of Abnormal Psychology, 107*(3), 508-519.

Fjeldheim CB, Nöthling J, Pretorius K, Basson M, Ganasen K, Heneke R, Cloete KH, Seedat S (2014). Trauma exposure, posttraumatic stress disorder and the effect of explanatory variables in paramedic trainees. *BMC Emergency Medicine* **14**, 11.

Fletcher D, Sarkar M (2013). Psychological resilience: A review and critique of definitions, concepts and theory. *European Psychologist* **18(1)**, 12-23.

Hedges L, Olkin I (1985). *Statistical Methods for Meta-Analysis*. Academic Press: New York.

Higgins JPT, Altman DG, Sterne JAC (2011). Chapter 8: Assessing risk of bias in included studies. In *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011] (ed. J. P. T. Higgins and S. Green). The Cochrane Collaboration. Available from www.handbook.cochrane.org.

***Ireland M, Malouff JM, Byrne B** (2007). The efficacy of written emotional expression in the reduction of psychological distress in police officers. *International Journal of Police Science & Management* **9(4)**, 303–311.

Javidi H, Yadollahie M (2012). Post-traumatic Stress Disorder. *International Journal of Occupational and Environmental Medicine* **3(1)**, 2–9.

Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ (1993). A longitudinal twin study of personality and major depression in women. *Arch Gen Psychiatry* 50: 853–862.

Kendler KS, Kuhn J, Prescott CA (2004). The interrelationship of neuroticism, sex, and stressful life events in the prediction of episodes of major depression. *American Journal of Psychiatry* **161(4)**, 631–636.

Kleim B, Westphal M (2011). Mental health in first responders: A review and recommendation for prevention and intervention strategies. *Traumatology* **17(4)**, 17–24.

Kop WJ, Appels AP, Mendes de Leon CF, de Swart HB, Bär FW (1994). Vital exhaustion predicts new cardiac events after successful coronary angioplasty. *Psychosomatic Medicine* **56(4)**, 281-7.

***Kuehl KS, Elliot DL, Goldberg L, Mackinnon DP, Vila BJ, Smith J, Miočević M, O'Rourke HP, Valente MJ, DeFrancesco C, Sleigh A, McGinnis W** (2014). The safety and health improvement: enhancing law enforcement departments study: Feasibility and findings. *Frontiers*

in Public Health **2(38)**, 1-7.

Kuehl KS, Elliot DL, MacKinnon DP, O'Rourke HP, Miočević M, Valente M, Sleigh A, Garg B, McGinnis W, Kuehl H (2016). The SHIELD (Safety and Health Improvement: Enhancing Law Enforcement Departments) study: Mixed methods longitudinal findings. *Journal of Occupational and Environmental Medicine* **58(5)**, 492–498.

Lawrence JW, Fauerbach JA (2003). Personality, coping, chronic stress, social support and PTSD symptoms among adult burn survivors: a path analysis. *Journal of Burn Care & Rehabilitation* **24(1)**, 63–72.

Leppin AL, Bora PR, Tilburt JC, Gionfriddo MR, Zeballos-Palacios C, Dulohery MM, Sood A, Erwin PJ, Pablo Brito J, Boehmer KR, Montori VM (2014). The efficacy of resiliency training programs: A systematic review and meta-analysis of randomized trials. *PLOS ONE* **9(10)**, e111420.

Lommen MJJ, Engelhard IM, van de Schoot R, van den Hout MA (2014). Anger: Cause or consequence of posttraumatic stress? A prospective study of Dutch soldiers. *Journal of Traumatic Stress* **27(2)**, 200-207.

Luthar SS, Cicchetti D (2000). The construct of resilience: Implications for interventions and social policy. *Development and Psychopathology* **12(4)**, 857-885.

Luthar SS, Cicchetti D, Becker B (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development* **71(3)**, 543-562.

Macedo T, Wilheim L, Gonçalves R, Coutinho ESF, Vilete L, Figueira I, Ventura P (2014). Building resilience for future adversity: a systematic review of interventions in non-clinical samples of adults. *BMC Psychiatry* **14**, 227.

Marshall GN, Miles JNV, Stewart SH (2010). Anxiety sensitivity and PTSD symptom severity are reciprocally related: Evidence from a longitudinal study of physical trauma survivors. *Journal of Abnormal Psychology* **119(1)**, 143–150.

McCaslin SE, Inslicht SS, Metzler TJ, Henn-Haase C, Maguen S, Neylan TC, Choucroun G, Marmar CR (2008). Trait dissociation predicts posttraumatic stress disorder symptoms in a prospective study of urban police officers. *Journal of Nervous and Mental Disease* **196(12)**, 912–

918.

***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.

McFarlane AC, Atchison M, Rafalowicz E, Papay P (1994). Physical Symptoms in Post-traumatic Stress Disorder. *Journal of Psychosomatic Research* **38(7)**, 715-726.

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLOS ONE* **6(7)**, e1000097.

New Economics Foundation (2016). Measuring Well-being: A guide for practitioners. https://b.3cdn.net/nefoundation/8d92cf44e70b3d16e6_rgm6bpd3i.pdf

Nolen-Hoekseman S, Morrow J (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: the 1989 Loma Prieta Earthquake. *Journal of Personality and Social Psychology* **61(1)**, 115–121.

***Norris R, Carroll D, Cochrane R** (1990). The effects of aerobic and anaerobic training on fitness, blood pressure and psychological stress and well-being. *Journal of Psychosomatic Research* **34(4)**, 367–375.

***Norvell N, Belles D** (1993). Psychological and physical benefits of circuit weight training in law enforcement personnel. *Journal of Consulting and Clinical Psychology* **61(3)**, 520–527.

Ong, A. D., Bergeman, C. S., & Boker, S. M. (2009). Resilience comes of age: Defining features in later adulthood. *Journal of Personality*, 77(6), 1777-1804.

Ozer, E.J., Best, S.R., Lipsey, T.L. and Weiss, D.S. (2003). Predictors of posttraumatic stress disorder and symptoms in adults: A meta-analysis. *Psychological Bulletin*, 129, 52-73.

O'Rourke N, Kupferschimdt AL, Claxton A, Smith JZ, Chappell N, Beattie BL (2010). Psychological resilience predicts depressive symptoms among spouses of persons with Alzheimer disease over time. Aging and Mental Health **14**, 984–993.

Peñalba V, McGuire H, Leite JR (2008). Psychosocial interventions for prevention of psychological disorders in law enforcement officers. *Cochrane Database of Systematic Reviews* **3**, CD005601.

***Ramey SL, Perkhounkova Y, Hein M, Bohr NL, Anderson A** (2017). Testing a resilience training program in police recruits: A pilot study. *Biological Research for Nursing* **19(4)**, 440–449.

Reichard AA, Jackson LL (2010). Occupational Injuries Among Emergency Responders. *American Journal of Industrial Medicine* **53**, 1-11.

Rueger SY, Malecki CK (2011). Effects of stress, attributional style and perceived parental support on depressive symptoms in early adolescence: a prospective analysis. *Journal of Clinical Child and Adolescent Psychology* **40(3)**, 347–359.

Rutter, M. (1985). Resilience in the face of adversity: Protective factors and resistance to psychiatric disorder. *The British Journal of Psychiatry*, 147(6), 598-611.

Rutter M (2006). Implications of resilience concepts for scientific understanding. *Annals of the New York Academy of Sciences* **1094**, 1-12.

Samadian F, Dalili N, Jamalian A (2016). Lifestyle Modifications to Prevent and Control Hypertension. *Iranian Journal of Kidney Disease* **10(5)**, 237–263.

***Skeffington PM, Rees CS, Mazzucchelli TG, Kane RT** (2016). The primary prevention of PTSD in firefighters: Preliminary results of an RCT with 12-month follow-up. *PLOS ONE* **11(7)**, e0155873.

Solomon Z, Weisenberg M, Schwarzwald J, Mikulincer M (1987). Posttraumatic stress disorder among frontline soliders with combat stress reaction: the 1982 Israeli experience. *American Journal of Psychiatry* **144(4)**, 448–454.

***Tanigoshi H, Kontos AP, Remley TPH** (2008). The effectiveness of individual wellness counseling on the wellness of law enforcement officers. *Journal of Counselling and Development* **86(1)**, 64–74.

Topper M, Emmelkamp PMG, Watkins E, Ehring T (2017). Prevention of anxiety disorders and depression by targeting excessive worry and rumination in adolescents and young adults: A randomized controlled trial. *Behaviour Research and Therapy* **90**, 123-136.

***Tuckey MR, Scott JE** (2014). Group critical incident stress debriefing with emergency services personnel: A randomized controlled trial. *Anxiety Stress & Coping* **27(1)**, 38–54.

Van Etten ML, Taylor S (1998). Comparative Efficacy of Treatments for Post-traumatic Stress

Disorder: A Meta-Analysis. *Clinical Psychology & Psychotherapy* **5**, 126–144.

Vanhove, AJ, Herian MN, Perez ALU, Harms PD, Lester PB (2015). Can resilience be developed at work? A meta-analytic review of resilience-building programme effectiveness. *Journal of Occupational and Organizational Psychology* **89**, 278-307.

Warburton DE, Nicol CW, Bredin SS (2006). Health benefits of physical activity: the evidence. *Canadian Medical Association Journal* **174**, 801–809.

Weiss DS, Marmar CR, Metzler TJ, Ronfeldt HM (1996). Predicting symptomatic distress in emergency services personnel. *Journal of Consulting and Clinical Psychology* **63(3)**, 361–368.

Wild, J., Hackmann, A., & Clark, D.M. (2008). Rescripting early memories linked to negative images in social phobia: A pilot study. *Behavior Therapy*, 39, 47-56.

Wild J, Smith KV, Thompson E, Béar F, Lommen MJJ, Ehlers A (2016). A prospective study of pre-trauma risk factors for post-traumatic stress disorder and depression. *Psychological Medicine* **46(12)**, 2571–2582.

***Wilson SA, Tinker RH, Becker LA, Logan CR** (2001). Stress management with law enforcement personnel: A controlled outcome study of EMDR versus a traditional stress management program. *International Journal of Stress Management* **8(3)**, 179–200.

Windle, G. (2011). What is resilience? A review and concept analysis. *Reviews in Clinical Gerontology*, 21(2), 152-169.

Yi, J. P., Vitaliano, P. P., Smith, R. E., Yi, J. C., & Weinger, K. (2008). The role of resilience on psychological adjustment and physical health in patients with diabetes. *British Journal of Health Psychology*, 13(2), 311-325.

Table 1. Study Characteristics.

Study	Design (no. of arms; control(s))	Country	Population	Sample size (intervention, control)	Gender (female:male)	Intervention	Outcomes (self-report measure)	Time points
Ângelou & Chambel (2013)[32]	QCT (2; TAU)	Portugal	Firefighters	104 (67, 37)	4:100	Leadership stress management workshop	Burnout (MBI-GS) Wellbeing (UWES)	Pre, 4mFU
Arnetz *et al.* (2013)[33]	QCT (2, TAU)	Sweden	Police	75 (37, 38)	24:51	Imagery and skills training program	Wellbeing (GHQ-12) Coping (non-validated scale*) Vital exhaustion (Maastricht questionnaire) Somatic symptoms (BSS) Sleep (KSQ)	Pre, post, 18mFU
Christopher *et al.* (2018)[38]	RCT (2, TAU)	USA	Police	61 (31, 30)	7:54	Mindfulness-Based Resilience Training	Alcohol use (PROMIS) Anxiety (PROMIS) Depression (PROMIS) Sleep difficulties (PROMIS) Suicidal ideation (CHRT) Stress (PSQ) Burnout (OLBI) Mindfulness (FFMQ-SF) Wellbeing (AAQ-II, SCS-SF, BPAG-SF, PROMIS) Resilience (CD-RISC)	Pre, post, 3mFU
Ireland *et al.* (2007)[34]	QCT (2, TAU)	Australia	Police	67 (28, 39)	28:39	Writing about personal emotions	Depression (DASS) Anxiety(DASS) Stress (DASS)	Pre, post
Kuehl *et al.* (2014)[42,§]	CRCT (2, TAU)	USA	Police	408 (204, 204)	154:254	Worksite health & safety wellness program	Stress (non-validated scale†) Depression (non-validated scale†) Wellbeing (SF-36) Burnout (MBI) Sleep (PROMIS, PSQI, KSS) Tobacco (non-validated scale‡) Alcohol use (non-validated scale‡) Musculoskeletal discomfort (CMDQ)	Pre, 6mFU
McCraty & Atkinson (2012)[39]	RCT (2, waitlist)	USA	Police	65 (29, 36)	10:55	Coherence Advantage program	Wellbeing (POQA) Physical stress (POQA)	Pre, post
Norris *et al.* (1990)[35]	QCT (3, TAU)	UK	Police	150 (50 aerobic, 50 anaerobic, 50 control)	0:150	Aerobic and anaerobic exercise training	Wellbeing (GHQ-28) Stress (JSQ) Quality of life (LSS)	Pre, post

Norvell & Belles (1993)[36]	QCT (3, waitlist + dropouts)	USA	Police	43 (14, 15 waitlist, 14 dropouts)	0:43	Circuit weight training	Stress (PSS) Wellbeing (SCL-90) Physical symptoms (CHIPS)	Pre, post
Ramey *et al.* (2017)[37]	QCT (2, placebo)	USA	Police	34 (17, 17)	6:28	Physiological coherence practice and telementor sessions	Stress (PSS) Vital exhaustion (Maastricht questionnaire 9-item version of Form B) PTSD (IES) Wellbeing (POQA) Resilience (RSES) Physical stress (POQA)	Pre, post
Skeffington *et al.* (2016)[43]	CRCT (2, TAU)	Australia	Firefighters	77 (30, 45)	4:73	Mental Agility and Psychological Strength training	PTSD (PCL-C) Depression (DASS) Anxiety (DASS) Stress (DASS) Coping (Brief COPE)	Pre, 6mFU, 12mFU
Tanigoshi *et al.* (2008)[40]	RCT (2, TAU)	USA	Police	51 (24, 27)	9:42	Wellness counselling	Wellbeing (5F-Wel)	Pre, post
Tuckey & Scott (2014)[44]	CRCT (3, TAU + placebo)	Australia	Firefighters	67 (20 CISD, 28 education,19 control)	6:61	Group CISD	Wellbeing (K-10) PTSD (IES-R) Quality of life (Q-LES-Q) Alcohol use (non-validated scale¶)	Pre, post
Wilson *et al.* (2001)[41]	RCT (2, placebo)	USA	Police	62 (33, 29)	13:49	EMDR	Wellbeing (SCL-90, SUDS, STAXI) PTSD (PTDS) Stress (PSI, JSS) Coping (CRI)	Pre, post, 6mFU

RCT=randomised controlled trial; **QCT**=quasi-randomised controlled trial; **CRCT**=cluster randomised controlled trial; **TAU**=treatment as usual; **CISD=**Critical Incident Stress Debriefing; **EMDR**=Eye Movement Desensitization and Reprocessing; **MBI-GS**=Maslach Burnout Inventory - General Survey; **UWES**=Utrecht Work Engagement Scale; **GHQ-12**=General Health Questionnaire-12; **Maastricht Questionnaire**; **BSS**=Bodily Symptoms Scale; **KSQ**=Karolinska Sleep Questionnaire; **PROMIS**=National Institutes of Health Patient-Reported Outcomes Information System; **CHRT**=Concise Health Risk Tracking scale; **PSQ**=Police Stress Questionnaire; **OLBI**=Oldenburg Burnout Inventory; **FFMQ-SF=**Five Facet Mindfulness Questionnaires-Short Form; **AAQ-II**=Acceptance and Action Questionnaire-II; **SCS-SF**=Self-Compassion Scale-Short Form; **DASS**=Depression Anxiety Stress Scales; **SF-36**=Short Form-36; **MBI=**Maslach Burnout Inventory; **PSQI**=Pittsburgh Sleep Quality Index; **KSS**=Karolinska Sleepiness Scale; **CMDQ**=Cornell Musculoskeletal Discomfort Questionnaire; **POQA**=Personal and Organizational Quality Assessment; **GHQ-28**=General Health Questionnaire-28; **JSQ**=Job Stress Questionnaire; **LSS**=Life Situation Survey; **PSS**=Perceived Stress Scale; **SCL-90**=Symptom Checklist-90; **CHIPS**=Cohen-Hoberman Inventory of Physical Symptoms; **9-item Maastricht questionnaire**; **IES**=Impact of Events Scale; **RSES**=Response to Stressful Experiences Scale; **PCL-C**=The PTSD Checklist – Civilian Version; **Brief COPE**=Brief Coping Orientations to Problems Experienced; **5F-Wel**=Five Factor Wellness Inventory; **K-10**=Kessler Psychological Distress Scale; **IES-R**=Impact of Events Scale – Revised; **Q-LES-Q**=Quality of Life Enjoyment and Satisfaction Questionnaire; **SUDS=**Subjective Units of Disturbance Scale; **STAXI=**State-Trait Anger Expression Inventory; **PTDS**=Posttraumatic Stress Diagnostic Scale; **PSI**=Police Stress Inventory; **JSS**=Job Stress Survey; **CRI**=Coping Responses Inventory; **mFU**=months follow up.

*Brief non-validated 3-item coping measure used.
†Used constructs with established reliability from authors' previous studies, but unclear what constructs are.
‡Used constructs with established reliability from authors' previous studies, but unclear what constructs are.
§The authors published a paper reporting results at six-months (Kuehl et al., 2014) and a final-results paper (Kuehl et al., 2016). Insufficient data was provided in the final-results paper, so all data included in this review are extracted from the paper reporting six-month findings.
¶Participants were asked how many standard alcohol drinks they had in the previous seven-day period.

Table 2. Intervention characteristics.

Study	Intervention Type	Intervention Content	Format	Modifiable risk factor targeted
Ângelou & Chambel (2013)[32]	Stress management	Leadership stress management workshop for the supervisors of participants, involving psycho-education, coping strategies, support systems, leadership roles, problem solving	3 consecutive days, each lasting 7 hours, led by the researchers	None*
Arnetz *et al.* (2013)[33]	Stress management	Imagery and skills training program involving psycho-education, relaxation training, guided imagery, mental practice of police tactical skills	10 weekly sessions, each lasting 1.5 hours, led by Special Forces officers who were trained by the researchers	Behavioural disengagement – participants practiced adaptive coping strategies (e.g. identifying specific goals and enacting plans to achieve them) during imaginal exposure to stressful incidents
Christopher *et al.* (2018)[38]	Stress management	Experiential and didactic exercises, including body scan, sitting and walking meditations, mindful movement, and group discussion. Participants were given an iPod Touch programmed with guided practices and monitoring software to supplement in-session content and support practice between sessions	8 weekly sessions, each lasting 2 hours except the extended 6-hour final class	None†
Ireland *et al.* (2007)[34]	Stress management	Writing about strong emotions related to work or not, and what participants planned to do about them	4 consecutive days, each lasting 15 minutes, written instructions provided	Emotional suppression – participants wrote about their strong emotions and what they would do as a result of the strong emotions
Kuehl *et al.* (2014)[42]	Stress management	Worksite health and safety wellness classes providing information about healthy eating, exercise, body weight, stress, sleep, other lifestyle factors	12 weekly sessions, each lasting 30 minutes, peer-led	None‡
McCraty & Atkinson (2012)[39]	Self-regulation	Group classes involving self-regulation skills (e.g., Heart Focused Breathing, Freeze Frame, Inner Ease, Prep Shift and Reset, Getting In Sync) and technology (emWave) for achieving physiological coherence	3 sessions spaced evenly over 1 month, each lasting 4 hours, led by trained instructors	None‡
Norris *et al.* (1990)[35]	Physical exercise	Aerobic exercise involved 5-10 minutes of warming up, 20-30 minutes of road running then 5-10 minutes cooling down stretches. The anaerobic exercise involved 5-10 minutes of warming up and 20-30 minutes of circuit training	At least 25 sessions, 3 times a week, each lasting 1 hour, led by trained instructors	Physical inactivity – participants regularly exercised
Norvell & Belles (1993)[36]	Physical exercise	Circuit weight training using 12 circuit machines in a gym	48 sessions, 3 times a week, each lasting 20 minutes, monitored by gym staff	Physical inactivity – participants regularly exercised

Ramey *et al.* (2017)[37]	Self-regulation	Initial group class covering psycho-education and self-regulation techniques to alter physiological coherence. All participants received flipcharts and booklets to guide home practice, but the intervention group also received the HeartMath Inner Balance application, an earlobe sensor and mentoring via four phone conferences sessions	5 sessions, once every 2-3 weeks, initial class lasting 2 hours led by researchers, phone conferences lasting 1 hour and led by mental health professionals	None‡
Skeffington *et al.* (2016)[43]	Stress management	Group class covering psycho-education (coping strategies, stress, PTSD), defusion exercises, practical skills, self-care	4 weekly sessions, each lasting 60 minutes, led by primary researcher	None*
Tanigoshi *et al.* (2008)[40]	Psychological	Individual wellness counselling based on Myers and Sweeney's (2004) Indivisible Self Model (an evidence-based wellness model that views wellness holistically across the life span), cognitive-behavioural intervention strategies	5 sessions, once every 2 weeks, each lasting 1 hour, led by trained counsellors	Behavioural disengagement – participants identified area of wellness they scored low in at the start of treatment and then created treatment goals and personalised wellness plans to work towards
Tuckey & Scott (2014)[44]	Debriefing	Group CISD sessions following the Mitchell (1983) and Mitchell and Everly (1993) seven-phase protocol: (1) Introduction, (2) Facts, (3) Thoughts, (4) Reactions, (5) Symptoms, (6) Education, and (7) Re-entry.	1 session lasting 90 minutes, led by trained mental health professionals and a peer supporter for the introduction and education phases, led by psychologists and social workers	None§
Wilson *et al.* (2001)[41]	Psychological	Initial clinical interview exploring job stress. Stressors identified in interview used in one-to-one EMDR sessions following the standard protocol (Shapiro, 1995)	2-hour interview followed by 3 sessions, each lasting 2 hours, led by psychologists and social workers with experience of working with police officers	None¶

CISD=Critical Incident Stress Debriefing; **EMDR**=Eye Movement Desensitization and Reprocessing

*Provided psychoeducation about stress, coping and mental health

†Targeted mental focus, sustained attention and a sense of personal and situational awareness

‡Targeted physiological coherence - a functional mode where a person's heart rhythm pattern becomes more ordered[39]

§Information provided about potential symptoms and coping strategies

¶Targeted memory processing

Table 3. The effectiveness of included studies.

Intervention	Outcome sub-categories	Time points (T)	Within-group comparisons				Between-group comparisons	
			T1 to T2 (Intervention group)	T1 to T2 (Control group)	T1 to T3 (Intervention group)	T1 to T3 (Control group)	T2	T3
Leadership stress management workshop[32]	Wellbeing (burnout - emotional exhaustion)	Pre (T1)	-0·09 (-0·43, 0·25)	-0·28 (-0·74, 0·17)			-0·13 (-0·53, 0·28)	
	Wellbeing (burnout - cynicism)	4mFU (T2)	-0·08 (-0·41, 0·26)	-0·11 (-0·56, 0·35)			0·08 (-0·33, 0·48)	
	Wellbeing (engagement - vigour)		0·00 (-0·34, 0·34)	-0·31 (-0·77, 0·15)			0·13 (-0·27, 0·53)	
	Wellbeing (engagement - dedication)		0·10 (-0·24, 0·44)	-0·15 (-0·61, 0·30)			0·05 (-0·35, 0·45)	
Imagery and skills training program[33†]	Wellbeing	Pre (T1)	Data not reported at T2	Data not reported at T2	-0·48 (-0·98, 0·03)	-0·09 (-0·57, 0·38)	Data not reported at T2	-0·34 (-0·83, 0·15)
	Coping	Post (T2)			0·49 (-0·01, 0·99)	0·12 (-0·36, 0·59)		0·95 (0·45, 1·45)*
	Vital exhaustion[‡]	18mFU (T3)			0·58 (0·08, 1·08)*	-0·13 (-0·61, 0·34)		0·60 (0·11, 1·09)*
	Sleep				0·26 (-0·24, 0·76)	-0·17 (-0·65, 0·30)		0·52 (0·03, 1·01)*
	Somatic symptoms (stomach problems and heart complaints)				0·19 (-0·31, 0·69)	-0·34 (-0·82, 0·13)		0·31 (-0·18, 0·80)
Mindfulness-Based Resilience Training[38]	Alcohol use	Pre (T1)	0·33 (-0·20, 0·87)	-0·06 (-0·59, 0·46)	0·10 (-0·43, 0·63)	-0·03 (-0·43, 0·63)	0·38 (-0·18, 0·93)	0·12 (-0·44, 0·68)
	Anxiety	Post (T2)	0·34 (-0·20, 0·87)	0·51 (-0·02, 1·04)	0·14 (-0·40, 0·67)	0·45 (-0·09, 0·98)	-0·02 (-0·57, 0·54)	-0·17 (-0·73, 0·39)
	Depression	3mFU (T3)	0·10 (-0·43, 0·64)	0·16 (-0·37, 0·68)	-0·08 (-0·61, 0·46)	0·29 (-0·24, 0·83)	0·10 (-0·46, 0·65)	-0·24 (-0·80, 0·32)
	Sleep difficulties		0·4 (-0·11, 0·96)	0·13 (-0·39, 0·66)	0·13 (-0·41, 0·66)	0·08 (-0·46, 0·60)	0·60 (0·04, 1·16)*	0·23 (-0·33, 0·79)
	Suicidal ideation		0·09 (-0·44, 0·62)	0·30 (-0·22, 0·83)	0·03, -0·50, 0·57)	0·28 (-0·26, 0·81)	-0·28 (-0·84, 0·28)	-0·39 (-0·96, 0·17)
	Organisational stress		0·28 (-0·26, 0·81)	-0·12 (-0·65, 0·40)	0·18 (-0·35, 0·72)	0·13 (-0·40, 0·66)	0·53 (-0·03, 1·09)	0·15 (-0·41, 0·71)
	Operational stress		0·21 (-0·32, 0·74)	0·06 (-0·47, 0·59)	0·16 (-0·38, 0·69)	0·03 (-0·50, 0·56)	0·09 (-0·46, 0·65)	0·05 (-0·51, 0·61)
	Burnout		0·50 (-0·04, 1·04)	-0·03 (-0·55, 0·50)	0·34 (-0·19, 0·88)	0·18 (-0·35, 0·72)	0·73 (0·17, 1·30)*	0·38 (-0·18, 0·94)
	Resilience		0·19 (-0·35, 0·72)	0·10 (-0·42, 0·63)	0·14 (-0·39, 0·68)	0·14 (-0·39, 0·67)	⁋	⁋
	Anger		0·17 (-0·37, 0·70)	0·28 (-0·25, 0·80)	0·04 (-0·49, 0·57)	0·38 (-0·15, 0·92)	0·09 (-0·47, 0·64)	-0·14 (-0·71, 0·42)
	Aggression		0·74 (0·20, 1·28)*	0·20 (-0·32, 0·73)	0·47 (-0·07, 1·01)	0·40 (-0·13, 0·94)	0·53 (-0·03, 1·10)	0·06 (-0·50, 0·62)
	Nonreactivity		0·69 (0·15, 1·23)*	0·10 (-0·42, 0·63)	0·29 (-0·25, 0·82)	0·38 (-0·15, 0·92)	0·63 (0·07, 1·20)*	0·03 (-0·53, 0·59)
	Nonjudging		0·55 (0·01, 1·09)*	0·61 (0·08, 1·14)*	0·21 (-0·33, 0·74)	0·30 (-0·23, 0·83)	-0·21 (-0·76, 0·35)	-0·29 (-0·85, 0·27
	Acting with awareness		0·04 (-0·49, 0·57)	0·05 (-0·48, 0·57)	0·17 (-0·36, 0·70)	0·32 (-0·21, 0·86)	0·01 (-0·55, 0·56)	-0·16 (-0·72, 0·40)
	Psychological flexibility		0·42 (-0·12, 0·95)	0·11 (-0·41, 0·64)	0·25 (-0·28, 0·78)	0·62 (0·09, 1·16)*	0·36 (-0·19, 0·92)	-0·32 (-0·88, 0·24)
	Self-compassion		0·17 (-0·37, 0·70)	0·20 (-0·33, 0·72)	0·25 (-0·29, 0·78)	0·34 (-0·19, 0·88)	⁋	⁋
Writing about personal emotions[34 †]	Mental health symptomatology (depression)	Pre (T1)	0·10 (-0·42, 0·63)	-0·28 (-0·73, 0·16)			0·20 (-0·28, 0·69)	
	Mental health symptomatology (anxiety)	Post (T2)	0·32 (-0·21, 0·85)	-0·31 (-0·75, 0·14)			0·23 (-0·25, 0·72)	
	Stress		0·56 (0·03, 1·09)*	-0·25 (-0·69, 0·19)			0·39 (-0·10, 0·88)	
Worksite health & safety wellness program[42 †]	Wellbeing (general health status)	Pre (T1)	0·69 (0·49, 0·89)*	0·12 (-0·08, 0·32)			0·34 (0·13, 0·55)*	
	Wellbeing (burnout)	6mFU (T2)	0·09 (-0·11, 0·29)	-0·07 (-0·27, 0·13)			-0·08 (-0·29, 0·13)	
	Stress		0·35 (0·15, 0·56)*	0·06 (-0·14, 0·26)			0·11 (0·16, 0·58)*	
	Mental health symptomatology (depression)		-0·06 (-0·26, 0·14)	0·26 (0·06, 0·46)*			-0·36 (-0·57, -0·15)*	
	Sleep (sleep quality)		0·64 (0·44, 0·84)*	-0·05 (-0·25, 0·15)			0·69 (0·48, 0·90)*	
	Sleep (sleep quantity)		0·53 (0·33, 0·73)*	0·03 (-0·17, 0·24)			0·65 (0·44, 0·86)*	
	Sleep (sleepiness)		0·29 (0·09, 0·50)*	0·50 (0·30, 0·70)*			-0·24 (-0·44, -0·03)*	
	Sleep (fatigue)		-0·20 (-0·40, 0·00)	-0·01 (-0·22, 0·19)			-0·55 (-0·76, -0·34)*	
	Somatic symptoms (musculoskeletal pain)		0·25 (0·05, 0·45)*	0·25 (0·05, 0·45)*			-0·11 (-0·32, 0·10)	
	Somatic symptoms (musculoskeletal pain with foot pain)		0·13 (-0·07, 0·33)	0·25 (0·05, 0·45)*			-0·22 (-0·43, -0·01)*	
	Alcohol use		-0·09 (-0·29, 0·11)	-0·03 (-0·23, 0·18)			-0·11 (-0·32, 0·10)	
	Tobacco use		-0·10 (-0·30, 0·10)	-0·07 (-0·28, 0·13)			0·35 (0·14, 0·56)*	
	Physical activity		0·49 (0·28, 0·69)*	0·56 (0·35, 0·76)*			0·13 (-0·08, 0·34)	
	Healthy eating		0·82 (0·62, 1·03)*	0·55 (0·35, 0·76)*			⁋	

Coherence Advantage program[39]	Wellbeing (global negative emotion)	Pre (T1) Post (T2)	0·41 (-0·12, 0·93)	-0·06 (-0·56, 0·43)			0·32 (-0·20, 0·83)	
	Wellbeing (anger)		0·20 (-0·33, 0·72)	0·12 (-0·38, 0·62)			0·17(-0·34, 0·68)	
	Wellbeing (distress)		0·60 (0·07, 1·13)*	0·04 (-0·46, 0·54)			0·24 (-0·27, 0·76)	
	Mental health symptomatology (depression)		0·21 (-0·31, 0·73)	-0·37 (-0·87, 0·13)			0·51 (0·00, 1·03)	
	Wellbeing (sadness)		0·35 (-0·17, 0·88)	-0·08 (-0·58, 0·42)			0·25 (-0·26, 0·76)	
	Wellbeing (fatigue)		0·48 (-0·05, 1·01)	-0·01 (-0·51, 0·48)			0·04 (-0·47, 0·55)	
	Wellbeing (positive emotion)		0·28 (-0·25, 0·80)	0·14 (-0·36, 0·64)			-0·04 (-0·55, 0·48)	
	Wellbeing (peacefulness)		0·38 (-0·14, 0·91)	-0·05 (-0·55, 0·45)			0·32 (-0·19, 0·84)	
	Wellbeing (vitality)		0·18 (-0·35, 0·70)	-0·09 (-0·59, 0·41)			0·19 (-0·33, 0·70)	
	Sleep (sleeplessness)		0·46 (-0·07, 0·99)	-0·20 (-0·70, 0·30)			0·10 (-0·41, 0·61)	
	Somatic symptoms (physical anxiety)		0·44 (-0·08· 0·97)	0·15 (-0·35, 0·65)			-0·16 (-0·67, 0·35)	
	Somatic symptoms (body aches)		0·15 (-0·38, 0·67)	-0·10 (-0·59, 0·40)			-0·03 (-0·54, 0·48)	
	Somatic symptoms (indigestion)		0·10 (-0·42, 0·63)	-0·18 (-0·68, 0·32)			-0·10 (-0·61, 0·41)	
Aerobic and anaerobic exercise training[35, †]	Wellbeing	Pre (T1) Post (T2)	Aerobic: 0·80 (0·27, 1·34)* Anaerobic: 2·71 (1·99, 3·42)*	-0·08 (-0·63, 0·48)			Aerobic vs anaerobic: 0·26 (-0·29, 0·80) Aerobic vs control: 0·83 (0·28, 1·39)* Anaerobic vs control: 0·63 (0·06, 1·20)*	
	Stress		Aerobic: 0·56 (0·03, 1·09)* Anaerobic: 0·34 (-0·23, 0·91)	-0·09 (-0·65, 0·46)			Aerobic vs anaerobic: ⁋ Aerobic vs control: 0·72 (0·17, 1·27)* Anaerobic vs control: ⁋	
	Quality of life		Aerobic: 0·50 (-0·03, 1·03) Anaerobic: 0·40 (-0·17, 0·96)	-0·62 (-1·18, -0·06)*			Aerobic vs anaerobic: 0·17 (-0·37, 0·72) Aerobic vs control:0·94 (0·38, 1·49)* Anaerobic vs control: 0·72 (0·15, 1·29)*	
Circuit weight training[36, †]	Wellbeing	Pre (T1) Post (T2)	1·06 (0·28, 1·84)*	Waitlist: 0·08 (-0·63, 0·80) Dropouts: 0·09 (-0·65, 0·83)			Intervention vs waitlist: 1·00 (0·24, 1·76)* Intervention vs dropouts: ⁋ Waitlist vs dropouts: ⁋	
	Stress		0·40 (-0·35, 1·15)	Waitlist: -0·02 (-0·79, 0·69) Dropouts: -0·05 (-0·79, 0·69)			Intervention vs waitlist: 0·61 (-0·13, 1·35) Intervention vs dropouts: ⁋ Waitlist vs dropouts: ⁋	
	Somatic symptoms		0·78 (0·02, 1·55)*	Waitlist: -0·03 (-0·75, 0·68) Dropouts: -0·13 (-0·88, 0·61)			Intervention vs waitlist: 1·36 (0·57, 2·15)* Intervention vs dropouts: ⁋ Waitlist vs dropouts: ⁋	
Physiological coherence practice and telementor sessions[37]	Wellbeing (emotional vitality)	Pre (T1) Post (T2)	0·22 (-0·46, 0·91)	0·12 (-0·56, 0·81)			0·00 (-0·68, 0·68)	
	Stress (perceived stress)		0·05 (-0·63, 0·73)	0·19 (-0·48, 0·87)			-0·06 (-0·74, 0·62)	
	Stress (organizational stress)		0·17 (-0·51, 0·86)	0·30 (-0·39, 0·98)			-0·27 (-0·96, 0·41)	
	Stress (emotional stress)		0·16 (-0·52, 0·85)	-0·20 (-0·88, 0·48)			-0·16 (-0·85, 0·52)	
	Mental health symptomatology (PTSD)		-0·30 (-0·98, 0·39)	-0·24 (-0·91, 0·43)			0·04 (-0·65, 0·72)	
	Vital exhaustion		-0·03 (-0·71, 0·65)	0·16 (-0·53, 0·84)			0·00 (-0·69, 0·69)	
	Resilience		0·02 (-0·66, 0·71)	0·16 (-0·52, 0·84)			-0·26 (-0·95, 0·42)	
	Somatic symptoms (physical stress)		0·00 (-0·68, 0·68)	0·00 (-0·68, 0·68)			-0·11 (-0·79, 0·57)	
	Somatic symptoms (health symptoms)		0·00 (-0·68, 0·68)	-0·17 (-0·85, 0·52)			0·15 (-0·53, 0·84)	
Mental Agility and Psychological Strength training[43]	Mental health symptomatology (PTSD)	Pre (T1) 6mFU (T2) 12mFU (T3)	-0·22 (-0·75, 0·30)	0·55 (0·10, 0·99)*	-0·08 (-0·64, 0·48)	0·22 (-0·24, 0·68)	-0·39 (-0·90, 0·12)	0·04 (-0·51, 0·59)
	Mental health symptomatology (depression)		-0·64 (-1·17, -0·11)*	0·38 (-0·06, 0·82)	-0·59 (-1·15, -0·02)*	-0·02 (-0·47, 0·44)	-0·55 (-1·06, -0·04)*	-0·16 (-0·72, 0·39)
	Mental health symptomatology (anxiety)		0·34 (-0·19, 0·86)	0·53 (0·09, 0·98)*	0·43 (-0·13, 0·99)	0·31 (-0·15, 0·77)	-0·15 (-0·66, 0·36)	0·14 (-0·41, 0·70)
	Stress		0·20 (-0·33, 0·72)	0·90 (0·45, 1·36)*	0·19 (-0·37, 0·75)	0·58 (0·12, 1·04)*	-0·73 (-1·25, -0·21)*	-0·43 (-0·99, 0·12)
	Coping (adaptive coping)		-0·44 (-0·96, 0·09)	-0·23 (-0·68, 0·21)	-0·69 (-1·26, -0·13)*	-0·25 (-0·71, 0·20)	-0·08 (-0·59, 0·43)	-0·03 (-0·58, 0·53)
	Coping (maladaptive coping)		0·26 (-0·26, 0·79)	0·57 (0·12, 1·01)*	0·16 (-0·40, 0·71)	0·55 (0·09, 1·02)*	-0·26 (-0·77, 0·24)	-0·37 (-0·93, 0·19)

Wellness counselling[40, †]	Wellbeing	Pre (T1) Post (T2)	0·74 (0·16, 1·31)*	0·00 (-0·53, 0·53)			0·52 (-0·03, 1·08)	
Group CISD[44, †]	Wellbeing	Pre (T1) Post (T2)	CISD: 0·18 (-0·44, 0·80) Education: 0·10 (-0·43, 0·62)	0·44 (-0·20, 1·08)			CISD vs Education: 0·22 (-0·36, 0·79) CISD vs control: -0·20 (-0·83, 0·43) Education vs control: -0·44 (-1·02, 0·15)	
	Mental health symptomatology (PTSD)		CISD: 0·40 (-0·23, 1·02) Education: 0·05 (-0·47, 0·58)	0·25 (-0·39, 0·89)			CISD vs Education: 0·11 (-0·47, 0·68) CISD vs control: -0·15 (-0·78, 0·47) Education vs control: -0·21 (-0·78, 0·47)	
	Quality of life		CISD: 0·22 (-0·40, 0·84) Education: -0·06 (-0·59, 0·46)	-0·82 (-1·48, -0·17)*			CISD vs Education: 0·25 (-0·32, 0·83) CISD vs control: -0·23 (-0·86, 0·40) Education vs control: -0·53 (-1·12, 0·05)	
	Alcohol use		CISD: 0·04 (-0·59, 0·66) Education: 0·01 (-0·51, 0·53)	-0·62 (-1·26, 0·03)			CISD vs Education: ⁋ CISD vs control: ⁋ Education vs control: -0·44 (-1·02, 0·15)	
EMDR[41, †]	Wellbeing (psychological functioning)	Pre (T1)	Insufficient data	Insufficient data	Data not reported at T3	Data not reported at T3	Insufficient data	Insufficient data
	Wellbeing (distress)	Post (T2)	2·17 (1·61, 2·72)*	0·60 (0·08, 1·12)*			1·47 (0·94, 2·01)*	Insufficient data
	Wellbeing (trait anger)	6mFU (T3)	Insufficient data	Insufficient data			0·62 (0·11, 1·12)*	1·20 (0·67, 1·72)*
	Wellbeing (state anger)		Insufficient data	Insufficient data			-0·12 (-0·62, 0·38)	0·58 (0·07, 1·08)*
	Stress (job stress)		0·22 (-0·26, 0·71)	-0·50 (-1·02, 0·02)			0·84 (0·33, 1·35)*	Insufficient data
	Stress (police stress)		Insufficient data	Insufficient data			Insufficient data	Insufficient data
	Mental health symptomatology (PTSD)		Insufficient data	Insufficient data			Insufficient data	Insufficient data
	Coping		Insufficient data	Insufficient data			Insufficient data	Insufficient data

Data are effect sizes (95% confidence intervals)

*Significant effect

†Targets a modifiable risk factor

‡Items were reverse coded so that higher scores indicated less vital exhaustion

§Total mean of all participants given at baseline (not split by condition)

⁋Significant differences between groups at baseline