Effects of mental health interventions for students in higher education are sustainable over time: a systematic review and meta-analysis of randomized controlled trials

Regina Winzer Correspond., Lene Lindberg, Karin Guldbrandsson, Anna Sidorchuk

1 Department of Public Health Sciences, Karolinska Institute, Stockholm, Sweden
2 Department of Living Conditions and Lifestyles, The Public Health Agency of Sweden, Solna, Sweden
3 Department of Public Health Sciences, Karolinska Institute, Solna, Sweden
4 Center for Epidemiology and Community Medicine, Stockholm County Council, Stockholm, Sweden
5 Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institute, Stockholm, Sweden

Corresponding Author: Regina Winzer
Email address: regina.winzer@folkhalsomyndigheten.se

Background. Symptoms of depression, anxiety, and distress are more common in undergraduates compared to age-matched peers. Mental ill health among students is associated with impaired academic achievement, worse occupational preparedness, and lower future occupational performance. Research on mental health promoting and mental ill health preventing interventions has shown promising short-term effects, though the sustainability of intervention benefits deserve closer attention. We aimed to identify, appraise and summarize existing data from RCTs reporting on whether the effects of mental health promoting and mental ill health preventing interventions sustained at least 3 months post-intervention, and to analyse how the effects vary for different outcomes in relation to follow-up length. Further, we aimed to assess whether the effect sustainability varied by intervention type, study-level determinants and of participant characteristics.

Material and methods. A systematic search in MEDLINE, PsycInfo, ERIC and Scopus was performed for RCTs published in 1995-2015 and reporting an assessment of mental ill health and positive mental health outcomes for, at least, 3 months of post-intervention follow-up. Random-effect modeling was utilized for quantitative synthesis of the existing evidence with standardized mean difference (Hedges’ g) used to estimate an aggregated effect size. Sustainability of the effects of interventions was analysed separately for 3-6 months, 7-12 months and 13-18 months of post-intervention follow-up. Results. 26 studies were eligible after reviewing 6571 citations. The pooled effects were mainly small, but significant for several categories of outcomes. Thus, for the combined mental ill health outcomes, symptom reduction sustained up to 7-12 months post-intervention [standardized mean difference (Hedges’ g) effect size (ES)=-0.28 (95%CI; -0.49, -0.08)]. Further, sustainability of symptom-reductions were evident for depression with
intervention effect lasting up to 13-18 months [ES=-0.30 (95%CI; -0.51, -0.08)], for anxiety up to 7-12 months [ES=-0.27 (95%CI; -0.54, -0.01)], and for stress up to 3-6 months [ES= -0.30 (95%CI;-0.58, -0.03)]. The effects of interventions to enhance positive mental health sustained up to 3-6 months for the combined positive mental health outcomes [ES=0.32 (95%CI; 0.05, 0.59)]. For enhanced active coping, sustainability up to 3-6 months was observed with a medium and significant effect [ES=0.75 (95%CI; 0.19, 1.30)].

**Discussion.** The evidence suggests long-term effect sustainability for mental ill health preventive interventions, specially, for interventions to reduce the symptoms of depression and symptoms of anxiety. Interventions to promote positive mental health offer promising, but shorter-lasting effects. Future research should focus on mental health organizational interventions to examine their potential for students in tertiary education.
Effects of mental health interventions for students in higher education are sustainable over time: A systematic review and meta-analysis of randomized controlled trials

Regina Winzer¹,², Lene Lindberg¹,³, Karin Gulbrandsson¹,², Anna Sidorchuk¹,⁴

1 Department of Public Health Sciences, Karolinska Institutet, Stockholm, Sweden
2 Department of Living Conditions and Lifestyles, The Public Health Agency of Sweden, Solna, Sweden
3 Center for Epidemiology and Community Medicine Stockholm County Council, Stockholm, Sweden
4 Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Regina Winzer¹,²
regina.winzer@folkhalsomyndigheten.se
Abstract

Background

Symptoms of depression, anxiety, and distress are more common in undergraduates compared to age-matched peers. Mental ill health among students is associated with impaired academic achievement, worse occupational preparedness, and lower future occupational performance. Research on mental health promoting and mental ill health preventing interventions has shown promising short-term effects, though the sustainability of intervention benefits deserve closer attention. We aimed to identify, appraise and summarize existing data from RCTs reporting on whether the effects of mental health promoting and mental ill health preventing interventions sustained at least 3 months post-intervention, and to analyse how the effects vary for different outcomes in relation to follow-up length. Further, we aimed to assess whether the effect sustainability varied by intervention type, study-level determinants and of participant characteristics.

Material and methods

A systematic search in MEDLINE, PsycInfo, ERIC and Scopus was performed for RCTs published in 1995-2015 and reporting an assessment of mental ill health and positive mental health outcomes for, at least, 3 months of post-intervention follow-up. Random-effect modeling was utilized for quantitative synthesis of the existing evidence with standardized mean difference
(Hedges’ g) used to estimate an aggregated effect size. Sustainability of the effects of interventions was analysed separately for 3-6 months, 7-12 months and 13-18 months of post-intervention follow-up.

**Results**

26 studies were eligible after reviewing 6571 citations. The pooled effects were mainly small, but significant for several categories of outcomes. Thus, for the combined mental ill health outcomes, symptom reduction sustained up to 7-12 months post-intervention [standardized mean difference (Hedges’ g) effect size (ES)= -0.28 (95%CI; -0.49, -0.08)]. Further, sustainability of symptom-reductions were evident for depression with intervention effect lasting up to 13-18 months [ES=-0.30 (95%CI; -0.51, -0.08)], for anxiety up to 7-12 months [ES=-0.27 (95%CI; -0.54, -0.01)], and for stress up to 3-6 months [ES= -0.30 (95%CI; -0.58, -0.03)]. The effects of interventions to enhance positive mental health sustained up to 3-6 months for the combined positive mental health outcomes [ES=0.32 (95%CI; 0.05, 0.59)]. For enhanced active coping, sustainability up to 3-6 months was observed with a medium and significant effect [ES=0.75 (95%CI; 0.19, 1.30)].

**Discussion**

The evidence suggests long-term effect sustainability for mental ill health preventive interventions, specially, for interventions to reduce the symptoms of depression and symptoms of anxiety. Interventions to promote positive mental health offer promising, but shorter-lasting
effects. Future research should focus on mental health organizational interventions to examine their potential for students in tertiary education.

INTRODUCTION

Mental health problems among students in higher education is an emerging public health issue and evidence-based prevention is essential (Christensson et al. 2010; Dahlin et al. 2011; Garlow et al. 2008; Hunt & Eisenberg 2010; Steptoe et al. 2007). Recent systematic reviews on student health raise concerns over high rates of mental ill health outcomes with pooled prevalence ranging between 27% and 34% for depression and depressive symptoms and reaching 11% for suicidal ideation (Ibrahim et al. 2013; Rotenstein et al. 2016; Tung et al. 2018). Also, a two-fold risk for suicide is shown during ongoing university studies compared to when having attained university studies (Lageborn et al. 2017). Elevated rates of mental ill health, namely symptoms of distress, anxiety and depression, in undergraduates appear to substantially exceed the corresponding estimates in age-matched peers (Cvetkovski et al. 2012; Dyrbye et al. 2006; Leahy et al. 2010; Winzer et al. 2014) and the general population (Ibrahim et al. 2013; Rotenstein et al. 2016). Female students, minority groups and students with financial problems constitute groups with higher risks (Cvetkovski et al. 2012; Eisenberg et al. 2013; Said et al. 2013). Once heightened at the beginning of the study period, the symptoms of anxiety and depression remain elevated over the academic years and at no time point drop down to pre-registration levels (Bewick et al. 2010). Mental ill health among students may potentially be
caused by heavy workload, insufficient feedback from teachers and worries about future endurancer/competence (Dahlin et al. 2005), but may also reflect the increase in deteriorated mental health among adolescents (Hunt & Eisenberg 2010). Mental ill health problems are often accompanied by decrements in positive mental health through lowered self-perception, inadequate social-emotional skills and poor interpersonal relationships (Conley et al. 2015). Moreover, perceived academic stress and burn-out are associated with impaired academic achievement (Andrews 2004; Keyes et al. 2012; Vaez & Laflamme 2008), worse occupational preparedness and lower occupational performance after the graduation (Rudman & Gustavsson 2012). In prevention science the health promotion approach constitutes a substantial ingredient of the integrative model for mental health intervention in youth (Weisz et al. 2005). Including aspects of positive mental health, i.e. emotional, psychological and social well-being (Westerhof & Keyes 2010) is a beneficial strategy in mental health interventions (Kobau et al. 2011). It has been shown that psychological assets, e.g. boostering positive emotions, coping strategies and compassion may help people to manage life’s challenges. Thus, promoting mental health and preventing mental ill health are two essential and complementary steps in reducing the burden of disease (Jane-Llopis 2007; Keyes 2007; World Health Organization (WHO) 2002).

Previous research on mental health promotion and mental ill health prevention has shown promising short-term effects of stress reduction techniques and meditation, self-hypnosis, cognitive behavioural and mindfulness interventions (Conley et al. 2015; Conley et al. 2017; Regehr et al. 2013; Shiralkar et al. 2013) as well as of technology based interventions (Conley et al. 2016; Davies et al. 2014; Farrer et al. 2013). As mental health problems persist during the study period (Bewick et al. 2010; Christensson et al. 2010) and negatively affect academic performance and future working capacity (Rudman & Gustavsson 2012; Vaez & Laflamme
2008), the sustainability of intervention benefits as well as its determinants and moderators deserve closer attention. Several reviews have approached the issue of intervention effect sustainability by averaging the effects reported for the longest follow-up periods, although a substantial variability in the ranges and means of the follow-up lengths made the comparisons difficult (Conley et al. 2015; Conley et al. 2016; Conley et al. 2017). The authors highlighted the need for in-depth investigation of the intervention benefit sustainability over variable postintervention follow-up periods since the effects may change their direction and strength over time (Conley at al. 2015). Therefore, to further address the nature of sustainability of intervention effects, in this review we aimed to systematically identify, appraise and summarize the existing data from randomized control trials (RCT) reporting on whether the effects of mental health promoting and mental ill health preventing interventions are sustained for at least 3 months of post-interventional follow-up. Further, we aimed to analyse how the direction and magnitude of the effects vary for different outcomes in relation to the lengths of follow-up and to assess whether effect sustainability varied by the types and major features of interventions, study-level determinants, and characteristics of participants.

MATERIALS AND METHODS

Eligibility criteria

The protocol was registered in PROSPERO, CRD42015029353 (Data, S1). The study followed the guidelines for conducting systematic review as suggested by the Cochrane handbook for systematic reviews of interventions (Higgins et al. 2008) and reported the study findings and procedure in relation to the sec statement (Moher et al. 2009) (Table, S5).
The PICO components (Population, Intervention, Comparator and Outcome) were developed after discussing eligibility criteria with stakeholders from the student health services: P= students in university settings; I= any types of mental health-promoting and mental ill health-preventing interventions; C= any types of active or inactive controls; O= (i) positive mental health, including well-being, coping, locus of control, resilience, self-esteem/self-compassion, stress management, academic achievement or academic performance, and (ii) mental ill health, including symptoms of anxiety, symptoms of depression, psychological distress, worry, fatigue, sleeping problems, and perceived stress. The study design was restricted to RCTs with at least 3 months of post-intervention follow-up. No language restrictions were initially applied. Studies focused on students with diagnosed psychiatric disorders and studies conducted in primary care settings were excluded.

Search strategy

In collaboration with librarians (CG, AW: see Acknowledgements), a sensitive search strategy was developed and adapted to the following databases: Medline (Ovid), PsycInfo (Ovid), Eric (Ovid) and Scopus. Gray literature was searched in Dissertations & Theses (ProQuest), Dart Europe, OpenGrey and Base Bielefeld. The searches were limited to studies published from January 1, 1995 to December 31, 2015. Key words and MESH terms are reported in the Supplemental Information (Data, S2 and Data, S3, respectively). Reference lists of the relevant reviews and studies selected for inclusion were manually scrutinized and the following journals were hand-searched from January 2012 to March 2016: College Student Journal, Journal of American College Health and Journal of College Counseling.
Study selection, data extraction and quality assessment

Screening was conducted independently by two authors (RW, KG) and two colleagues (AF, AM: see Acknowledgement). The eligibility of each article was initially evaluated by the title and abstract and, if found appropriate, followed by full-text examination. At this stage only English-language publications were assessed. This resulted in a loss of 20 publications in Chinese (k=15), Japanese (k=3), Korean (k=1) and Spanish (k=1). Gray literature was taken into consideration when accessible free of charge. Any disagreements were resolved through panel discussions. Studies selected for inclusion were examined for potential overlap in study populations, which was not found.

Data extracted from the articles included first author, country of origin, setting, funding, inclusion and exclusion criteria, characteristics of the intervention and comparison groups (age, gender, ethnicity), characteristics of the intervention (type, format, delivery level, length of session, duration), type of comparison, outcome definition and measurement scale, sample size, post-intervention length of follow-up, percent of withdrawals at each measurement point and study quality (described below). If a study reported multiple outcomes and/or if outcomes were assessed at multiple follow-up time points, quantitative data (means and standard deviations (SD)) were extracted separately for each outcome at each follow-up period. The same approach was utilized for multi-armed RTCs, from which separate extraction was performed for each intervention-comparison pair. When data were missing in the original reports we contacted authors for further clarification.

As suggested by Conley, et al. (Conley et al. 2015), original interventions were grouped into: (i) cognitive behaviour therapy (CBT)-related if focusing on identifying and
changing unhelpful cognitions, behaviors and emotional regulation; (ii) mind-body-related, i.e.,
interventions that facilitate the mind's capacity to affect bodily function and symptoms; and (iii)
psycho-educational-related if focusing on information, discussion and didactic communication
on, e.g. stress-reduction and coping. Categorization was based on the original definitions, if
provided, and otherwise by us. Level of delivery was considered as universal if intervention
targeted students without reported mental ill health symptoms and as selective if provided for
those with adverse mental health symptoms. Interventions were further divided into group or
individual format. Comparators were sub-divided into active controls (i.e., another type of
intervention) and inactive controls (waitlist controls, placebo-controls, “living as usual” and no
intervention). Study outcomes were classified in two major categories: mental ill health
outcomes consisting of anxiety symptoms, depressive symptoms, psychological distress, stress,
self-reported worry, passive coping, and deteriorated quality of sleep; and positive mental health
and academic performance outcomes including self-esteem, self-compassion, self-efficacy,
mental or subjective well-being, resilience, active coping, happiness, stress management and
academic performance.

The quality of selected trials was assessed independently by three authors (RW, KG, LL) and colleagues (AF, AM, SB: see Acknowledgements) using the Effective Public Health Practice Project Quality Assessment Tool (EPHPP), as recommended by the Cochrane Collaboration for public health reviews (Higgins et al. 2008). The EPHPP assesses selection bias, study-design, confounders, blinding, data collection method and withdrawals and dropouts to yield the study quality as either strong, moderate or weak. Discrepancies in quality assessment were resolved by discussions with one of the three reviewers not involved in the review process.
Statistical analysis

Because of the variety of instruments used for measuring outcomes, a standardized mean difference using Hedges’ $g$ was chosen as a common effect size (ES) for conducting quantitative synthesis. ES was calculated separately at each post-intervention follow-up time point as a difference in means between intervention and control group, divided by the pooled within-group SD and incorporating a correction factor for small sample sizes (Borenstein et al. 2009). One trial reported Hedges’ $g$ as the study ES (Braithwaite & Fincham 2009), while for other studies it was calculated from the available raw data. Throughout the recalculations we kept the original direction of scales indicating the improvement of outcome measures. Thus, for mental ill health outcomes ESs below zero pointed to superiority of the intervention group over the controls, while for positive mental health and academic performance outcomes, ESs above zero indicated that the results favored the intervention. For one study where follow-up means, but not SDs, were provided and the intervention effect was indicated as “non-significant” (Chiauzzi et al. 2008), we set ES to zero. To ease interpretation of the magnitude of Hedges’ $g$, we applied Cohen’s convention (Cohen 1992) and defined the ES as small (0.2), medium (0.5), and large (0.8).

Precautions were taken to overcome unit-of-analysis error and avoid using multiple assessment of the same construct (Higgins et al. 2008). If more than one ES was reported in a given study for the same outcome at a given follow-up point (e.g., for depression assessed by both the Hamilton Depression Rating Scale and Beck Depression Inventory), we averaged ESs to obtain the single outcome measure per intervention at each measurement point (Higgins 2006; Jones & Johnston 2000; Kanji et al. 2006; Peden et al. 2001; Seligman et al. 1999). ESs were also averaged within the trials with multiple interventions of similar nature, i.e. if interventions
belonged to the same category (Chiauzzi et al. 2008; Mak et al. 2015). A similar approach was applied to studies with multiple comparisons (Chiauzzi et al. 2008; Rohde et al. 2014; Yang et al. 2014). An exception was the study by Kanji N et al. (Kanji et al. 2006), where two control groups – an attention control and a time control – were included separately as considered to be different in approach and content and, thus, representing active and inactive comparisons, respectively.

Meta-analysis was conducted for all specific outcomes originally reported and for outcomes combined within mental ill health and positive mental health and academic performance categories. To analyze the combined outcomes, we applied a hierarchical approach for selecting outcomes from the studies reporting more than one from the same category. The hierarchy was based on descending order of outcome reporting, i.e. from the most often reported to the least often reported. For mental ill health outcomes the hierarchical selection was ordered as: depressive symptoms, anxiety symptoms, stress, psychological distress, self-reported worry, quality of sleep and passive coping. For positive mental health and academic performance outcomes the order was: self-esteem, academic performance, self-efficacy, self-compassion, mental or subjective well-being, resilience, stress management, active coping, and happiness.

Because of the initial assumptions of between-study heterogeneity, a random-effects model incorporating both within- and between-study variability was used for quantitative synthesis. To assess the sustainability of intervention effect over time and address the variety of the follow-up lengths reported in the original studies, we categorized the post-intervention follow-ups as 3-6 months, 7-12 months, and 13-18 months. Each of the included studies reported outcome measures for at least one of these categories and quantitative synthesis was conducted separately for each category. If a given study provided several outcome measures falling in the
same length category (e.g., for both 3 and 6 month follow-ups), the ES for the follow-up close to
the upper boundary (i.e., 6 months) was chosen (Kanji et al. 2006; Seligman et al. 2007; Vazquez
et al. 2012). Only one study (Seligman et al. 1999) assessed outcomes at follow-up periods
longer than 18 months and the measurements of those periods (i.e., 24 months, 30 months and 36
months) were not included in the meta-analysis. Finally, to obtain comparability between trials,
if a study provided results based on both imputed data (i.e., intention-to-treat analysis) and non-
imputed data (i.e., follow-up completers) (Reavley et al. 2014) or if both crude and adjusted ESs
were available (Chase et al. 2013) we favored the imputed and crude measures for the main
analysis leaving the latter (non-imputed and adjusted measures) for sensitivity analysis.

We evaluated statistical heterogeneity among the studies using $Q$ and $I^2$ statistics.
For $Q$, p-value <0.1 was considered as representative of statistically significant heterogeneity,
and $I^2$ values of 25%, 50% and 75% were indicating low, moderate and high heterogeneity,
respectively (Higgins et al. 2003).

The subgroup analyses were performed by stratifying the main analysis by a priori
identified moderators related to interventions (category of intervention, delivery level, type of
format, type of controls), study-level moderators (initial study size, study quality) and
moderators related to participant characteristics (gender, country). The analyses were performed
if at least two studies were included in each subgroup. Mixed-methodology was applied with
random-effect modeling used for within-group pooling, while between-group differences were
assessed with fixed-effect model. Leave-one-out influence analysis was conducted to assess the
potential impact of individual studies on the overall pooled ES by omitting one study at a time
(Tobias 1999). Following the approach suggested by Hart et al (Hart et al. 2012), sensitivity
analysis was conducted to assess whether the overall pooled ES differed if the lowest or the
highest original ES was selected from the studies with multiple outcome assessments or multiple interventions or comparisons. In meta-analyses with three or more studies included, we assessed publication bias by funnel plots, Egger’s regression asymmetry test, and the Begg-Mazumdar adjusted rank correlation test (Begg & Mazumdar 1994; Egger et al. 1997).

All statistical analyses were performed using STATA version 13.1 (StataCorp, College Station, TX). P-values < 0.05 were considered statistically significant. All statistical tests were two-sided.

RESULTS

After removing the duplicates, 6571 records were available for title and abstract screening. Among these, 6519 records were excluded as not meeting the PICO-criteria, leaving 52 articles for full-text examination. Further evaluation excluded another 26 studies: post-intervention follow-up less than 3 months (k=11), not enough data to calculate ES (k=6), not a RCT (k=5), population not relevant (k=3), and outcome not relevant (k=1). A selection process yielded a final number of 26 RCTs to be included in the meta-analysis (Fig. 1).

Figure 1 Flow diagram of the study selection process.

Characteristics of included studies
Table 1 summarizes the characteristics of studies eligible for inclusion. Among the 26 RCTs (Braithwaite & Fincham 2009; Chase et al. 2013; Cheng et al. 2015; Chiauzzi et al. 2008; Erogul et al. 2014; Fontana et al. 1999; Franklin & Franklin 2012; Gortner et al. 2006; Hamdan-Mansour et al. 2009; Higgins 2006; Jones & Johnston 2000; Kanji et al. 2006; Kattelmann et al. 2014; Kenardy et al. 2006; Li et al. 2015; Mak et al. 2015; Pachankis & Goldfried 2010; Peden et al. 2001; Reavley et al. 2014; Rohde et al. 2014; Seligman et al. 1999; Seligman et al. 2007; Shapiro et al. 2011; Vazquez et al. 2012; Yang et al. 2014; Zheng et al. 2015), CBT-related interventions were assessed in 11 studies, while mind-body-related and psycho-educational-related interventions were assessed in 10 and 5 studies, respectively. Universal and selective delivery levels were equally present (k=13 for both). Face-to-face group format was the most common (k=16). At least one mental ill health outcome was assessed in 24 studies, while at least one positive mental health outcome was appraised in 14 studies. Twenty-three trials reported at least one outcome measurement during 3-6 months post-intervention follow-up, with 8 and 5 trials reporting corresponding measurements during 7-12 months and 13-18 months follow-ups. None of the interventions had an organizational approach. More detailed characteristics of the selected studies are presented on-line (Table S1). The study quality assessed in all 26 RCTs varied between strong (k=4), moderate (k=12) and week (k=10). When subdivided by the outcome categories, 24 trials with at least one mental ill health outcome revealed their study quality as strong (k=4), moderate (k=11) and weak (k=9), and 14 trials with at least one positive mental health outcome and academic performance of strong (k=3), moderate (k=6) and weak (n=5) quality. Across all studies included in the analysis, selection bias was the most commonly assessed weakness component (n=21). (Table S2).
Table 1 Summary of descriptive characteristics of randomized controlled trials included in the systematic review and meta-analysis.

Effects and sustainability over time

Interventions preventing mental ill health

As presented in Table 2, for the combined mental ill health outcomes an aggregated ES for all preventive interventions yielded a superiority of interventions over the comparisons at 3-6 months and 7-12 months of post-intervention follow-up, although the effects were small. Pooled ES did not reach statistical significance for follow-up periods of 13-18 months. High-to-moderate heterogeneity was detected at all three follow-up periods with $I^2$ of 79.5%, 74.2%, and 58.2%, respectively. Publication bias was evident for studies with 3-6 months follow-up (Egger’s test p-value=0.013), though not for studies with longer follow-up periods (7-12 months: p-value=0.151; 13-18 months: p-value=0.141), (Figure, S1). Influence analysis revealed no indication that individual RCTs, if omitted, would significantly influence the observed overall ESs (data not shown). As previously noted, sensitivity analyses were performed for studies with multiple outcome measures reported for the same follow-up and for studies with multiple interventions or comparisons. Pooling together the highest ESs originally reported for these studies did not alter the results of the main analysis 3-6 months: ES= -0.28 (95% CI -0.44, -0.12); 7-12 months: ES= -0.35 (95% CI -0.60, -0.10); 13-18 months: ES= -0.23 (95% CI -0.54, 0.08)). Neither were the results influenced when the lowest originally reported ESs were used (3-6 months: ES= -0.24 (95% CI -0.39, -0.09); 7-12 months: ES= -0.23 (95% CI -0.43, -0.03); 13-18 months: ES= -0.06 (95% CI -0.18, 0.06)). Only one study reported the results using both
imputed and non-imputed data (Reavley et al. 2014), with the former included in the main meta-
analysis. Alternative inclusion of the latter did not change the overall ES (data not shown).

In sub-group analyses for the combined mental ill health outcomes, studies employing CBT-related interventions revealed significant pooled ESs for 3-6 month and 13-18 month follow-ups (Table 2 and Figure 2). Less consistent results were observed for mind-body-related interventions. No superiority of intervention group appeared among studies with psycho-educational interventions. Pooled ESs for universal preventive interventions yielded significant results for follow-up up to 7-12 months. Less consistency appeared in the aggregated results for selective interventions and interventions conducted face-to-face in groups. Trials with small sample size and trials comprising more than 60% females yielded significant effects for up to 7-12 months of follow-up. The small numbers of studies might explain the lack of consistency in the results of other sub-group comparisons. The high heterogeneity seen for studies with 3-6 months of follow-up might reflect differences in delivery level (p-value for Q between sub-groups=0.006) and study size (p<0.001), while for studies with follow-up periods of 7-12 months, heterogeneity could be explained by differences in type of comparison (p<0.001), study size (p=0.01), and country where the RCT was conducted (p=0.003). We were unable to detect between-group differences for trials with follow-up of 13-18 months because of the small number of studies in the sub-groups.

Assessment of the specific mental ill health outcomes, revealed a sustainable effect of all interventions combined lasting up to 13-18 months for symptoms of depression (ES= -0.30 (95% CI -0.51, -0.08)), (Table, S3). For symptoms of anxiety sustainability was observed up to 7-12 months (ES= -0.27 (95% CI; -0.54, -0.01)). Only one study assessed the effect of interventions targeting anxiety during 13-18 months of post-intervention follow-up and, hence,
we were unable to perform meta-analysis. For symptoms of stress, reductions lasted up to 3-6 months post intervention (ES=-0.30 (95%CI; -0.58, -0.03)). Other comparisons were either inconclusive or quantitative synthesis was not performed because of the small number of studies.

Table 2 Meta-analysis and sub-group analyses for hierarchically selected mental ill health outcomes, stratified by the length of post interventional follow-up periods.

Figure 2 The effects of mental ill health preventing interventions on hierarchically selected mental ill health outcomes stratified by the length of post interventional follow-up periods.

Interventions promoting mental health and academic performance – a paucity of outcomes

Table 3 presents overall ESs for the combined positive mental health and academic performance outcomes with rather limited data available, in particular, for the follow-up periods longer than 3-6 months. All interventions combined showed superiority over the controls during 3-6 months of follow-up with small, but significant pooled ES. For longer follow-up periods the results were inconclusive. High heterogeneity was detected when studies with 3-6 months follow-up were pooled ($I^2=86.5\%$). Because of the small number of studies, publication bias were only assessed for studies with 3-6 months of follow-up and were detected (Egger’s test p-value=0.03) (Figure, S5). Influence analysis indicated that four individual studies (Erogul et al. 2014; Hamdan-Mansour et al. 2009; Pachankis & Goldfried 2010; Peden et al. 2001), if omitted, would drop the significant overall ES for the studies with 3-6 months follow-up to borderline significance (data
not shown). Overall ESs at follow-ups of 7-12 and 13-18 months remain non-significant regardless of individual study influences (data not shown). Sensitivity analyses pooling the highest ESs originally reported for studies with multiple outcome assessment or multiple interventions or comparison groups showed no alteration to the overall results at 3-6 months follow-up (ES= 0.32 (95% CI 0.06, 0.58)), but made the overall ESs for 7-12 months follow-ups significant (ES= 0.53 (95% CI 0.20, 0.87)), as well as for 13-18 months follow-up (ES= 0.53 (95% CI 0.21, 0.86)). However, only 2 studies were assessed within each category of 7-12 and 13-18 months follow-ups. Use of the lowest originally reported ES did not affect the results of the main analysis (3-6 months: ES= 0.32 (95% CI 0.06, 0.58); 7-12 months: ES= 0.16 (95%CI -0.18, 0.50); 13-18 months: ES= 0.16 (95% CI -0.17, 0.49)). One study reported both crude and adjusted outcome assessment (Chase et al. 2013). Use of the adjusted ES for sensitivity analysis did not alter the results observed in the main analysis.

Sub-group analyses for the combined positive mental health and academic performance outcomes were performed only for studies with 3-6 months follow-up (Table 3 and Figure 3). Superiority of interventions over comparisons was shown for CBT-related interventions, selective delivery level, face-to-face group format, RCTs with inactive comparisons, studies with small sample size and trials conducted in US. Between-group difference was significant for delivery level (p<0.001), format type (p<0.001), study size (p<0.001) and gender mix (p=0.02). Sub-group analyses for studies with longer follow-up revealed either non-significant results or were impossible to conduct owing to the small number of trials in the sub-groups.

Because of lack of data on the specific positive mental health and academic performance outcomes, only studies on active coping, self-esteem and self-efficacy with 3-6
months follow-up were quantitatively assessed (Table, S4). Sustainability of the intervention
effect was observed for active coping (ES=0.75(95%CI; 0.19, 1.30)) with no significant effects
shown for other outcomes.

Table 3 Meta-analysis and sub-group analyses for hierarchically selected positive mental health
and academic performance outcomes stratified by the length of post interventional follow-up
periods.

Figure 3 The effects of mental health promoting interventions on hierarchically selected positive
mental health and academic performance outcomes stratified by the length of post interventional
follow-up periods.

DISCUSSION

Our systematic review and meta-analysis showed sustainability of the benefits of mental health
interventions targeting students in higher education, though in most of the analyses, the pooled
ESs yielded significant, but small overall effects. For the combined mental ill health outcomes,
the observed effects across all preventive interventions were sustained for up to 7-12 months
post-intervention. Sustainability of effects was most pronounced for interventions designed to
reduce the symptoms of depression, for which the superiority of intervention groups over the
comparisons remained significant for up to 13-18 months post-intervention. For the combined
positive mental health and academic performance outcomes, aggregated results across all
promotion interventions revealed slightly shorter, but still evident sustained effects, which
remained significant at post-intervention follow-up of 3-6 months.

To our knowledge, this is the first systematic review and meta-analysis focusing
primarily on the sustainability of the effects of mental health promoting and mental ill health
preventing interventions among students in higher education and analyzing different categories
of follow-up duration. A direct comparison to the existing literature was therefore difficult as
other reviews mostly assessed the effects measured at the completion of interventions. The
closest comparisons are three reviews by Conley et al. on universal and indicated mental health
prevention programs (Conley et al. 2015; Conley et al. 2017) and technology-delivered
preventive interventions (Conley et al. 2016). The reviews by Conley et al. assessed the effects
of interventions across all types of adjustment outcomes in university students at the longest
follow-up period reported, which varied from 2 to 52 weeks (Conley et al. 2015), 13 to 52 weeks
(Conley et al. 2016) and 4 to 157 weeks (Conley et al. 2017). The first review (Conley et al.
2015) showed the duration of follow-up to be negatively correlated with aggregated ES across
mental ill health and positive mental health outcomes combined as well as no effect for psycho-
educational interventions. A similar tendency for the effects of intervention to become non-
significant as the duration of follow-up increases was observed in our study, though the
sustainability of effects differed between ill-health and positive mental health outcomes. As in
Conley’s review (Conley et al. 2015), no effects of psycho-educational interventions on any
outcomes were evident in our data, regardless of the duration of follow-up. The second review
(Conley et al. 2016) reported a significant effect of universal interventions at any follow-up periods ranging between 13 to 52 weeks as well as a positive effect of selective interventions during the follow-up periods of 2-26 weeks. Similarly, in our study mental ill health outcomes were reduced by universal interventions for up to 7-12 months of follow-up and by selective interventions at follow-ups of up to 3-6 months, although our results on positive mental health and academic performance outcomes were less conclusive. Similar to the third review (Conley et al. 2017), our results indicated that the most sustainable effects were observed for interventions designed to reduce the symptoms of depression and symptoms of anxiety.

Although our literature search for intervention studies was not limited to psychological interventions, only this type was retrieved. The scarcity of organizational mental health promoting interventions was verified by a scoping review (Enns et al. 2016). However, an exception may be a recent systematic review on learning environment interventions for medical student well-being, suggesting changes to curriculum (Wasson et al. 2016). Their results support previous findings suggesting that to maximize the effectiveness of mental health promotion, all levels of delivery must contribute, i.e., not just individual and group levels, but also structural, and societal levels (Hamilton & Bhatti 1996). To further improve the sustainability of student mental health promotion, psychological interventions may be combined with a whole-setting approach, as endorsed by the WHO initiative Health Promoting Universities (HPU) (World Health Organization 1995).

Limitations
Systematic reviews on student mental health have indicated lack of follow-up data on outcome assessment as a major obstacle for determining the long-term effect of interventions (Conley et al. 2015; Conley et al. 2016; Davies et al. 2014; Farrer et al. 2013). Likewise, the scarcity of studies assessing the effects of interventions at post-interventional follow-ups of longer than 3 months along with a substantial variability in the lengths of follow-ups reported in the original studies should be considered as major limitations of our review. In particularly, the lack of original evidence affected our analysis of positive mental health outcomes as it restricted us to mainly aggregating the effects of interventions with 3-6 months of follow-up. Other limitations must also be considered. First, in most cases low numbers of studies in sub-groups prevented us from exploring the moderating effect of types of interventions, study-level determinants and participant characteristics during follow-up periods longer than 6 months making the results of sub-group analyses tentative. This also precluded us from conducting in-depth investigation of sources for heterogeneity, which was found to be mostly high. Second, our intention to analyse two dimensions of mental health, that resulted in combining the original outcomes into the “mental ill health” and “positive mental health and academic performance” outcome categories, with a hierarchical approach applied, could have boosted heterogeneity. In a subsequent analyses, we attempted to reduce heterogeneity by pooling together the studies with the same specific outcomes reported, though for several outcomes it was not possible due to data scarcity. Third, the evidence was insufficient to obtain any aggregated ESs for the specific outcomes, in particular, for self-reported worries, passive coping, academic performance, self-compassion, mental and subjective well-being, resilience and happiness rating. Fourth, a substantial variability exists in measurements instruments and, in several cases, the same outcome was measured by different scales. We tried to address this limitation by choosing Hedges’ $g$ as an ES
and by investigating how sensitive the aggregated results were to our initial approach of combining the original ESs in cases of multiple outcome measures or in multi-armed RCTs. The sensitivity analyses proved the robustness of our findings for mental ill health, though for positive mental health outcomes the use of the lowest ESs from the original studies altered the results for studies with 7-12 and 13-18 months of follow-up. Fifth, more than 30% of the original studies were assessed as being of weak quality. To address this issue, we conducted sub-group analyses stratifying the trials by study quality. For both categories of outcomes, these analyses revealed inconclusive results when trials with insufficient quality were pooled that should be accounted when interpreting our results. Furthermore, selection bias as the most commonly identified weakness. This bias, whether induced by the investigators or caused by self-selection may have resulted in either underestimation or overestimation of the original ES and therefore could affect the aggregated results. Finally, the results should be seen in the context of the presence of publication bias among the studies with 3-6 months of follow-up and of our inability to assess publication bias for positive mental health outcomes at follow-ups longer than 6 months, which may have resulted from our restriction to English-language publications at the final stage of selection.

CONCLUSIONS

Despite the limitations, the evidence suggests long-term effect sustainability for mental ill health preventive interventions, in particular, for interventions to reduce the symptoms of depression and symptoms of anxiety. Interventions designed to promote positive mental health offer promising, but shorter-lasting effects. As the research field of health promoting interventions for
students expands, future studies may improve our attempts to establish the effectiveness and sustainability of those interventions, e.g., ascertaining the effects for specific positive mental health outcomes. In addition, future research should also focus on mental health organizational interventions to investigate their potential for students in tertiary education.

Acknowledgements

We gratefully acknowledge Anders Wändahl and Carl Gornitzki, University Library, Karolinska Institutet for advice and support and for providing us with electronic searches; our colleagues, Annika Frykholm, Anna Månsdotter and Sven Bremberg who contributed to the screening and extraction of data-base searches; and stake-holders from Student Health Services for valuable input for practice.

References

Andrews. 2004. The relation of depression and anxiety to life-stress and achievement in students. The British journal of psychology 95:509. 10.1348/0007126042369802

Begg CB, and Mazumdar M. 1994. Operating characteristics of a rank correlation test for publication bias. Biometrics:1088-1101.

Bewick B, Koutsoportou G, Miles J, Slaa E, and Barkham M. 2010. Changes in undergraduate students' psychological well-being as they progress through university. Studies in Higher Education 35:633-645.

Borenstein M, Hedges LV, Higgins J, and Rothstein HR. 2009. Effect sizes based on means. In: Borenstein M, Hedges L, Higgins J, and Rothstein H, eds. Introduction to Meta-Analysis. Chichester: Wiley, 21-32.

Braithwaite SR, and Fincham FD. 2009. A randomized clinical trial of a computer based preventive intervention: replication and extension of ePREP. Journal of Family Psychology 23:32-38. http://dx.doi.org/10.1037/a0014061

Chase JA, Houmanfar R, Hayes SC, Ward TA, Vilardaga JP, and Follette V. 2013. Values are not just goals: Online ACT-based values training adds to goal setting in improving undergraduate college
532 student performance. *Journal of Contextual Behavioral Science* 2:79-84.
533 10.1016/j.jcbs.2013.08.002
534 Cheng M, Hasche L, Huang H, and Su XS. 2015. The effectiveness of a meaning-centered
535 psychoeducational group intervention for Chinese college students. *Social Behavior and
536 Personality* 43:741-756. [http://dx.doi.org/10.2224/sbp.2015.43.5.741](http://dx.doi.org/10.2224/sbp.2015.43.5.741)
537 Chiauzzi E, Brevard J, Thurn C, Decembrele S, and Lord S. 2008. "MyStudentBody-Stress: An online stress
management intervention for college students" Erratum. *Journal of Health Communication
538 13:827. [http://dx.doi.org/10.1080/10810730802619404](http://dx.doi.org/10.1080/10810730802619404)
539 Christensson A, Runeson B, Dickman PW, and Vaez M. 2010. Change in depressive symptoms over
540 higher education and professional establishment - a longitudinal investigation in a national
541 cohort of Swedish nursing students. *BMC Public Health* 10:343. 10.1186/1471-2458-10-343
542 Cohen J. 1992. A power primer. *Psychol Bull* 112:155-159.
543 Conley CS, Durlak JA, and Kirsch AC. 2015. A Meta-analysis of universal mental health prevention
544 programs for higher education students. *Prev Sci* 16:487-507. 10.1007/s11121-015-0543-1
545 Conley CS, Durlak JA, Shapiro JB, Kirsch AC, and Zahniser E. 2016. A Meta-Analysis of the Impact of
546 Universal and Indicated Preventive Technology-Delivered Interventions for Higher Education
547 Students. *Prev Sci* 17:659-678. 10.1007/s11121-016-0662-3
548 Conley CS, Shapiro JB, Kirsch AC, and Durlak JA. 2017. A meta-analysis of indicated mental health
549 prevention programs for at-risk higher education students. *J Couns Psychol* 64:121-140.
550 10.1037/cou0000190
551 Cvetkovski S, Reavley NJ, and Jorm AF. 2012. The prevalence and correlates of psychological distress in
552 Australian tertiary students compared to their community peers. *Aust N Z J Psychiatry* 46:457-
553 467. 10.1177/0004867411435290
554 Dahlin M, Joneborg N, and Runeson B. 2005. Stress and depression among medical students: a cross-
555 sectional study. *Med Educ* 39:594-604. 10.1111/j.1365-2929.2005.02176.x
556 Dahlin M, Nilsson C, Stotzer E, and Runeson B. 2011. Mental distress, alcohol use and help-seeking
557 among medical and business students: a cross-sectional comparative study. *BMC Med Educ
558 11:92. 10.1186/1472-6920-11-92
559 Davies EB, Morriss R, and Glazebrook C. 2014. Computer-delivered and web-based interventions to
560 improve depression, anxiety, and psychological well-being of university students: a systematic
561 review and meta-analysis. *J Med Internet Res* 16:e130. 10.2196/jmir.3142
562 Dyrbye LN, Thomas MR, and Shanafelt TD. 2006. Systematic review of depression, anxiety, and other
563 indicators of psychological distress among U.S. and Canadian medical students. *Acad Med
564 81:354-373.
565 Egger M, Davey Smith G, Schneider M, and Minder C. 1997. Bias in meta-analysis detected by a simple,
566 graphical test. *BMJ* 315:629 - 634.
567 Eisenberg D, Hunt J, and Speer N. 2013. Mental health in American colleges and universities: variation
568 across student subgroups and across campuses. *J Nerv Ment Dis* 201:60-67.
569 10.1097/NMD.0b013e31827ab077
570 Enns J, Holmqvist M, Wener P, Halas G, Rothney J, Schultz A, Goertzen L, and Katz A. 2016. Mapping
571 interventions that promote mental health in the general population: A scoping review of
572 reviews. *Preventive Medicine* 87:70-80.
573 Erogul M, Singer G, McIntyre T, and Stefanov DG. 2014. Abridged Mindfulness Intervention to Support
574 Wellness in First-Year Medical Students. *Teaching and Learning in Medicine* 26:350-356.
575 10.1080/10401334.2014.945025
576 Farrer L, Gulliver A, Chan JK, Batterham PJ, Reynolds J, Callear A, Tait R, Bennett K, and Griffiths KM.
577 2013. Technology-based interventions for mental health in tertiary students: systematic review.
578 *J Med Internet Res* 15:e101. 10.2196/jmir.2639
Fontana AM, Hyra D, Godfrey L, and Cermak L. 1999. Impact of a peer-led stress inoculation training intervention on state anxiety and heart rate in college students. *Journal of Applied Biobehavioral Research* 4:45-63. http://dx.doi.org/10.1111/j.1751-9861.1999.tb00054.x

Franklin J, and Franklin A. 2012. The long-term independently assessed benefits of coaching: A controlled 18-month follow-up study of two methods. *International Coaching Psychology Review* 7:33-38.

Garlow SJ, Rosenberg J, Moore JD, Haas AP, Koestner B, Hendin H, and Nemeroff CB. 2008. Depression, desperation, and suicidal ideation in college students: results from the American Foundation for Suicide Prevention College Screening Project at Emory University. *Depress Anxiety* 25:482-488. 10.1002/da.20321

Gortner E-M, Rude SS, and Pennebaker JW. 2006. Benefits of Expressive Writing in Lowering Rumination and Depressive Symptoms. *Behavior Therapy* 37:292-303. http://dx.doi.org/10.1016/j.beth.2006.01.004

Hamdan-Mansour AM, Puskar K, and Bandak AG. 2009. Effectiveness of cognitive-behavioral therapy on depressive symptomatology, stress and coping strategies among Jordanian university students. *Issues in Mental Health Nursing* 30:188-196. http://dx.doi.org/10.1080/01612840802694577

Hamilton N, and Bhatti T. 1996. Population health promotion: An integrated model of population health and health promotion. In: Canada PHAo, editor. Ottawa: Health Promotion Development Division.

Hart SL, Hoyt MA, Diefenbach M, Anderson DR, Kilbourn KM, Craft LL, Steel JL, Cuijpers P, Mohr DC, and Berendsen M. 2012. Meta-analysis of efficacy of interventions for elevated depressive symptoms in adults diagnosed with cancer. *Journal of the National Cancer Institute*.

Higgins DM. 2006. Preventing generalized anxiety disorder in an at-risk sample of college students: A brief cognitive-behavioral approach 3235646 Ph.D. The University of Maine.

Higgins J, Thompson S, Deeks J, and Altman D. 2003. Measuring inconsistency in meta-analyses. *BMJ* 327:557 - 560.

Higgins JPT, Green S, and Cochrane Collaboration. 2008. *Cochrane handbook for systematic reviews of interventions*. Chichester, England ; Hoboken, NJ: Wiley-Blackwell.

Hope V, and Henderson M. 2014. Medical student depression, anxiety and distress outside North America: a systematic review. *Med Educ* 48:963-979. 10.1111/medu.12512

Hunt J, and Eisenberg D. 2010. Mental health problems and help-seeking behavior among college students. *J Adolesc Health* 46:3-10. 10.1016/j.jadohealth.2009.08.008

Ibrahim AK, Kelly SJ, Adams CE, and Glazebrook C. 2013. A systematic review of studies of depression prevalence in university students. *J Psychiatr Res* 47:391-400. 10.1016/j.jpsychires.2012.11.015

Jane-Llopis E. 2007. Mental health promotion: concepts and strategies for reaching the population. *Health Promot J Austr* 18:191-197.

Jones MC, and Johnston DW. 2000. Evaluating the impact of a worksite stress management programme for distressed student nurses: A randomised controlled trial. *Psychology & Health* 15:689-706. 10.1080/08870440008405480

Kanji N, White A, and Ernst E. 2006. Autogenic training to reduce anxiety in nursing students: a randomized controlled trial. *J Adv Nurs* 53:729-735. 10.1111/j.1365-2648.2006.03779.x

Kattelmann KK, Bredbenner CB, White AA, Greene GW, Hoerr SL, Kidd T, Colby S, Horacek TM, Phillips BW, Koenings MM, Brown ON, Olfert MD, Shelnutt KP, and Morrell JS. 2014. The effects of Young Adults Eating and Active for Health (YEAH): a theory-based Web-delivered intervention. *Journal of Nutrition Education & Behavior* 46:S27-41. http://dx.doi.org/10.1016/j.jneb.2014.08.007
Kenardy J, McCafferty K, and Rosa V. 2006. Internet-delivered indicated prevention for anxiety disorders: Six-month follow-up. Clinical Psychologist 10:39-42.
http://dx.doi.org/10.1080/1328420050378746

Keyes CL, Eisenberg D, Perry GS, Dube SR, Kroenke K, and Ddingra SS. 2012. The relationship of level of positive mental health with current mental disorders in predicting suicidal behavior and academic impairment in college students. J Am Coll Health 60:126-133.
10.1080/07448481.2011.608393

Keyes CLM. 2007. Promoting and protecting mental health as flourishing: a complementary strategy for improving national mental health. The American Psychologist 62:95-108.

Kobau R, Seligman ME, Peterson C, Diener E, Zack MM, Chapman D, and Thompson W. 2011. Mental health promotion in public health: perspectives and strategies from positive psychology. Am J Public Health 101:e1-9. 10.2105/ajph.2010.300083

Lageborn CT, Ljung R, Vaez M, and Dahlin M. 2017. Ongoing university studies and the risk of suicide: a register-based nationwide cohort study of 5 million young and middle-aged individuals in Sweden, 1993–2011. BMJ Open 7:e014264.

Leahy CM, Peterson RF, Wilson IG, Newbury JW, Tonkin AL, and Turnbull D. 2010. Distress levels and self-reported treatment rates for medicine, law, psychology and mechanical engineering tertiary students: cross-sectional study. Aust N Z J Psychiatry 44:608-615. 10.3109/00048671003649052

Li M, Fang Q, Li J, Zheng X, Tao J, Yan X, Lin Q, Lan X, Chen B, Zheng G, and Chen L. 2015. The Effect of Chinese Traditional Exercise-Baduanjin on Physical and Psychological Well-Being of College Students: A Randomized Controlled Trial. PLoS ONE [Electronic Resource] 10:e0130544.
http://dx.doi.org/10.1371/journal.pone.0130544

Mak WW, Chan AT, Cheung EY, Lin CL, and Ngai KC. 2015. Enhancing Web-based mindfulness training for mental health promotion with the health action process approach: randomized controlled trial. J Med Internet Res 17:e8. http://dx.doi.org/10.2196/jmir.3746

Moher D, Liberati A, Tetzlaff J, Altman DG, and Group P. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6:e1000097.
10.1371/journal.pmed.1000097

Pachankis JE, and Goldfried MR. 2010. Expressive writing for gay-related stress: psychosocial benefits and mechanisms underlying improvement. Journal of Consulting & Clinical Psychology 78:98-110. http://dx.doi.org/10.1037/a0017580

Peden AR, Rayens MK, Hall LA, and Beebe LH. 2001. Preventing depression in high-risk college women: a report of an 18-month follow-up. J Am Coll Health 49:299-306. 10.1080/07448480109596316

Reavley NJ, McCann TV, Cvetkovski S, and Jorm AF. 2014. A multifaceted intervention to improve mental health literacy in students of a multicampus university: a cluster randomised trial. Social Psychiatry & Psychiatric Epidemiology 49:1655-1666. http://dx.doi.org/10.1007/s00127-014-0880-6

Regehr C, Glancy D, and Pitts A. 2013. Interventions to reduce stress in university students: A review and meta-analysis. Journal of Affective Disorders 148:1-11.
http://dx.doi.org/10.1016/j.jad.2012.11.026

Rohde P, Stice E, Shaw H, and Gau JM. 2014. Cognitive-behavioral group depression prevention compared to bibliotherapy and brochure control: nonsignificant effects in pilot effectiveness trial with college students. Behaviour Research & Therapy 55:48-53.
http://dx.doi.org/10.1016/j.brat.2014.02.003

Rotenstein LS, Ramos MA, Torre M, Segal JB, Peluso MJ, Guille C, Sen S, and Mata DA. 2016. Prevalence of Depression, Depressive Symptoms, and Suicidal Ideation Among Medical Students: A Systematic Review and Meta-Analysis. JAMA 316:2214-2236. 10.1001/jama.2016.17324
Rudman A, and Gustavsson JP. 2012. Burnout during nursing education predicts lower occupational preparedness and future clinical performance: a longitudinal study. *Int J Nurs Stud* 49:988-1001. 10.1016/j.ijnurstu.2012.03.010

Said D, Kypri K, and Bowman J. 2013. Risk factors for mental disorder among university students in Australia: findings from a web-based cross-sectional survey. *Soc Psychiatry Psychiatr Epidemiol* 48:935-944. 10.1007/s00127-012-0574-x

Seligman ME, Schulman P, DeRubeis RJ, and Hollon SD. 1999. The prevention of depression and anxiety. *Prevention & Treatment* 2:No Pagination Specified. http://dx.doi.org/10.1037/1522-3736.2.1.28a

Seligman ME, Schulman P, and Tryon AM. 2007. Group prevention of depression and anxiety symptoms. *Behaviour Research & Therapy* 45:1111-1126.

Shapiro SL, Brown KW, Thoresen C, and Plante TG. 2011. The moderation of Mindfulness-based stress reduction effects by trait mindfulness: results from a randomized controlled trial. *Journal of Clinical Psychology* 67:267-277. http://dx.doi.org/10.1002/jclp.20761

Shiralkar MT, Harris TB, Eddins-Folensbee FF, and Coverdale JH. 2013. A systematic review of stress-management programs for medical students. *Acad Psychiatry* 37:158-164. 10.1176/appi.ap.12010003

Steptoe A, Tsuda A, Tanaka Y, and Wardle J. 2007. Depressive symptoms, socio-economic background, sense of control, and cultural factors in university students from 23 countries. *Int J Behav Med* 14:97-107.

Tobias A. 1999. Assessing the influence of a single study in the meta-analysis estimate. *Stata Technical Bulletin* 8.

Tung YJ, Lo KKH, Ho RCM, and Tam WSW. 2018. Prevalence of depression among nursing students: A systematic review and meta-analysis. *Nurse Educ Today* 63:119-129. 10.1016/j.nedt.2018.01.009

Vaez M, and Laflamme L. 2008. Experienced stress, psychological symptoms, self-rated health and academic achievement: A longitudinal study of Swedish university students *Social Behavior and Personality* 36:183-195. 10.2224/sbp.2008.36.2.183

Wasson LT, Cusmano A, Meli L, Louh I, Falzon L, Hampsey M, Young G, Shaffer J, and Davidson KW. 2016. Association between learning environment interventions and medical student well-being: a systematic review. *JAMA* 316:2237-2252.

Vazquez FL, Torres A, Blanco V, Diaz O, Otero P, and Hermida E. 2012. Comparison of relaxation training with a cognitive-behavioural intervention for indicated prevention of depression in university students: a randomized controlled trial. *Journal of Psychiatric Research* 46:1456-1463. http://dx.doi.org/10.1016/j.jpsychires.2012.08.007

Weisz JR, Sandler IN, Durlak JA, and Anton BS. 2005. Promoting and protecting youth mental health through evidence-based prevention and treatment. *Am Psychol* 60:628-648. 10.1037/0003-066x.60.6.628

Westerhof GJ, and Keys CL. 2010. Mental Illness and Mental Health: The Two Continua Model Across the Lifespan. *J Adult Dev* 17:110-119. 10.1007/s10804-009-9082-y

Winzer R, Lindblad F, Sorjonen K, and Lindberg L. 2014. Positive versus negative mental health in emerging adulthood: a national cross-sectional survey. *BMC Public Health* 14:1238. 10.1186/1471-2458-14-1238

World Health Organization. 1995. Health Promoting Universities Available at http://www.who.int/healthy_settings/types/universities/en/ (accessed August, 8 2016).

World Health Organization (WHO). 2002. Prevention and Promotion in Mental Health. Geneva.
Yang W, Ding Z, Dai T, Peng F, and Zhang JX. 2014. Attention bias modification training in individuals with depressive symptoms: A randomized controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry*: No Pagination Specified. [http://dx.doi.org/10.1016/j.jbtep.2014.08.005](http://dx.doi.org/10.1016/j.jbtep.2014.08.005)

Zheng G, Lan X, Li M, Ling K, Lin H, Chen L, Tao J, Li J, Zheng X, Chen B, and Fang Q. 2015. Effectiveness of Tai Chi on Physical and Psychological Health of College Students: Results of a Randomized Controlled Trial. *PLoS ONE [Electronic Resource]* 10:e0132605. [http://dx.doi.org/10.1371/journal.pone.0132605](http://dx.doi.org/10.1371/journal.pone.0132605)
Table 1 (on next page)

Summary of study characteristics of randomized controlled trials included in the systematic review and meta-analysis.

* E.g. a different variant of the same intervention, a different intervention.

b E.g. no intervention, “living as usual”, a waiting list control

c Percentage does not add to 100 because studies could fall into multiple categories.

d Assessed by The Effective Public Health Practice Project Quality Assessment Tool (EPHPP).
Table 1 Summary of descriptive characteristics of randomized controlled trials included in systematic review and meta-analysis

|                                           | Number of comparisons (k) | %    |
|------------------------------------------|---------------------------|------|
| **Total number of studies**              | 26                        | 100  |
| **Participant characteristics**          |                           |      |
| Gender mix, Study population             |                           |      |
| Approx. even (40 – 60 females)           | 6                         | 23.0 |
| More than 60% females                    | 17                        | 65.0 |
| More than 60% males                      | 1                         | 4.0  |
| Not reported                              | 2                         | 8.0  |
| **Region (countries)**                   |                           |      |
| US                                        | 14                        | 54.0 |
| Australia                                 | 3                         | 11.5 |
| Europe (UK, Scotland, Spain)             | 3                         | 11.5 |
| East Asia and Pacific (China + Hong Kong)| 5                         | 19.0 |
| Middle East and North Africa (Jordan)    | 1                         | 4.0  |
| **Intervention characteristics**         |                           |      |
| Intervention classification               |                           |      |
| CBT-related                               | 11                        | 42.0 |
| Mind-body-related                        | 10                        | 39.0 |
| Psycho-educational-related               | 5                         | 19.0 |
| Enlarged with Material/Home-work/Training/Booster |                   |      |
| Yes                                       | 15                        | 58.0 |
| No/Unclear                                | 11                        | 42.0 |
| **Type of delivery**                     |                           |      |
| Universal                                 | 13                        | 50.0 |
| Selective                                 | 13                        | 50.0 |
| **Type of format**                       |                           |      |
| Internet-based individual                 | 5                         | 19.0 |
| Internet-based individual and in groups  | 2                         | 8.0  |
| Face-to-face individual                   | 2                         | 8.0  |
| Face-to-face in group                     | 16                        | 61.0 |
| Face-to-face in pairs                     | 1                         | 4.0  |
| **Length of intervention**               |                           |      |
| < 1 week                                  | 4                         | 15.0 |
| 1 - 4 weeks                               | 1                         | 4.0  |
| 5 - 7 weeks                               | 7                         | 27.0 |
| 8 weeks                                   | 8                         | 31.0 |
| 9-12 weeks                                | 5                         | 19.0 |
| 13-16 weeks                               | 0                         | 0.0  |
| > 16 weeks                                | 1                         | 4.0  |
| **Comparison condition**                 |                           |      |
| Active control a                          | 8                         | 31.0 |
| Inactive control b                        | 18                        | 69.0 |
| **Study characteristics**                |                           |      |
| Length of follow-up c                    |                           |      |
| 3 months                                  | 13                        | 50.0 |
| 4-6 months                                | 14                        | 54.0 |
| Duration         | Count | Percentage |
|------------------|-------|------------|
| 7-9 months       | 3     | 11.5       |
| 10-12 months     | 7     | 27.0       |
| 13-15 months     | 0     | 0.0        |
| >15 months       | 5     | 19.0       |

| Study size (participants) | Count | Percentage |
|---------------------------|-------|------------|
| n≤ 100                    | 16    | 62.0       |
| n> 100                    | 10    | 38.0       |

| Study quality | Count | Percentage |
|---------------|-------|------------|
| Weak          | 10    | 39.0       |
| Moderate      | 12    | 46.0       |
| Strong        | 4     | 15.0       |

---

a. E.g. a different variant of the same intervention, a different intervention.
b. E.g. no intervention, “living as usual”, a waiting list control.
c. Percentage does not add to 100 because studies could fall into multiple categories.
d. Assessed by The Effective Public Health Practice Project Quality Assessment Tool (EPHPP).
Table 2 (on next page)

Meta-analysis and sub-group analyses for hierarchically selected mental ill health outcomes, stratified by the length of post interventional follow-up periods.

k; number of studies; n/a; not applicable

(*) p<0.1 * p<0.05, ** p< 0.01, ***p<0.001

The format type “face-to-face in pair” was not utilized for mental ill health outcomes. [p]
Table 2 Meta-analysis and sub-group analyses for hierarchically selected mental ill health outcomes, stratified by the length of post interventional follow-up periods.

| Variables                         | Length of post intervention follow-up periods (months) |
|-----------------------------------|-------------------------------------------------------|
|                                   | 3-6                     | 7-12       | 13-18        |
| **All interventions** (k)         | 21                      | 9          | 3            |
| Hedges’ g (95% CI)               | -0.28 (-0.44, -0.12)   | -0.28 (-0.49, -0.08) | -0.17 (-0.39, 0.05) |
| Q / I²                           | 97.50*** / 79.5%        | 31.00*** / 74.2%  | 4.78(*) / 58.2%  |
| **Subgroup analyses**             |                         |            |              |
| **Type of interventions**         |                         |            |              |
| CBT-related (k)                  | 11                      | 4          | 2            |
| Hedges’ g (95% CI)               | -0.40 (-0.64, -0.16)   | -0.12 (-0.51, 0.16) | -0.30 (-0.51, -0.08) |
| Q / I²                           | 44.60*** / 77.6%        | 7.61(*) / 60.6%  | 0.28 / 0.0%  |
| Mind-body related (k)            | 9                       | 3          | 0            |
| Hedges’ g (95% CI)               | -0.20 (-0.44, 0.04)    | -0.43 (-0.66, -0.20) | ---          |
| Q / I²                           | 33.30*** / 76.0%        | 1.00 / 0.0%  | ---          |
| Psychoeducational (k)            | 1                       | 2          | 1            |
| Hedges’ g (95% CI)               | 0.09 (-0.05, 0.23)     | -0.64 (-1.83, 0.54) | -0.02 (-0.16, 0.16) |
| Q / I²                           | ---                     | 16.58*** / 94.0%  | ---          |
| **Delivery level**                |                         |            |              |
| Universal (k)                    | 8                       | 5          | 1            |
| Hedges’ g (95% CI)               | -0.23 (-0.46, -0.01)   | -0.46 (-0.83, -0.09) | -0.02 (-0.16, 0.12) |
| Q / I²                           | 34.23*** / 79.6%        | 23.04*** / 82.6%  | ---          |
| Selective (k)                    | 13                      | 4          | 2            |
| Hedges’ g (95% CI)               | -0.31 (-0.54, -0.08)   | -0.12 (-0.39, 0.16) | -0.30 (-0.51, -0.08) |
| Q / I²                           | 55.74*** / 78.5%        | 7.61(*) / 60.6%  | 0.28 / 0.0%  |
| **Format type**                  |                         |            |              |
| Face-to-face in group (k)        | 15                      | 6          | 2            |
| Hedges’ g (95% CI)               | -0.35 (-0.54, -0.16)   | -0.20 (-0.42, 0.03) | -0.30 (-0.51, -0.08) |
| Q / I²                           | 55.20*** / 74.6%        | 11.68* / 57.2%  | 0.28 / 0.0%  |
| Face-to-face individual (k)      | 2                       | 0          | 0            |
| Hedges’ g (95% CI)               | 0.06 (-0.30, 0.42)     | ---        | ---          |
| Q / I²                           | 1.91 / 47.5%            | ---        | ---          |
| Internet-based individual (k)    | 2                       | 2          | 0            |
| Hedges’ g (95% CI)               | -0.28 (-1.09, 0.53)    | -0.64 (-1.83, 0.54) | ---          |
| Q / I²                           | 6.26* / 84.0%           | 16.58*** / 94.0%  | ---          |
| Internet-based individual and in groups (k) | 2 | 1 | 1 |
| Hedges’ g (95% CI)               | -0.30 (-1.10, 0.50)    | -0.45 (-0.84, -0.07) | -0.02 (-0.16, 0.12) |
| Q / I²                           | 15.17*** / 93.4%        | ---        | ---          |
| **Type of comparison**           |                         |            |              |
| Active (k)                       | 2                       | 2          | 0            |
| Hedges’ g (95% CI)               | -0.34 (-1.21, 0.53)    | -0.88 (-1.59, -0.18) | ---          |
| Q / I²                           | 12.40*** / 91.9%        | 4.45* / 77.5%  | ---          |
| Inactive (k)                     | 19                      | 7          | 3            |
| Hedges’ g (95% CI)               | -0.28 (-0.44, -0.11)   | -0.14 (-0.28, 0.01) | -0.17 (-0.39, 0.05) |
| Q / I²                           | 83.38*** / 78.4%        | 10.17 / 41.0%  | 4.78 / 58.2%  |
| **Study quality**                |                         |            |              |
| Strong (k)                       | 2                       | 2          | 1            |
| Hedges’ g (95% CI)               | -0.22 (-0.63, 0.19)    | -0.17 (-0.34, 0.01) | -0.26 (-0.52, -0.01) |
| Q / I²                           | 2.50 / 60.0%            | 0.06 / 0.0%  | ---          |
|                      | Moderate (k) | Weak (k) | Not reported (k) |
|----------------------|-------------|----------|------------------|
| **Hedges’ g (95% CI)** | -0.26 (-0.44, -0.09) | -0.30 (-0.62, 0.02) | -0.25 (-0.94, 0.44) |
| **Q / I²**            | 21.12* / 57.4% | 21.82*** / 86.2% | 7.73** / 74.1% |

| Study size            | **100 participants or less (k)** | **More than 100 participants (k)** |
|-----------------------|----------------------------------|-----------------------------------|
| **Hedges’ g (95% CI)**| -0.45 (-0.68, -0.23)            | -0.01 (-0.10, 0.08)               |
| **Q / I²**            | 44.08***/70.5%                   | 7.07 / 15.1%                      |
| **Group difference Q (df) / p for Q** | 2.63 (2) / 0.27 | 46.3 (1) / <0.001 |

| Participants’ gender mix | **Approx. even (40% – 60% females) (k)** | **More than 60% females (k)** | **More than 60% males (k)** | **Not reported (k)** |
|--------------------------|------------------------------------------|-------------------------------|-----------------------------|---------------------|
| **Hedges’ g (95% CI)**   | -0.31 (-0.65, 0.04)                    | -0.32 (-0.52, -0.11)          | 0.30 (-0.17, 0.77)         | -0.25 (-0.94, 0.44) |
| **Q / I²**               | 5.40 / 63.0%                           | 80.27***/83.3%                | ---                         | 7.73** / 74.1%      |
| **Group difference Q (df) / p for Q** | 46.3 (1) / <0.001 | --- | --- | --- |

| Country                  | **US (k)** | **Other countries (k)** |
|--------------------------|------------|-------------------------|
| **Hedges’ g (95% CI)**   | -0.22 (-0.47, 0.03) | -0.33 (-0.55, -0.11) |
| **Q / I²**               | 35.73***/74.8% | 60.57***/83.5% |
| **Group difference Q (df) / p for Q** | 1.15 (1) / 0.28 | 1.35 (1) / 0.38 |

k; number of studies; n/a; not applicable

(* p<0.1 * p<0.05, ** p< 0.01, ***p<0.001

The format type “face-to-face in pair” was not utilized for mental ill health outcomes.

5
6
7
8
9
10
11
Table 3 (on next page)

Meta-analysis and sub-group analyses for hierarchically selected positive mental health and academic performance outcomes stratified by the length of post interventional follow-up periods.

k; number of studies; n/a; not applicable (*) p<0.1 * p<0.05, ** p< 0.01, ***p<0.001. The format type “Internet-based individual and in groups” was not utilized for positive mental health and academic performance outcomes.
| Variables                                                                 | Length of post intervention follow-up periods (months) |
|--------------------------------------------------------------------------|------------------------------------------------------|
|                                                                          | 3-6       | 7-12      | 13-18     |
| All interventions (k)                                                   | 11        | 2         | 2         |
| Hedges’ g (95% CI)                                                      | 0.32 (0.05, 0.59)       | 0.34 (-0.05, 0.73)       | 0.33 (-0.06, 0.72)       |
| Q / I                                                                   | 73.8*** / 86.5%       | 0.00 / 0.0%       | 0.04 / 0.0%       |
| **Subgroup analyses**                                                   |                                                      |
| **Type of interventions**                                               |                                                      |
| CBT-related (k)                                                         | 4         | 0         | 1         |
| Hedges’ g (95% CI)                                                      | 0.52 (0.06, 0.98)       | ---       | 0.29 (-0.29, 0.87)       |
| Q / I                                                                   | 24.5*** / 87.8%       | ---       | ---       |
| Mind-body related (k)                                                   | 6         | 1         | 0         |
| Hedges’ g (95% CI)                                                      | 0.23 (-0.16, 0.61)       | 0.35 (-0.35, 1.05)       | ---       |
| Q / I                                                                   | 41.68*** / 88.0%       | ---       | ---       |
| Psychoeducational (k)                                                   | 1         | 1         | 1         |
| Hedges’ g (95% CI)                                                      | 0.10 (-0.38, 0.58)       | 0.34 (-0.14, 0.82)       | 0.37 (-0.17, 0.91)       |
| Q / I                                                                   | ---       | ---       | ---       |

Group difference Q (df) / p for Q

| Delivery level | 3-6 | 7-12 | 13-18 |
|----------------|-----|------|-------|
| Universal (k)  | 5   | 2    | 1     |
| Hedges’ g (95% CI) | 0.05 (-0.23, 0.33)       | 0.34 (-0.05, 0.74)       | 0.37 (-0.17, 0.91)       |
| Q / I | 21.75** / 77.0%       | 0.00 / 0.0%       | ---       |
| Selective (k)  | 5   | 0    | 1     |
| Hedges’ g (95% CI) | 0.64 (0.18, 1.09)       | ---       | 0.29 (-0.29, 0.87)       |
| Q / I | 34.8*** / 88.5%       | ---       | ---       |

Group difference Q (df) / p for Q

| 17.26 (1) / <0.001 | n/a | n/a       |

| Format type       | 3-6 | 7-12 | 13-18 |
|-------------------|-----|------|-------|
| Face-to-face in groups (k) | 6   | 1    | 1     |
| Hedges’ g (95% CI) | 0.53 (0.16, 0.91)       | 0.35 (-0.35, 1.05)       | 0.29 (-0.29, 0.87)       |
| Q / I | 27.34*** / 81.7%       | ---       | ---       |
| Face-to-face individual (k) | 2   | 0    | 0     |
| Hedges’ g (95% CI) | 0.51 (-0.63, 1.66)       | ---       | ---       |
| Q / I | 16.17*** / 93.8%       | ---       | ---       |
| Face-to-face in pairs (k) | 0   | 1    | 1     |
| Hedges’ g (95% CI) | ---       | 0.34 (-0.14, 0.82)       | 0.37 (-0.17, 0.91)       |
| Q / I | ---       | ---       | ---       |
| Internet-based individual (k) | 3   | 0    | 0     |
| Hedges’ g (95% CI) | -0.14 (-0.36, 0.07)       | ---       | ---       |
| Q / I | 4.42 / 54.8%       | ---       | ---       |

Group difference Q (df) / p for Q

| 25.88 (2) / <0.001 | n/a | n/a       |

| Type of comparison | 3-6 | 7-12 | 13-18 |
|--------------------|-----|------|-------|
| Active (k)         | 1   | 0    | 0     |
| Hedges’ g (95% CI) | -0.11 (-0.50, 0.28)       | ---       | ---       |
| Q / I | ---       | ---       | ---       |
| Inactive (k)       | 10  | 2    | 2     |
| Hedges’ g (95% CI) | 0.36 (0.08, 0.65)       | 0.34 (-0.05, 0.74)       | 0.33 (-0.06, 0.73)       |
| Q / I | 72.53*** / 87.6%       | 0.00 / 0.0%       | 0.04 / 0.0%       |

Group difference Q (df) / p for Q

| n/a | n/a | n/a       |

| Study quality | 3-6 | 7-12 | 13-18 |
|---------------|-----|------|-------|
| Strong (k)    | 2   | 1    | 0     |
| Study size                        | Hedges’ g (95% CI) | Q / I² | Group difference Q (df) / p for Q |
|----------------------------------|--------------------|--------|----------------------------------|
| 100 participants or less (k)     | 0.49 (-0.53, 1.51) | 0.35 (-0.35, 1.05) | 0.53 (2) / 0.77 / n/a / n/a |
| Moderate (k)                     | 12.33*** / 91.9%   | ---    | ---                              |
| Hedges’ g (95% CI)               | 0.18 (-0.11, 0.47) | ---    | ---                              |
| Q / I²                           | 17.02** / 76.5%    | ---    | ---                              |
| Weak (k)                         | 4                  | 0      | 0                                |
| Hedges’ g (95% CI)               | 0.48 (-0.30, 1.25) | 0.34 (-0.14, 0.82) | 43.94*** / 93.2% / 0.04 / 0.0% |
| Q / I²                           | ---                |        |                                  |
| Group difference Q (df) / p for Q| 0.53 (2) / 0.77    | n/a    | n/a                              |
| Study size                        |                    |        |                                  |
| 100 participants or less (k)     | 0.84 (0.46, 1.23)  | 0.35 (-0.35, 1.05) | 0.53 (2) / 0.77 / n/a / n/a |
| Hedges’ g (95% CI)               | 11.97* / 66.6%     | ---    | ---                              |
| Q / I²                           |                     |        |                                  |
| More than 100 participants (k)   | 6                  | 1      | 1                                |
| Hedges’ g (95% CI)               | -0.04 (-0.20, 0.13)| 0.34 (-0.14, 0.82)| 0.37 (-0.017, 0.91) |
| Q / I²                           | 12.01* / 58.4%     | ---    | ---                              |
| Group difference Q (df) / p for Q| 49.84 (1) / <0.001 | n/a    | n/a                              |
| Participants’ gender mix         |                    |        |                                  |
| Approx. even (40% – 60% females) (k) | 0.49 (-0.06, 1.03) | 0.35 (-0.14, 0.82) | 0.37 (-0.017, 0.91)|
| Hedges’ g (95% CI)               | 22.88*** / 86.9%   | ---    | ---                              |
| Q / I²                           |                     |        |                                  |
| More than 60% females (k)        | 6                  | 1      | 1                                |
| Hedges’ g (95% CI)               | 0.10 (-0.19, 0.40) | 0.35 (-0.35, 1.05) | 0.29 (-0.29, 0.87) |
| Q / I²                           | 29.73*** / 83.2%   | ---    | ---                              |
| More than 60% males (k)          | 1                  | 0      | 0                                |
| Hedges’ g (95% CI)               | 1.12 (0.62, 1.63)  | ---    | ---                              |
| Q / I²                           | ---                | ---    | ---                              |
| Not reported (k)                 | 0                  | 0      | 0                                |
| Hedges’ g (95% CI)               | ---                | ---    | ---                              |
| Q / I²                           | ---                | ---    | ---                              |
| Group difference Q (df) / p for Q| 5.18 (1) / 0.02    | n/a    | n/a                              |
| Country                          |                    |        |                                  |
| US (k)                           | 6                  | 1      | 1                                |
| Hedges’ g (95% CI)               | 0.52 (0.11, 0.93)  | 0.35 (-0.35, 1.05) | 0.29 (-0.29, 0.87) |
| Q / I²                           | 38.79*** / 87.1%   | ---    | ---                              |
| Other countries (k)              | 5                  | 1      | 1                                |
| Hedges’ g (95% CI)               | 0.09 (-0.25, 0.44) | 0.34 (-0.14, 0.82) | 0.37 (-0.017, 0.91) |
| Q / I²                           | 24.24*** / 83.5%   | ---    | ---                              |
| Group difference Q (df) / p for Q| 10.78 (1) / 0.001  | n/a    | n/a                              |

k; number of studies; n/a; not applicable

(*) p<0.1 * p<0.05, ** p< 0.01, ***p<0.001

The format type “Internet-based individual and in groups” was not utilized for positive mental health and academic performance outcomes.
Flow diagram of the study selection process.

PRISMA flow diagram From: Moher D, Liberati A, Tetzlaff J, Altman DG, P referred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097
Records identified through database searching (n = 6004)
- Medline (Ovid) (n=961)
- PsycInfo (Ovid) (n=2462)
- Eric (Ovid) (n=420)
- Scopus (n=2161)

Records identified through other sources: Reference lists of included studies, relevant Journals (n=27)

Record after duplicates removed (n = 5016)

Search update for Medline, PsycInfo, Eric, Scopus (n=352)
- Gray literature searches
  - Dart Europe (n=310)
  - Open Grey (n=250)
  - Base Bielefeld (n=144)
  - Dissertations & Theses (n=499)

Records screened by title/abstract (n=6571)

Records excluded (n = 6519) as not fulfilling inclusion- and exclusion criteria

Full-text articles assessed for eligibility (n = 52)

Full-text articles excluded (n = 26)
- Reasons:
  - Post-intervention follow-up < 3 months (n=11)
  - Not enough data to calculate ES (n=6)
  - Not RCT (n=5)
  - Not relevant population (n=3)
  - Not relevant outcome (n=1)

Studies included in quantitative synthesis (meta-analysis) (n = 26)
Figure 2

The effects of mental ill health preventing interventions on hierarchically selected mental ill health outcomes stratified by the length of post interventional follow-up periods.

Lines represent standardized difference in means (Hedges’g) and 95% confidence intervals (CI); the size of the box represents the weight of each study.
Figure 3

The effects of mental health promoting interventions on hierarchically selected positive mental health and academic performance outcomes stratified by the length of post interventional follow-up periods.

The effect sizes of all interventions and combined subtotals for positive mental health and academic performance outcomes by the length of follow-up. Lines represent standardized difference in means (Hedges’g) and 95% confidence intervals (CI); the size of the box represents the weight of each study.
