DOI:
10.1111/jcpp.12809

The Journal of Child Psychology and Psychiatry
ACAMH The Association for Child and Adolescent Mental Health

Cognitive Bias Modification of Interpretations in Youth and its Effect on Anxiety: A Meta-analysis

Journal:	*Journal of Child Psychology and Psychiatry*
Manuscript ID	JCPP-RR-2016-00481.R3
Manuscript Type:	Research Review
Date Submitted by the Author:	n/a
Complete List of Authors:	Krebs, Georgina; King's College London, Institute of Psychiatry, Psychology and Neuroscience Pile, Victoria; King\'s College London, Psychology Department Grant, Sean; RAND Corporation, RAND Corporation Degli-Esposti, Michelle; University of Oxford, Experimental Psychology Montgomery, Paul; University of Oxford, Centre for Evidence-Based Intervention Lau, Jennifer; King\'s College London, Psychology Department;
Key Words:	Adolescence, Cognitive training, Anxiety, Information processing, School children
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
Supplementary material S13-S18.zip	

Running head: Meta-analysis of CBM-I in youth and its effect on anxiety

Cognitive Bias Modification of Interpretations in Youth and its Effect on Anxiety: A Meta-analysis

Georgina Krebs[a, b], Victoria Pile[b], Sean Grant[c,d], Michelle Degli-Esposti[e], Paul Montgomery[c] & Jennifer Y F Lau[b]*.

[a] OCD & Related Disorders Clinic for Young People, South London and Maudsley NHS Foundation Trust, Denmark Hill, London, SE5 8AZ, UK

[b] Institute of Psychiatry, Psychology & Neuroscience, King's College London, De Crespigny Park, London, SE5 8AF, UK.

[c] University of Oxford, Centre for Evidence-Based Intervention, 32 Wellington Square, Oxford, OX1 2ER, UK.

[d] RAND Corporation, 1776 Main Street, Santa Monica, CA 90401, USA.

[e] Department of Experimental Psychology, University of Oxford, 9 South Parks Road, Oxford, OX1 3UD, UK

* Correspondence to: Dr Jennifer Lau, King's College London, Institute of Psychiatry, Psychology & Neuroscience, Department of Psychology, Box PO77, Henry Wellcome Building, De Crespigny Park, London, SE5 8AF, UK; Telephone: +44 (0)20 7848 0678; Email: jennifer.lau@kcl.ac.uk

Abstract

Background: Emerging evidence suggests that Cognitive Bias Modification of Interpretations (CBM-I) is effective in altering interpretation biases and reducing anxiety in adults. Less is known about the impact of CBM-I in young people but some recent findings, including a meta-analysis of combined cognitive bias modification of interpretation and attention techniques, have cast doubt on its clinical utility. Given the current debate, this meta-analysis sought to establish the independent effects of CBM-I on interpretations biases and anxiety in youth.

Methods: Studies were identified through a systematic literature search of PsycINFO, Ovid Medline, PsycARTICLES, Web of Science, and Embase between January 1992 and March 2017. Eligible studies aimed to target interpretation biases; did not combine CBM-I with another intervention; included a control condition; randomly allocated participants to conditions; assessed interpretation bias and/or anxiety as an outcome; included individuals up to age 18; and did not present previously reported data. Reference lists of included articles were checked for further eligible studies, and authors were contacted for unpublished data.

Results: We identified 26 studies meeting eligibility criteria that included in the meta-analysis. CBM-I had moderate effects on negative and positive interpretations (g=-0.70 and g=-0.52 respectively) and a small but significant effect on anxiety assessed after training (g=-0.17) and after a stressor (g= -0.34). No significant moderators were identified.

Conclusions: In contrast to previous meta-analytic findings, our results indicate that CBM-I has potential but weak anxiolytic effects in youth. Our findings suggest that it may be premature to disregard the potential value of CBM-I research and further research in this field is warranted.

Key words: cognitive bias modification; interpretation bias training; anxiety; children; adolescents

Abbreviations: Cognitive bias modification (CBM); cognitive bias modification of interpretations (CBM-I).

Introduction

Anxiety disorders are the most common and functionally impairing psychiatric condition affecting children and adolescents (Merikangas et al., 2010; Wood, 2006). Left untreated, anxiety disorders typically persist into adulthood where they have been ranked as the sixth leading cause of disability globally (Baxter, Vos, Scott, Ferrari, & Whiteford, 2014). Approximately half of young people with anxiety disorders do not recover with current first-line psychological treatment, namely cognitive behavioural therapy (CBT), and about half of those who show an initial response subsequently relapse (Ginsburg, Becker, Keeton, & et al., 2014). Furthermore, accessing evidence-based treatments for anxiety is difficult (Kendall, Settipani, & Cummings, 2012). Hence, there is an urgent need to improve therapeutic outcomes and access for anxious youth by developing novel "standalone" or "adjunct" interventions. Cognitive Bias Modification of Interpretations (CBM-I) has been suggested as one such possibility.

CBM-I first emerged as a method for testing the causal link between interpretation biases and anxiety and mood primarily in analogue samples with varying levels of anxiety (Mathews & Mackintosh, 2000). The procedure involves teaching participants to generate benign or positive interpretations of ambiguous stimuli (usually ambiguous scenarios) through repeated training trials. Promising early results in the capacity of this training tool to reduce anxiety, albeit in analogue samples, has sparked interest in the clinical utility of Cognitive Bias Modification (CBM), including CBM-I specifically. This interest has partly arisen because the computerized format of these techniques means that they could represent a lower-cost and more easily disseminated intervention compared to existing, more costly therapies. Claims around the effectiveness of CBM-I have received mixed empirical support in adult analogue and clinical populations, which may in part reflect the significant heterogeneity between studies. The large number of studies in this area has enabled combining data using meta-analytic techniques (Cristea, Kok, & Cuijpers, 2015a; Hallion & Ruscio, 2011; Menne-Lothmann et al., 2014). Two meta-analyses have examined the impact of CBM-I in combination with attention bias modification (ABM) (Cristea et al., 2015a; Hallion & Ruscio, 2011), while the third assessed the effects of CBM-I in isolation (Menne-Lothmann et al., 2014). Of note, findings from these meta-analyses suggest that CBM-I may yield greater effects on biases and symptom reduction than ABM. Hallion and Ruscio found that CBM-I had a greater effect on the targeted biases than ABM, although there was no differential effect on affective symptoms (Hallion & Ruscio, 2011). Cristea and colleagues found that only CBM-I, not ABM, had a significant impact on anxiety and depression (Cristea et al., 2015a). Although Menne-Lothmann et al. (2014) did not compare CBM-I and ABM, they did find a small but significant effects of CBM-I alone on biases and on mood (when compared to negative training) (Menne-Lothmann et al., 2014). Interestingly, in the study by Hallion & Ruscio (2011), CBM was found to exert a greater effect on anxiety compared to depression (Hallion & Ruscio, 2011) although this was not reported by a later meta-analysis (Cristea et al., 2015a). Moreover, the effect of training on mood was only reliably detected when symptoms were assessed after exposure to a stressor, which is in keeping with diathesis-stress conceptualisations of cognitive biases (e.g. MacLeod, Campbell, Rutherford, & Wilson, 2004). Taken together, the results of these meta-analyses suggest that CBM-I may have modest effects on negative affect, particularly anxiety, in adult samples.

Less is known about the effect of CBM-I on childhood and adolescent anxiety, despite implications for early intervention. From a theoretical perspective, CBM-I training could yield stronger effects in youth, particularly in adolescents. Cognitive processing styles that are similar to the ones being targeted by CBM-I may develop during childhood and stabilise and mature across adolescence (e.g. Lau & Eley, 2008; Lau, Rijsdijk, & Eley, 2006;

Nolen-Hoeksema, Girgus, & Seligman, 1992), and may therefore be more amenable to modification during adolescence than adulthood. Whether CBM-I is as beneficial for children, compared to adults, is more difficult to predict. On the one hand, CBM-I involves a simple learning mechanism, which is not dissimilar to how children first acquire fears through associative learning (Benjet et al., 2010). Specifically, children may acquire fears by pairing neutral stimuli with aversive outcomes, for example by modelling their parents. CBM-I also pairs neutral, ambiguous stimuli with benign outcomes, and could be argued to reflect reinforcement learning and therefore be more appropriate for children. However, there is also some suggestion that cognitive styles are not yet mature in childhood. For example, cognitive styles moderate the effects of stress on affective symptoms in adolescence but not in childhood (Cole & Turner, 1993; Turner & Cole, 1994), and play less of an important role in predicting anxiety (Rudy, Davis, & Matthews, 2012) and mediating treatment effects (Kendall et al., 2016) compared to other cognitive factors in children. Thus, they may be less amenable to change.

There has been one meta-analysis assessing the effect of CBM-I together with ABM in children and adolescents across a range of mental health outcomes (Cristea, Mogoase, David, & Cuijpers, 2015b). This meta-analysis drew on 23 studies but only 13 evaluated CBM-I alone. While CBM-I and ABM training yielded significant effects on post-test cognitive biases relative to control training conditions, no significant effects were found on mental health outcomes including anxiety. Comparing effect sizes for CBM-I versus ABM found no difference between training type on mental health measures, but only CBM-I had a significant effect on targeted biases. Importantly, this study did not report the effects of CBM-I alone on anxiety specifically, only on combined mental health outcomes. The authors concluded that CBM is unlikely to have any clinical utility in non-adult populations.

Before the conclusions of Cristea and colleagues (Cristea et al, 2015b) regarding

CBM in youth are accepted, a number of factors should be considered. First, as mentioned above, this meta-analysis did not examine the effects of CBM-I alone on anxiety specifically. In light of the meta-analytic evidence in adult populations that: a) CBM-I may exert greater effects on affective symptoms than ABM (Cristea et al., 2015a), and b) CBM may have a greater impact on anxiety than other mood states (Hallion & Ruscio, 2011), it follows that CBM-I could still have a significant effect on anxiety in youth. Second, in their meta-analysis, Cristea and colleagues did not examine the impact of CBM on emotional reactivity (Cristea et al., 2015b). Diathesis-stress models conceptualise cognitive biases as being latent vulnerabilities that only exert an effect on affective state when the individual encounters a stressor (MacLeod et al., 2004). It therefore remains possible that CBM in youth could have a significant impact on anxiety after exposure to a stressor. Indeed, Hallion & Ruscio (2011) found that CBM only had reliable effects on anxiety in adults after exposure to a stressor. Third, the meta-analysis by Cristea et al. only included 13 studies that evaluated of CBM-alone and may therefore have lacked statistical power to detect small effects.

The current meta-analysis aimed to extend the previous meta-analysis by Cristea and colleagues (2015). Specifically, the primary aim was to determine the extent to which CBM-I alone modifies negative and positive interpretations in children and adolescents and to establish whether CBM-I is associated with immediate changes in anxiety. We focused solely on anxiety as an outcome because: a) there may be differential effects of CBM-I on anxiety versus depression, and from a theoretical and clinical perspective it is important to understand the impact of CBM-I on anxiety specifically; b) Hallion and Ruscio (2011) found evidence, albeit tentative, that CBM-I may have greater effects on anxiety than depression; c) a larger number of CBM-I studies in youth have examined anxiety as an outcome compared to depression, thereby affording us greater statistical power. In order to maxmise statistical power, we examined the impact of CBM-I on anxiety in unselected community samples with varying levels of anxiety (i.e. analogue samples), as well as participants with elevated levels of anxiety at baseline. This decision was made since most CBM-I studies in youth have been conducted in unselected community samples (Lau, 2013). Moreover, because anxious behaviours are likely to vary on a continuum from symptoms to disorder, with similar cognitive correlates characterising both, examining the modification of interpretations in analogue samples could inform their modification of clinically-significant anxiety in samples meeting diagnostic criteria. The second aim was to test the extent to which CBM-I is associated with changes to stress reactivity, as indexed by attenuations in anxiety following exposure to a challenging or stressful experience. Finally, we aimed to explore the influence of potential moderators on the effect of CBM-I. We chose *a priori* to examine four moderator variables that were hypothesised to be associated with the effect of CBM-I: 1) type of control condition (i.e. negative versus neutral versus no training); 2) number of training trials; 3) gender; and 4) age.

Method

This manuscript was developed in accordance with the Preferred Reported Items for Systematic Reviews and Meta-Analysis (PRISMA) Statement (Moher, Liberati, Tetzlaff, Altman, & The, 2009).

Literature Search and Selection criteria

Databases (PsycINFO, Ovid Medline, PsycARTICLES, Web of Science, and Embase) were originally searched in May 2014, with an updated search in November 2015, using multipurpose (.mp) searches with the following terms: "interpret* bias AND training"; "interpret* bias AND modif*"; "child"; "adolescent"; "young person"; "youth"; and "pediatric/paediatric" for publications between January 1992 and March 2017. This search was supplemented by reviewing reference lists and by correspondence with authors of included studies. Titles and abstracts were screened separately by two of the authors (JL and VP) to investigate whether the article focused on CBM-I training (eligibility criterion 1) in children and adolescents (eligibility criterion 6). Articles that appeared to meet these criteria were retained for full text review by both authors to assess whether they met the full set of eligibility criteria. All articles meeting eligibility criteria were included. Reference lists of included articles were checked for further eligible studies, and authors were contacted for unpublished data.

Eligibility criteria were as follows: 1) the study aimed to modify interpretation biases; 2) the CBM-I intervention was delivered in isolation and not combined with another intervention; 3) the study included a control group consisting of either negative or neutral (i.e. no contingency) CBM-I or no training; 4) participants were randomly allocated to condition; 5) interpretation biases and/or mood state were assessed after the intervention; 6) participants were children or adolescents up to 18 years of age; and 7) data had not been previously reported as part of another paper that was also deemed eligible for inclusion in the current meta-analysis. Only English language studies were eligible; studies were not restricted by the length of follow-up period following CBM-I or publication type (e.g. peer-reviewed publication, doctoral thesis, unpublished manuscript).

Coding of Data

Data on four outcome measures were collected: (1) positive and (2) negative interpretation bias post-training; (3) anxiety post-training; and (4) anxiety after a stressor administered post-training. Means and standard deviations of raw scores for the dependent variables, as well as sample size per intervention group, were extracted from each manuscript. Where means were not available, *t* values were extracted. If studies did not report the data necessary to calculate an effect size or transformed data were reported, the data was requested from authors. The majority of authors responded in these instances; only anxiety outcomes for one study had to be excluded due to missing data (Klein et al., 2015). To investigate sources of heterogeneity, additional variables were coded: age; gender; the number of training trials; and the nature of the control group. All manuscripts were coded by the first author (GK); 58% of codings (15 out of 26 manuscripts) were checked by the last author (JL).

Risk of Bias Within Studies

Risk of bias within individual studies was minimised by including randomisation to training condition as a selection criterion for eligibility, but three coders (including VP) also formally assessed all included studies using the Risk of Bias tool developed by the Cochrane Collaboration (Higgins et al., 2011). Where disagreement occurred between the coders, it was discussed with SG or JL and a conclusion across coders was reached. Each study was assessed on the following criteria: 1) adequacy of sequence generation; 2) adequacy of allocation concealment; 3) adequacy of blinding providers and participants; 4) blinding of outcome assessment; 5) adequacy of methods used to address incomplete outcome data; and 6) evidence of selective outcome reporting. The tool categorises individual studies as either 'low', 'unclear', or 'high' risk of bias. Coding was based on guidelines provided in Chapter 8 of the Cochrane Handbook for Systematic Reviews of Interventions (Higgin et al., 2011), but as interventions such as CBM-I may deviate somewhat from the typical psychosocial interventions that are discussed, several coding decisions are noteworthy. Criterion 3 (blinding of personnel and participants) was coded as high risk of bias if the personnel were not blind and it was considered likely that this would influence the outcome measurements or, as was the case for the majority of the studies, it seemed likely that blinding of participants was broken and that this would influence the outcome measurements. This decision was made because it seemed likely that participants could implicitly understand the aims of the training by virtue of the repetitive nature of CBM-I and the high level of similarity between the training tools and the primary assessment measure. For example, a previous study found that 94% of participants correctly guessed the purpose of the CBM intervention (Chan Reynolds & Lau, 2015). Criterion 4 (blinding of outcome assessors) was coded as low risk of bias where the outcome assessors were blinded or when outcome assessors were not present at the assessment (e.g. measures were completed by the participant at home alone). However, a rating of unclear risk of bias was made if the outcome assessor was present but the primary outcome was a computerised and/or self-administered measure. This decision was made because it was expected that outcomes would largely be assessed using self-administered measures, and is not clear whether lack of assessor blinding would influence the way in which participants completed measures. Criterion 5 (handling incomplete outcome data) was rated as low risk of bias if there was no missing outcome data (or less than 2%), when the missing outcome data was balanced across groups, when missing data was judged as unlikely to be related to the outcomes (e.g. technical issues), or intent-to-treat analyses were conducted.

Risk of bias Across Studies

Publication bias was informally assessed by visually inspecting the presence of asymmetry in funnel plots for each outcome variable generated in Review Manager (RevMan) version 5 (The Cochrane Collaboration, 2014). Asymmetry was formally evaluated using Egger tests of publication bias (Egger, 1997). If significant evidence for potential publication bias was identified, we planned to use the Duval-Tweedie trim and fill procedure (Duval & Tweedie, 2000) to estimate the overall effect size for each outcome after adjusting for publication bias. These analyses were conducted in Stata version 14 (StataCorp, 2015) using the *metabias* and *metatrim* commands, respectively.

Power calculation

Power calculations were conducted to determine the number of studies required to have sufficient statistical power to detect effects (Borenstein et al. 2009). We conducted two sets of power analyses corresponding to two different expected effect sizes. First, we assumed a small effect size of 0.3, in line with convention (Borenstein et al., 2009) and previous studies in the field (Cristea et al., 2015b), and a medium level of between-study heterogeneity (τ^2; Borenstein et al., 2009). Results indicated that 12 studies with a mean sample size of 50 (25 participants per condition) would have 80% power to detect an effect of *d*=0.3 at the 0.05 alpha level. Alternatively, 10 studies with a mean sample size of 60 (30 participants per arm) or 9 studies with a mean sample size of 66 (33 participants per arm) would be needed. Second, we repeated these analyses with a smaller effect size estimate of 0.2, in light of previous meta-analytic data showing an effect size of 0.17 of cognitive bias modification on anxiety (Cristea et al., 2015b). These calculations showed that 26 studies with a mean sample size of 50 (25 participants per condition) would have 80% power to detect an effect of *d*=0.3 at the 0.05 alpha level, assuming a medium level of between-study heterogeneity. Alternatively, 22 studies with a mean sample size of 60 (30 participants per arm) or 19 studies with a mean sample size of 70 (35 participants per arm) would be needed.

Meta-Analytic Procedures

Pooled effect sizes were calculated and forest plots produced using RevMan version 5. The standardized mean difference was calculated for each individual study, per outcome, in order to indicate the difference between the CBM-I and comparison group post-training. If a study included multiple measures for the same outcome, an average effect size was calculated. Hedge's *g* was then calculated across studies for each outcome: negative bias, positive bias, post-training anxiety, and post-stressor anxiety. A random effects model was used for all outcomes because heterogeneity was expected *a priori* across studies. Heterogeneity was assessed using the I^2 statistic (Higgins, Thompson, Deeks, & Altman, 2003).

A secondary aim was to examine potential moderators that were identified *a priori*. Subgroup analyses were conducted using RevMan for categorical moderator variables (e.g., nature of comparison condition), whereas meta-regressions were conducted using Stata version 14 for continuous moderator variables (e.g., number of training trials) (Harbord & Higgins, 2008). A previous review article highlighted that studies of CBM-I in youth tend to recruit either children or adolescents of a relatively narrow age range (Lau, 2013), and we therefore expected age to be bi-modally distributed across studies. Hence, rather than examine age in a meta-regression, we categorized studies as including children or adolescents (see Table 1), and conducted subgroup analyses to examine possible moderator effects. Studies were classified as "adolescent" if they exclusively included young people aged 12 years or older and "child" if they included young people under 12 years of age (actual age ranges per study are shown in Table 1). There were three exceptions: 1) Burnett-Heyes et al. (2017) included participants aged 11-15 years and this was classified as an "adolescent" study since the mean age was 14 years; 2) Lester et al. (2011a) included children aged 7-15 years and this was classified as a "child" study because the mean was 11 years; and 3) Stoddard et al. (2016) included participants aged 9-17 years and this was classified as an "adolescent" study as the mean age was 14 years.

Lastly, an exploratory sub-group analysis was conducted to explore test whether CBM-I had a differential effect on anxious versus non-anxious participants.

Results

Study Selection

Our search identified 577 citations. After the removal of 138 duplicates, the search produced a total of 439 articles. Titles and abstracts were obtained for these articles and screened using the inclusion criteria 1 and 7 as outlined above. This led to 41 articles being identified as potentially eligible for inclusion, and were reviewed against the full eligibility criteria. Following review of the full texts, a further 15 studies were excluded, leaving a final total of 26 studies for inclusion in the meta-analysis. Figure 1 summarises the number of articles identified at each stage of the retrieval process and the reasons for exclusions.

INSERT FIGURE 1

Study Characteristics

Study characteristics are shown in Table 1. This meta-analysis included data from 1786 participants aged 6-18 years from across 26 studies of whom 821 were males and participants received between 15 and 135 CBM-I training trials. All studies were published in peer-reviewed journals. Most studies were conducted with non-clinical unselected community samples. Fifteen studies included children and eleven included adolescents. Almost all studies used an ambiguous scenarios CBM-I paradigm, although the administration format varied between studies, with some studies presenting materials on computer screens and others presenting them on printed cards. Stimuli were generally developmentally-relevant, varying across studies according to the age of participants (e.g. scenarios involving animals for children, scenarios involving romantic relationships for adolescents). Across both child and adolescent studies most used single-session training although the number of training trials varied. The majority compared CBM-I to negative interpretation training. Most studies assessed the impact of CBM-I on negative interpretation bias, positive interpretation bias, and anxiety post-training, but only seven reported anxiety following exposure to a stressor. The majority of studies assessed state anxiety using a visual analogue scale.

INSERT TABLE 1

Risk of Bias Within Studies

All 26 studies were assessed for risk of bias. As many studies did not provide information required for assessing whether certain criteria were met, overall, the risk of bias was unclear (see Figure 2).

'Random sequence generation' and 'allocation concealment' were predominately rated as unclear risk of bias as there was usually insufficient information provided to permit a judgement. 'Blinding of participants and personnel' was rated as being likely to have high risk of bias for all studies. Personnel were rarely blinded to training condition due to the nature of the intervention and while most studies aimed to blind participants, a measure of contingency awareness was rarely included. Outcome assessors were not blind in the majority of studies but since all studies relied on computerised and/or self-completed outcomes measures, this was rated as unclear risk of bias. The majority of studies were rated as low risk of bias with respect to handling incomplete outcome data. Although no study reported intent-to-treat analyses, levels of attrition were very low which may reflect the fact that most studies comprised a single session. For 'selective reporting', all studies were rated as unclear risk of bias. No published protocols were referred to in the study manuscripts and none were identified in trial registration databases (clinicaltrials.gov; ISRCTN) , hence it was not possible to assess risk of bias for selective reporting.

INSERT FIGURE 2

Risk of Bias Across Studies

Visual inspection of funnel plots identified some asymmetry for negative interpretations and anxiety post-stressor, providing evidence for possible publication bias for these outcomes (see Figures S1-S4 in the supplementary material). Egger tests indicated significant asymmetry for negative interpretations (Egger test = -2.90, SE = 1.09, p=0.01), but not for positive interpretations (Egger test = -1.70, SE = 1.45, p=0.26), anxiety post-training (Egger test = -1.30, SE = 1.05, p=0.23) or anxiety post-stressor (Egger test = -4.51, SE = 2.34, p=0.11). Using the Duval-Tweedie trim and fill procedure, no evidence of publication bias was found for any of the four outcomes, and therefore adjusted effect sizes were not calculated.

Statistical power

We identified 25 studies with a mean of 33 participants reporting negative interpretations as an outcome, 18 studies with a mean of 32 participants reporting positive interpretations as an outcome, 17 studies with a mean of 36 participants reporting anxiety post-training as an outcome and 7 studies with a mean of 33 participants reporting anxiety post-stressor as an outcome. Thus, according to our power analysis we were adequately powered to detect effect sizes of 0.3 for negative interpretations, positive interpretations and anxiety post-training but not anxiety post-stressor. However, we did not have 80% power to detect smaller effect sizes of 0.2 or lower for positive interpretations, anxiety post-training and anxiety post-stressor.

Effect of CBM-I on Interpretation Biases and Anxiety

In total, 25 studies provided data on the effects of CBM-I versus comparison on a measure of post-training negative interpretation bias. The overall effect size was moderate to large (g = -0.70; 95% CI -0.80 to -0.53), indicating that the CBM-I group displayed significantly fewer negative interpretations than the control group. The level of heterogeneity was substantial (I^2 = 64%). The effects sizes per study are shown in the forest plot in Figure 3.

Eighteen studies included a measure of post-training positive interpretation bias. The overall effect size of CBM-I versus control on positive interpretations was moderate (g = -0.52; 95% CI -0.72 to -0.32), showing that the CBM-I group had significantly more positive interpretations than the control group. Again, the level of heterogeneity was substantial (I^2 = 60%). The effect sizes per study are shown in the forest plot in Figure S5.

Seventeen studies provided data on a measure of anxiety immediately post-training. The overall effect size was statistically significant, yet small (g = -0.17; 95% CI -0.31 to -0.02), indicating that the CBM-I group were significantly less anxious than the comparison group after completing the training. The level of heterogeneity was moderate (I^2 = 42%). The effects sizes per study are shown in the forest plot in Figure 4. Only seven studies measured anxiety after exposure to a stressor. The overall effect size of CBM-I versus comparison on post-stressor anxiety was small (g = -0.34; 95% CI -0.60 to -0.08), with a moderate level of heterogeneity (I^2 = 47%). This indicates that the CBM-I group were less anxious in response to a stressor than the control group. The effects sizes per study are shown in the forest plot in Figure S6.

INSERT FIGURES 3 & 4

Sensitivity analyses were conducted in relation to all outcomes, excluding outliers. Outliers were defined as studies with 95% confidence intervals that did not overlap with the 95% confidence interval for the pooled effect size. For negative interpretations, three studies were excluded (Chan, Reynold & Lau, 2015; Lau, Belli & Chopra, 2012; Muris et al., 2009) and the overall effect size remained largely unchanged (g = -0.68; 95% CI -0.84 to -0.52) with a lower level of heterogeneity (I^2 = 52%). For positive interpretations, two studies were excluded (Lau, Belli & Chopra, 2012; Vassilopoulos et al., 2009). Again, the overall effect size was largely unchanged (g = -0.50; 95% CI -0.66 to -0.34) with a lower level of heterogeneity (I^2 = 37%). For anxiety post-training, one study was excluded (Vassilopoulos et al., 2009), reducing the overall effect size (g = -0.13; 95% CI -0.26 to 0.00). The level of heterogeneity was lower (I^2 = 22%). For anxiety post-stressor there were no outliers.

Moderator Analyses

Moderator analyses were conducted in relation to outcomes on measures of negative interpretations, positive interpretations, and anxiety post-training, but not anxiety post-stressor because too few studies assessed this. Results are presented in Table S1-S2 and Figures S7-S18. A subgroup analysis was conducted to examine the impact of control condition (negative training versus neutral training versus no training) on outcomes. There was no statistically significant effect of control condition on negative interpretations, positive interpretation or anxiety post-training. However, the effect of CBM-I on positive interpretations was only statistically significant when compared to negative training (7 studies) or neutral training (7 studies), and was not significant when compared to no training (4 studies). Similarly, the influence of CBM-I on anxiety post-training was only significant when compared to negative training (7 studies), and was not significant when compared to neutral training (6 studies) or no training (4 studies). A second subgroup analysis revealed no overall statistically significant effect of age group (child versus adolescent) on any outcome. However, the effect of CBM-I on anxiety post-training was only significant among children (10 studies), not adolescents (7 studies). Finally, two separate meta-regressions revealed no significant effect of the number of training trials (range 15-720 trials) or gender (percentage of males; range 9.5-100%) on any outcome measure.

Exploratory analyses

A further subgroup analysis was conducted in order to test whether CBM-I had a differential effect on anxious versus non-anxious participants. There was no significant effect of baseline anxiety status on negative interpretations (χ^2 = 2.25, df =1, p = .13, I^2 = 56%), positive interpretations (χ^2 = .07, df =1, p = .80, I^2 = 0%), or anxiety post-training (χ^2 = .02, df =1, p = .90, I^2 = 0%). The analyses were not conducted for in relation to anxiety post-stressor because too few studies assessed this outcome.

Discussion

The current study represents the first meta-analysis of CBM-I alone in children and adolescents and included data from 1786 participants across 26 studies. We aimed to establish: the extent to which CBM-I reduces negative interpretations and increases positive interpretations in youth; the impact of CBM-I on anxiety; and the factors that might moderate the effects of CBM-I. This meta-analysis is an updated but also more focused investigation of interpretation bias modification compared to an earlier meta-analysis that examined fewer CBM-I studies, mainly in combination with ABM, and investigated effects on mental health outcomes more generally (Cristea et al., 2015b).

Our results indicate that CBM-I has a statistically significant moderate effect on both decreasing negative interpretations and boosting positive interpretations, in line with previous findings in youth (Cristea et al., 2015b). With respect to the impact of CBM-I on anxiety, we found a small but significant effect on self-reported anxiety immediately following training, consistent with adult findings (Hallion & Ruscio, 2011; Menne-Lothmann et al., 2014). The effect of CBM-I on anxiety was non-significantly larger when anxiety was measured after exposure to an anxiety-provoking situation. However, as only seven studies had included a measure of anxiety post-stressor, we may have been underpowered to detect differences (Hallion & Ruscio, 2011). Our finding that CBM-I has significant, albeit small, effects on anxiety is in contrast to the conclusions reached by the previous meta-analysis of CBM-I and ABM in young people (Cristea et al., 2015b). This discrepancy many reflect the fact that we focussed on CBM-I, which may be more effective than ABM (Cristea et al., 2015a), and included more pure CBM-I studies.

The current results are unlikely to be explained by publication bias since the Duval-Tweedie trim and fill procedure did not identify evidence of such bias for any outcome. Similarly, our results are unlikely to be driven by outliers. Sensitivity analyses indicated that the effects of CBM-I on negative and positive interpretations were largely unchanged and remained significant after excluding outliers. However, we found that the effect on anxiety post-training remained small after excluding one outlying study, and that the overall effect size was no longer statistically significant (p=.05).

Although we found no significant moderating effect of control condition on any outcome, the effects of CBM-I were only statistically significant across all outcomes when compared with negative training (i.e. they were not consistently significant when compared to neutral training or no training). These findings are in line with those of Menne-Lothman et al. (2014) and raise the question of whether CBM-I is genuinely improving interpretation biases and anxiety, or whether the between-group effects are mainly driven by the impact of negative training having the reverse effect on interpretations and anxiety. We did not find statistically significant moderating effects of age, gender or number of training trials on CBM-I with respect to any outcome. Our results are at odds with some previous findings in adults, but consistent with others. For example, Menne-Lothmann et al. (2014) found a significant effect of the number of training trials whereas Hallion and Ruscio (2011) and Cristea and colleagues (2015) did not. Of interest, although there was no statistically significant moderating effect of age, we found that the effect of CBM-I on anxiety was only significant among children (10 studies, n = 803 participants) and not adolescents (7 studies, n = 431 participants). While CBM-I may be more effective at reducing anxiety in younger populations, this finding may be confounded by methodological differences between child and adolescent studies. For example, in adolescent studies CBM-I tends to involve actively generating an interpretation by completing a single word fragment. In contrast, in child studies CBM-I typically involves selection of an interpretation from two alternatives that are presented, with selection of the positive being reinforced via feedback. It is possible that the latter is a more powerful training method as it encourages participants to select positive interpretations over competing negative interpretations.

While these findings are somewhat more promising than the earlier meta-analysis, it is important to note that compared to more established treatment packages, such as CBT, the effect size of CBM-I on anxiety is small and may not be clinically meaningful. This difference is perhaps not surprising as most CBM-I studies have been conducted with unselected analogue participants with less potential to reduce anxiety, and most have been single session experiments that primarily aimed to used CBM-I test mechanisms underlying anxiety rather than aiming to evaluate CBM-I as a clinical intervention. Nevertheless, our findings raise the question of why CBM-I does not have a more substantial impact on anxiety, given that it appears to successfully modify interpretation biases. There are a number of possible explanations for this observation. First, CBM-I effects on interpretation bias may be over-inflated. In most studies the outcome measure of interpretation is very similar to training materials, raising the possibility of demand effects – a possibility that was also discussed by Cristea and colleagues (2015) in their earlier meta-analysis of CBM-I procedures. Second, if interpretations biases only play a small role in anxiety, targeting interpretation biases in isolation may not be adequate. Instead, targeting multiple cognitive biases may enable stronger training effects. Indeed, established treatment protocols such as CBT involve multiple techniques of therapeutic change, of which challenging interpretations is just one aspect. Moreover, such techniques are often tailored to individual patient needs. CBT allows the incorporation of therapeutic techniques based on a shared understanding of what biases may be contributing and maintaining an individual's distress. In contrast, CBM-I is less flexible and less sensitive to such individual differences in its implementation. Individual differences in which cognitive biases are driving a disorder could also mean that CBM-I is more effective for some compared to other individuals. Based on these reasons, it is perhaps unrealistic to expect that modifying interpretations alone and in the current rigid format would yield equivalent or superior effects. At best, one might consider CBM to be a complementary adjunct treatment. Third, there may be a temporal lag between a change in interpretation bias and its impact on anxiety. Consistent with this hypothesis, changes in emotional information-processing have been found to precede and predict later changes in symptoms among anxious patients receiving CBT (Reinecke, Waldenmaier, Cooper, &

Harmer, 2013). It may be that time is needed for repeated practice of this new style of processing information, and for consolidation to occur.

The current findings should be considered in the context of a number of limitations. First, according to our power analyses we had less than 80% power to detect small effect sizes for anxiety post-training and anxiety post-stressor, and therefore these findings should be interpreted with caution. For example, although we found a statistically significant effect size of 0.17 for anxiety post-training, this finding could be spurious since being underpowered can give rise to type I as well as type II errors (Button et al., 2013). On the other hand, some of our non-significant findings (e.g. failure to find any significant moderator effects) could reflect type II errors. Second, we found a significant level of heterogeneity with respect to interpretation bias and anxiety data, raising the question of whether summary effect sizes are meaningful. Future research should seek to establish methodological and clinical characteristics that account for the substantial variation between studies. Third, overall studies were assessed as being at unclear risk of bias, principally due to a lack of documentation (Higgins et al., 2011). Thus, our finding that CBM-I has a significant effect in reducing anxiety could be a product of methodological biases within studies. Furthermore, while the discrepancy between the current findings and the results of the previous meta-analysis of CBM in youth (Cristea et al., 2015b) could indicate that CBM-I has a greater effect on anxiety than other mental health outcomes, it is also possible that anxiety studies have a higher risk of bias. Future studies should adopt more rigorous methodologies to reduce risk of bias and ensure that necessary information is included in publications to allow for risk of bias assessments. Improvements should include use of: random sequence generation to determine randomisation; assessment regarding blindness of participants (contingency awareness); blind outcome assessments; intent-to-treat analyses; and publication of study protocol. A fourth limitation is that the majority of studies in this meta-analysis examined CBM-I in unselected, analogue samples, which may limit the generalisability of our findings. Our exploratory analysis did not show a differential effect of baseline anxiety status on CBM-I outcomes but is likely have been underpowered. A final but nonetheless serious limitation was that few studies included psychometrically validated measures of anxiety (5 out of 17 post-training, 1 out of 7 post-stressor), with the majority using VASs. Furthermore, in a proportion of studies (5 out of 17 post-training, 2 out of 6 post-stressor), VASs for anxiety and low mood were combined to give a measure of negative affect. Although, there is some evidence that VASs have reasonable psychometric properties with respect to the measurement of state anxiety (Abend, Dan, Maoz, Raz & Bar-Hain, 2014), future studies should prioritise use of validated, anxiety-specific symptom measures.

A key question for future research is whether and how the effects of CBM-I can be enhanced in youth. Although, effect sizes for psychological therapies tend to be larger in initial, smaller studies and decrease with larger, more robust studies, there is nevertheless reason to believe that the effects for CBM-I could *potentially* increase in the future for two main reasons. First, the CBM-I procedures may be refined and improved. For example, 19 out of 27 studies included in this meta-analysis, involved single-session CBM-I, but adult studies tentatively suggest that multiple-session has significantly larger effects on symptoms than single-session CBM-I (Hallion & Ruscio, 2011), warranting further investigation in youth. Secondly, of the 26 studies included in the current meta-analysis, 20 were conducted with non-anxious individuals. Only three were conducted among clinically anxious participants (Fu, Du, Au & Lau, 2013; Klein et al., 2015; Orchard et al., 2017) and four with participants scoring above average on an anxiety measure (Fu, Du, Au & Lau, 2015; Vassilopoulos et al., 2009; Vassilopoulos, Blackwell et al., 2014; White et al., 2016). Although our exploratory analyses did not reveal significantly greater effects of CBM-I on anxious individuals compared to unselected samples, this may reflect the small number of studies included and it remains plausible that greater effects of CBM-I on anxiety will be obtained among individuals with clinical levels of anxiety symptoms where there is more potential for change. Further research is needed to investigate the effects of CBM-I in young people with anxiety disorders. In addition, future research should to look at the longer-term impact of CBM-I on anxiety in order to: a) test the hypothesis that changes in anxiety manifest after a lag; b) to establish durability of effects which is important in informing the possible clinical utility of CBM-I; and c) to test the hypothesis that CBM-I may modify reactivity to stress. In summary, a key priority is to conduct systematic, large-scale studies with clinical samples, longer-term follow-ups, and more robust and valid measures of interpretation bias and anxiety both immediately after training and in response to a psychological challenge. Only once these have been conducted can CBM-I effects be fully assessed.

Notwithstanding the limitations outlined above, this study represents the first systematic evaluation of the impact of CBM-I in young people. Our results suggest that even where the majority of studies include unselected, analogue samples, CBM-I is effective at modifying interpretation biases, at least within the domain targeted during training. We found preliminary evidence that CBM-I may have a small but significant effect in reducing anxiety in young people, and the effect sizes were of a similar magnitude to those found in adults (Cristea et al., 2015a; Hallion & Ruscio, 2011; Menne-Lothmann et al., 2014). Although the effects of CBM-I on anxiety are small, it is crucial to keep in mind that this field of research is still at an early stage, particularly in child and adolescent populations. More research is therefore warranted to establish the extent to which CBM-I has potential value as a method for advancing theoretical understanding and its clinical utility.

References

(Studies included in the meta-analysis are marked with an asterix)

Abend, R., Dan, O., Maoz, K., Raz, S., & Bar-Haim, Y. (2014). Reliability, validity and sensitivity of a computerized visual analog scale measuring state anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, *45*(4), 447-453.

Baxter, A., Vos, T., Scott, K., Ferrari, A., & Whiteford, H. (2014). The global burden of anxiety disorders in 2010. *Psychological Medicine*, *44(*11), 2363-2374.

*Belli S.R. & Lau, J.Y.F. (2014). Cognitive Bias Modification Training in Adolescents: Persistence of Training Effects. *Cognitive Therapy and Research*, *38*(6), 640-651. doi: 10.1007/s10608-014-9627-7

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. Hoboken, GB: Wiley.

*Burnett Heyes, S., Pictet, A., Mitchell, H., Raeder, S. M., Lau, J. Y. F., Holmes, E. A., & Blackwell, S. E. (2016). Mental Imagery-Based Training to Modify Mood and Cognitive Bias in Adolescents: Effects of Valence and Perspective. *Cognitive Therapy and Research*, 1-16.

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365-376.

*Chan, S.W.Y., Lau, J.Y., Reynolds, S.A. (2015). Is Cognitive Bias Modification training truly beneficial for adolescents? *Journal of Child Psychology and Psychiatry, 56*(11), 1239-48.

Cole, D. A., & Turner, J. E., Jr. (1993). Models of cognitive mediation and moderation in child depression. *J Abnormal Psychology, 102*(2), 271-281.

Collaboration, T. C. (2014). Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre.

Cristea, I. A., Kok, R. N., & Cuijpers, P. (2015a). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *The British Journal of Psychiatry, 206*(1), 7-16. doi: 10.1192/bjp.bp.114.146761

Cristea, I. A., Mogoase, C., David, D., & Cuijpers, P. (2015b). Practitioner Review: Cognitive bias modification for mental health problems in children and adolescents: a meta-analysis. *Journal of Child Psychology and Psychiatry, 56*(7), 723-734. doi: 10.1111/jcpp.12383

*De Winter, S., Bosmans, G., & Salemink, E. (2017). Exploring the causal effect of interpretation bias on attachment expectations. *Child Development*, *88*(1), 131-140.

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455-463.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. . *British Medical Journal, 315*(7109), 629-634.

*Fu, X., Du, Y., Au, S., & Lau JY (in press). Single-session Cognitive Bias Modification of Interpretations Training in High-anxious Adolescents. *Journal of Cognitive Psychotherapy.*

*Fu, X., Du, Y., Au, S. & Lau, J.Y.F. (2013). Reducing negative interpretations in adolescents with anxiety disorders: A preliminary study investigating the effects of a single session of cognitive bias modification training. *Developmental Cognitive Neuroscience* 4, 29-37. doi: http://dx.doi.org/10.1016/j.dcn.2012.11.003

Ginsburg, G. S., Becker, E. M., Keeton, C. P., & et al. (2014). Naturalistic follow-up of youths treated for pediatric anxiety disorders. *JAMA Psychiatry, 71*(3), 310-318. doi: 10.1001/jamapsychiatry.2013.4186

Hallion, L. S., & Ruscio, A. M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin, 137*(6), 940.

Harbord, R. M., & Higgins, J. P. (2008). Meta-regression in Stata. *Meta, 8*(4), 493-519.

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., . . . Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal, 343*, d5928. doi: 10.1136/bmj.d5928

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557-560. doi: 10.1136/bmj.327.7414.557

James, A., Soler, A., & Weatherall, R. (2005). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database Systematic Review, 19*(4).

Kendall, P. C., Cummings, C. M., Villabo, M. A., Narayanan, M. K., Treadwell, K., Birmaher, B., . . . Albano, A. M. (2016). Mediators of change in the Child/Adolescent Anxiety Multimodal Treatment Study. *Journal of Consulting and Clinical Psychology, 84*(1), 1-14. doi: 10.1037/a0039773

Kendall, P. C., Settipani, C. A., & Cummings, C. M. (2012). No Need to Worry: The Promising Future of Child Anxiety Research. *Journal of Clinical Child & Adolescent Psychology, 41*(1), 103-115. doi: 10.1080/15374416.2012.632352

Koster, E. H., & Bernstein, A. (2015). Introduction to the special issue on Cognitive bias modification: Taking a step back to move forward? *J Behavior Therapy and Experimental Psychiatry, 49*(Pt A), 1-4. doi: 10.1016/j.jbtep.2015.05.006

*Klein, A. M., Rapee, R. M., Hudson, J. L., Schniering, C. A., Wuthrich, V. M., Kangas, M., Lyneham, H. J., Souren, P. M. & Rinck, M. (2015). Interpretation modification training reduces social anxiety in clinically anxious children. *Behaviour Research and Therapy*, *75*, 78-84.

Lau, J. Y. (2013). Cognitive bias modification of interpretations: A viable treatment for child and adolescent anxiety? *Behaviour Research and Therapy, 51*(10), 614-622.

*Lau, J. Y., Belli, S. R., & Chopra, R. B. (2013). Cognitive bias modification training in adolescents reduces anxiety to a psychological challenge. *Clinical Child Psychology and Psychiatry, 18*(3), 322-333. doi: 10.1177/1359104512455183

Lau, J. Y., & Eley, T. C. (2008). Attributional style as a risk marker of genetic effects for adolescent depressive symptoms. *Journal of Abnormal Psychology, 117*(4), 849-859. doi: 10.1037/a0013943

*Lau, J. Y., Molyneaux, E., Telman, M. D., & Belli, S. (2011). The plasticity of adolescent cognitions: data from a novel cognitive bias modification training task. *Child Psychiatry & Human Development, 42*(6), 679-693. doi: 10.1007/s10578-011-0244-3

Lau, J. Y., Rijsdijk, F., & Eley, T. C. (2006). I think, therefore I am: a twin study of attributional style in adolescents. *Journal of Child Psychology and Psychiatry, 47*(7), 696-703. doi: 10.1111/j.1469-7610.2005.01532.x

*Lester, K.J., Field, A.P. & Muris, P. (2011a). Experimental modification of interpretation bias regarding social and animal fear in children. *Journal of Anxiety Disorders* 25, 697-705. doi: http://dx.doi.org/10.1016/j.janxdis.2011.03.006

*Lester, K.J., Field, A.P. & Muris, P. (2011b). Experimental Modification of Interpretation Bias about Animal Fear in Young Children: Effects on Cognition, Avoidance Behavior, Anxiety Vulnerability, and Physiological Responding. *Journal of Clinical Child & Adolescent Psychology*, 40(6), 864-877. doi: 10.1080/15374416.2011.618449

*Lothmann, C., Holmes, E. A., Chan, S. W. Y., & Lau, J. Y. F. (2011). Cognitive bias modification training in adolescents: effects on interpretation biases and mood. *Journal of Child Psychology and Psychiatry, 52*(1), 24-32. doi: 10.1111/j.1469-7610.2010.02286.x

*Muris, P., Huijding, J., Mayer, B. & Hameetman, M. (2008). A Space Odyssey: Experimental Manipulation of Threat Perception and Anxiety-Related Interpretation Bias in Children. *Child Psychiatry and Human Development*, 39(4), 469-480. doi: 10.1007/s10578-008-0103-z

*Muris, P., Huijding, J., Mayer, B., Remmerswaal, D. & Vreden, S. (2009). Ground control to Major Tom: Experimental manipulation of anxiety-related interpretation bias by means of the "space odyssey" paradigm and effects on avoidance tendencies in children. *Journal of Anxiety Disorders*, 23(3), 333-340. doi: http://dx.doi.org/10.1016/j.janxdis.2009.01.004

MacLeod, C., Campbell, E., Rutherford, L., & Wilson, E. (2004). The causal status of anxiety-linked attentional and interpretive bias. In J. Yiend (Ed.), *Cognition, emotion and psychopathology: Theoretical, empirical and clinical directions* (pp. 172–189). New York, NY: Cambridge University Press.

Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology, 109*(4), 602-615.

Menne-Lothmann, C., Viechtbauer, W., Höhn, P., Kasanova, Z., Haller, S. P., Drukker, M., . . . Lau, J. Y. F. (2014). How to Boost Positive Interpretations? A Meta-Analysis of the Effectiveness of Cognitive Bias Modification for Interpretation. *PLoS ONE, 9*(6), e100925. doi: 10.1371/journal.pone.0100925

Merikangas, K. R., He, J. P., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., . . . Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication--Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry, 49*(10), 980-989. doi: 10.1016/j.jaac.2010.05.017

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The, P. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine, 6*(7), e1000097. doi: 10.1371/journal.pmed.1000097

Nolen-Hoeksema, S., Girgus, J. S., & Seligman, M. E. (1992). Predictors and consequences of childhood depressive symptoms: a 5-year longitudinal study. *Journal of Abnormal Psychology, 101*(3), 405-422.

*Orchard, F., Apetroaia, A., Clarke, K., & Creswell, C. (2017). Cognitive bias modification of interpretation in children with social anxiety disorder. *Journal of Anxiety Disorders*, *45*, 1-8.

Reinecke, A., Waldenmaier, L., Cooper, M. J., & Harmer, C. J. (2013). Changes in Automatic Threat Processing Precede and Predict Clinical Changes with Exposure-Based Cognitive-Behavior Therapy for Panic Disorder. *Biological Psychiatry, 73*(11), 1064-1070. doi: http://dx.doi.org/10.1016/j.biopsych.2013.02.005

Rudy, B. M., Davis, T. E., 3rd, & Matthews, R. A. (2012). The relationship among self-efficacy, negative self-referent cognitions, and social anxiety in children: a multiple mediator model. *Behavior Therapy, 43*(3), 619-628. doi: 10.1016/j.beth.2011.11.003

*Salemink, E., & Wiers, R. W. (2011). Modifying Threat-related Interpretive Bias in Adolescents. *Journal of Abnormal Child Psychology, 39*(7), 967-976. doi: 10.1007/s10802-011-9523-5

StataCorp. 2015. *Stata Statistical Software: Release 14*. College Station, TX: StataCorp LP.

*Stoddard, J., Sharif-Askary, B., Harkins, E.A., Frank, H.R., Brotman, M.A., Penton-Voak, I.S., Maoz, K., Bar-Haim, Y., Munafo, M., Pine, D.S. & Leibenluft, E. (2016). An open pilot study of training hostile interpretation bias to treat disruptive mood dysregulation disorder. *Journal of Child and Adolescent Psychopharmacology*, *26*(1), 49-57.

*Telman, M. D., Holmes, E. A., & Lau, J. Y. F. (2013). Modifying Adolescent Interpretation Biases Through Cognitive Training: Effects on Negative Affect and Stress Appraisals. *Child Psychiatry & Human Development, 44*(5), 602-611. doi: 10.1007/s10578-013-0386-6

Turner, J. E., Jr., & Cole, D. A. (1994). Developmental differences in cognitive diatheses for child depression. *Journal of Abnormal Child Psychology, 22*(1), 15-32.

*Vassilopoulos, S. P., & Brouzos, A. (2016). Cognitive Bias Modification of Interpretations in Children: Processing Information About Ambiguous Social Events in a Duo. *Journal of Child and Family Studies*, *25*(1), 299-307.

*Vassilopoulos, S., & Moberly, N. (2013). Cognitive Bias Modification in Pre-adolescent Children: Inducing an Interpretation Bias Affects Self-imagery. *Cognitive Therapy and Research, 37*(3), 547-556. doi: 10.1007/s10608-012-9481-4

*Vassilopoulos, S. P., Banerjee, R., & Prantzalou, C. (2009). Experimental modification of interpretation bias in socially anxious children: Changes in interpretation, anticipated interpersonal anxiety, and social anxiety symptoms. *Behaviour Research and Therapy, 47*(12), 1085-1089. doi: http://dx.doi.org/10.1016/j.brat.2009.07.018

*Vassilopoulos, S. P., Brouzos, A., & Andreou, E. (2014). A Multi-Session Attribution Modification Program for Children with Aggressive Behaviour: Changes in Attributions, Emotional Reaction Estimates, and Self-Reported Aggression. *Behavioural and Cognitive Psychotherapy, FirstView*, 1-11. doi: doi:10.1017/S1352465814000149

*Vassilopoulos, S.P., Blackwell, S.E., Misailidi, P., Kyritsi, A. & Ayfanti, M. (2014). The differential effects of written and spoken presentation for the modification of interpretation and judgmental bias in children. *Behavioural and Cognitive Psychotherapy*, 42(5), 535-554. doi: 10.1017/s1352465813000301

*Vassilopoulos, S.P., Moberly, N.J. & Lau, J.Y.F. (2015). Cognitive bias modification training in children reduces anxiety during anticipatory processing of social evaluation. *International Journal of Cognitive Therapy*, 8(4).

*Vassilopoulos, S.P., Moberly, N.J. & Zisimatou, G. (2013). Experimentally Modifying Interpretations for Positive and Negative Social Scenarios in Children: A Preliminary Investigation. *Behavioural and Cognitive Psychotherapy*, 41(1), 103-116. doi:10.1017/S13524658120005

*White, L. K., Suway, J. G., Pine, D. S., Field, A. P., Lester, K. J., Muris, P., Bar-Haim, Yar. & Fox, N. A. (2016). The cognitive and emotional effects of cognitive bias modification in interpretations in behaviorally inhibited youth. *Journal of Experimental Psychopathology*, *7* (3). 499-510.

Wood, J. J. (2006). Effect of anxiety reduction on children's school performance and social adjustment. *Developmental Psychology, 42*(2), 345-349.

Table 1: Study characteristics

Study	Age range	% male	Mental health status	Training paradigm	No. of training sessions	Total no. of training trials	Control condition	Outcome measures			
								Negative interpretation bias	**Positive interpretation bias**	**Anxiety post-training**	**Anxiety post-stressor**
Belli & Lau (2014)	Adolescents (12-18 yrs)	20.3	Healthy	Ambiguous situations (social)	Single	50	Neutral training	Recognition test	Recognition test	VAS	N/A
Burnett Hayes et al (2017)	Adolescents (11-16 yrs)	100	Healthy	Mental imagery	Two	20	Neutral training	Recognition test; Scrambled sentences tasks; Pleasantness ratings of pictures	Recognition test;	VAS	N/A
Chan, Reynolds & Lau (2015)	Adolescents (16-18 yrs)	9.5	Healthy	Ambiguous situations	Two	80	Neutral training	Recognition test	Recognition test	N/A	STAI-S
De Winter et al (2017)	Children (8-12 yrs)	44.9	Healthy	Ambiguous situations (attachment-related)	Single	42	Neutral training	Recognition test	Recognition test	N/A	N/A
Fu, Du, Au & Lau (2013)	Adolescents (12-17 yrs)	46.4	Social phobia or GAD	Ambiguous situations	Single	50	Neutral training	Recognition test	Recognition test	VAS	N/A
Fu, Du, Au & Lau (2015)	Adolescents (12-18 yrs)	51.0	Selected for high anxiety	Ambiguous situations	Single	50	Neutral training	Recognition test	Recognition test	VAS	N/A

Klein et al (2015)	Children (7-12 yrs)	51.8	Anxiety disorder	Ambiguous situations	14	140	Neutral training	Ambiguous vignettes	N/A	N/A	N/A
Lau, Belli & Chopra (2013)	Adolescents (12-18 yrs)	50.0	Healthy	Ambiguous situations	Single	50	Negative training	Recognition test	Recognition test	VAS	VAS
Lau et al (2011)	Adolescents (13-18 yrs)	36.0	Healthy	Ambiguous situations	Single	50	Negative training	Recognition test	Recognition test	VAS	N/A
Lester et al (2011a)	Children (7-15 yrs)	56.7	Healthy	Ambiguous situations (animals)	Single	30	Negative training	Ambiguous vignettes	N/A	VAS	VAS
Lester et al (2011b)	Children (6-11 yrs)	40.8	Healthy	Ambiguous situations (animals)	Single	30	Negative training	Ambiguous vignettes	N/A	VAS	VAS
Lothmann et al (2011)	Adolescents (13-17 yrs)	46.3	Healthy	Ambiguous situations	Single	50	Negative training	Recognition test	Recognition test	N/A	N/A
Muris et al (2008)	Children (8-12 yrs)	48.6	Healthy	Ambiguous situations (space)	Single	30	Negative training	Ambiguous vignettes	N/A	N/A	N/A
Muris et al (2009)	Children (9-13 yrs)	53.3	Healthy	Ambiguous situations (space)	Single	30	Negative training	Ambiguous vignettes	N/A	N/A	N/A
Orchard et al (2017)	Children (7-12 yrs)	42.9	Social anxiety disorder	Ambiguous situations (social)	Three	45	No training	Ambiguous vignettes	Ambiguous vignettes	SCAS-SP (child and parent versions)	N/A

Salemink & Wiers (2011)	Adolescents (14-16 yrs)	46.5	Healthy	Ambiguous situations (social)	Single	40	Neutral training	Recognition test	Recognition test	STAI-C	N/A
Stoddard et al (2016): Experiment 2	Adolescents (9-17 years)	26.3	Healthy	Ambiguous facial expressions	Four	720	Neutral training	Ambiguous faces	N/A	N/A	N/A
Telman et al (2013)	Adolescents (15-18 yrs)	21.7	Healthy	Ambiguous situations	Single	50	Negative training	Recognition test	Recognition test	N/A	N/A
Vassilopoulos & Brouzos (2017)	Children (10-11 yrs)	52.6	Healthy	Ambiguous situations (social; administered to pairs of peers)	Single	20	No training	Ambiguous vignettes	Ambiguous vignettes	SASC-R	N/A
Vassilopoulos et al (2009)	Children (10-11 yrs)	18.7	Selected for high social anxiety	Ambiguous situations (social)	Three	45	No training	Ambiguous vignettes	Ambiguous vignettes	SASC-R	VAS
Vassilopoulos, Blackwell et al (2014)	Children (10-12 yrs)	50.0	Selected for high social anxiety	Ambiguous situations (social)	Single	15	Negative training	Ambiguous vignettes	Ambiguous vignettes	VAS	VAS
Vassilopoulos, Brouzos & Andreau (2014)	Children (10-12 yrs)	88.2	Selected for aggressive behaviour	Ambiguous situations (social)	Three	45	No training	Ambiguous vignettes	Ambiguous vignettes	N/A	N/A
Vassilopoulos & Moberly (2013)	Children (10-12 yrs)	42.6	Healthy	Ambiguous situations (social)	Single	20	Negative training	N/A	N/A	VAS	N/A
Vassilopoulos,	Children	42.7	Healthy	Ambiguous	Single	30	Negative	Ambiguous	Ambiguous	VAS	N/A

Moberly & Lau (2015)	(10-12 yrs)			situations (social)			training	vignettes	vignettes		
Vassilopoulos, Moberly & Zisimatou (2013)	Children (10-13 yrs)	39.9	Healthy	Ambiguous situations (social)	Three	48	No training	NSECQ	PSEDQ	SASC-R	N/A
White et al (2016)	Children (9-12 yrs)	64.4	Selected for high BI	Ambiguous situations	Single	50	Neutral training	Ambiguous vignettes	N/A	VAS	VAS

Notes: 'Healthy' indicates that study included an unselected sample of young people; VAS = visual analogue scale; STAI-C = State-Trait Anxiety Inventory for Children; SCAS = Spence Children's Anxiety Scale; SCAS-SP = Social Phobia subscale of Spence Children's Anxiety Scale ; STAI-S = State-Trait Anxiety Inventory-State version; SASC-R = Social Anxiety Scale for Children-Revised; NSECQ = Negative Social Events Catastrophization Questionnaire; PSEDQ = Positive Social Events Discounting Questionnaire; BI = Behavioural inhibition

Figure 1: PRISMA diagram of selection of studies

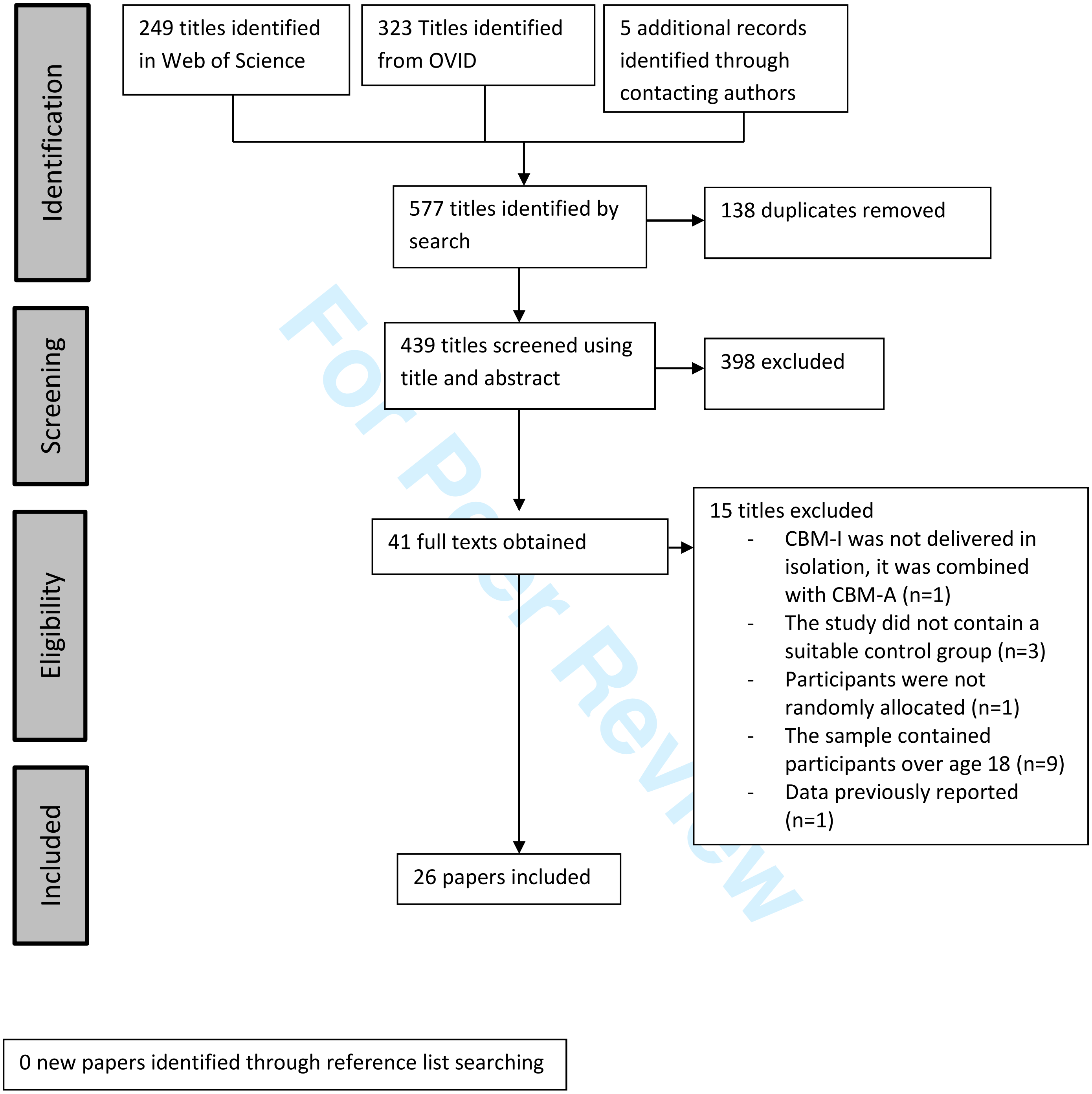

Figure 2: Summary of risk of bias across studies per criterion

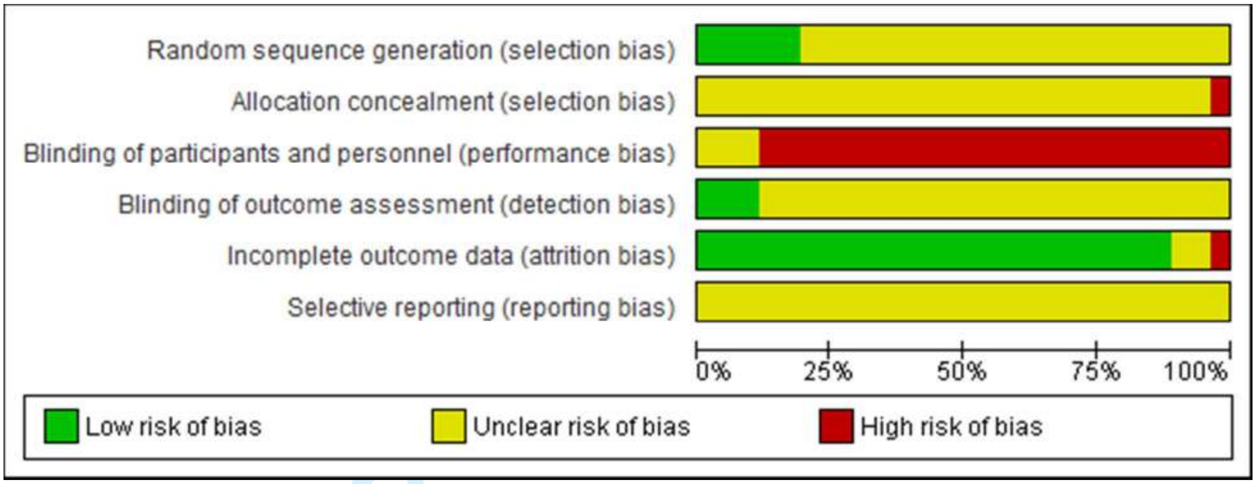

Figure 3: Forest plot of effect size of CBM-I versus control on negative interpretation bias

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Belli & Lau (2014)	-0.6775	0.30003	24	23	3.8%	-0.68 [-1.27, -0.09]
Burnett Heyes et al (2017)	-0.3953	0.269	28	29	4.2%	-0.40 [-0.92, 0.13]
Chan, Reynolds & Lau (2015)	0.1488	0.24844	33	32	4.4%	0.15 [-0.34, 0.64]
De Winter et al (2017)	-0.485	0.289923	25	24	3.9%	-0.48 [-1.05, 0.08]
Fu, Du, Au & Lau (2013)	-1.0739	0.41804	16	11	2.7%	-1.07 [-1.89, -0.25]
Fu, Du, Au & Lau (2015)	-0.8873	0.24533	37	36	4.4%	-0.89 [-1.37, -0.41]
Klein et al (2015)	-0.2035	0.22023	40	43	4.7%	-0.20 [-0.64, 0.23]
Lau, Belli & Chopra (2013)	-2.0586	0.3911	20	20	2.9%	-2.06 [-2.83, -1.29]
Lau, Molyneaux et al (2011)	-1.0349	0.35543	17	19	3.2%	-1.03 [-1.73, -0.34]
Lester et al (2011a)	-0.6085	0.20158	51	52	5.0%	-0.61 [-1.00, -0.21]
Lester et al (2011b)	-1.1295	0.26311	34	33	4.2%	-1.13 [-1.65, -0.61]
Lothmann et al (2011)	-1.1182	0.26482	32	34	4.2%	-1.12 [-1.64, -0.60]
Muris et al (2008)	-0.8196	0.24898	36	34	4.4%	-0.82 [-1.31, -0.33]
Muris et al (2009)	-0.0076	0.18281	63	57	5.2%	-0.01 [-0.37, 0.35]
Orchard et al (2017)	-0.4463	0.273061	28	27	4.1%	-0.45 [-0.98, 0.09]
Salemink & Wiers (2011)	-0.8785	0.17783	73	66	5.3%	-0.88 [-1.23, -0.53]
Stoddard et al (2016): Experiment 2	-1.6179	0.533648	8	11	2.0%	-1.62 [-2.66, -0.57]
Telman et al (2013)	-1.038	0.31411	23	23	3.7%	-1.04 [-1.65, -0.42]
Vassilopoulos & Brouzos (2016)	-0.6843	0.334235	20	18	3.4%	-0.68 [-1.34, -0.03]
Vassilopoulos et al (2009)	-0.7939	0.31686	22	21	3.6%	-0.79 [-1.41, -0.17]
Vassilopoulos, Blackwell et al (2014)	-0.2865	0.20903	53	41	4.9%	-0.29 [-0.70, 0.12]
Vassilopoulos, Brouzos & Andreau (2014)	-1.067	0.54439	16	18	1.9%	-1.07 [-2.13, -0.00]
Vassilopoulos, Moberly & Lau (2015)	-1.1073	0.22918	39	50	4.6%	-1.11 [-1.56, -0.66]
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4648	0.16385	77	76	5.4%	-0.46 [-0.79, -0.14]
White et al (2016)	-0.4396	0.301786	23	22	3.8%	-0.44 [-1.03, 0.15]
Total (95% CI)			**838**	**820**	**100.0%**	**-0.70 [-0.88, -0.53]**

Heterogeneity: $Tau^2 = 0.12$; $Chi^2 = 67.57$, $df = 24$ ($P < 0.00001$); $I^2 = 64\%$

Test for overall effect: $Z = 7.84$ ($P < 0.00001$)

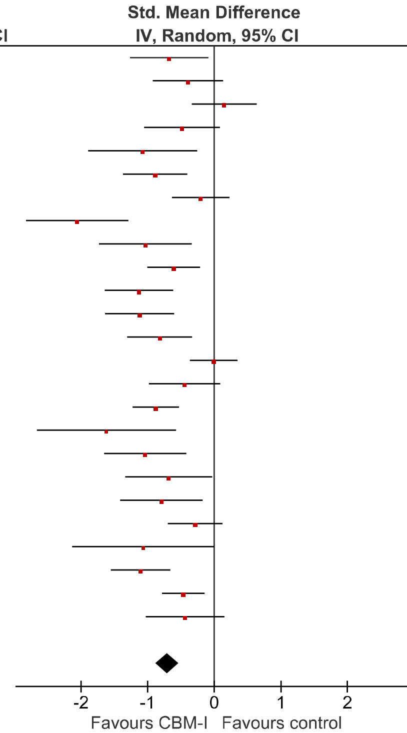

Figure 4: Forest plot of effect size of CBM-I versus control on anxiety post-training

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Belli & Lau (2014)	-0.3061	0.29349	24	23	4.5%	-0.31 [-0.88, 0.27]
Burnett Heyes et al (2017)	-0.0545	0.13023	29	30	10.3%	-0.05 [-0.31, 0.20]
Fu, Du, Au & Lau (2013)	0.1196	0.3822	16	12	3.0%	0.12 [-0.63, 0.87]
Fu, Du, Au & Lau (2015)	-0.298	0.23541	37	36	6.0%	-0.30 [-0.76, 0.16]
Lau, Belli & Chopra (2013)	0.2562	0.31753	20	20	4.0%	0.26 [-0.37, 0.88]
Lau, Molyneaux et al (2011)	-0.4415	0.33786	17	19	3.7%	-0.44 [-1.10, 0.22]
Lester et al (2011a)	-0.0565	0.197117	51	52	7.3%	-0.06 [-0.44, 0.33]
Lester et al (2011b)	-0.5796	0.2494	34	33	5.6%	-0.58 [-1.07, -0.09]
Orchard et al (2017)	0.4449	0.271	29	27	5.0%	0.44 [-0.09, 0.98]
Salemink & Wiers (2011)	0.1264	0.16457	73	75	8.6%	0.13 [-0.20, 0.45]
Vassilopoulos & Brouzos (2016)	-0.3688	0.327628	20	18	3.8%	-0.37 [-1.01, 0.27]
Vassilopoulos & Moberly (2013)	-0.4386	0.18908	61	54	7.6%	-0.44 [-0.81, -0.07]
Vassilopoulos et al (2009)	-1.0288	0.32462	22	21	3.9%	-1.03 [-1.67, -0.39]
Vassilopoulos, Blackwell et al (2014)	-0.1773	0.20839	53	41	6.9%	-0.18 [-0.59, 0.23]
Vassilopoulos, Moberly & Lau (2015)	-0.3549	0.21528	39	50	6.6%	-0.35 [-0.78, 0.07]
Vassilopoulos, Moberly & Zisimatou (2013)	0.0667	0.16173	77	76	8.8%	0.07 [-0.25, 0.38]
White et al (2016)	-0.1693	0.29875	23	22	4.4%	-0.17 [-0.75, 0.42]
Total (95% CI)			**625**	**609**	**100.0%**	**-0.17 [-0.31, -0.02]**

Heterogeneity: Tau² = 0.04; Chi² = 27.73, df = 16 (P = 0.03); I² = 42%
Test for overall effect: Z = 2.25 (P = 0.02)

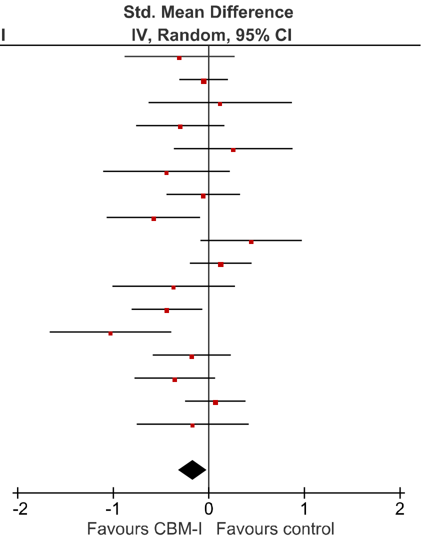

Table S1: Results of subgroup analyses

Moderator variable	**Outcome variable**	**Subgroup**	**Hedge's g (95% CI)**	***z***	***p***	I^2	**Test for subgroup differences** χ^2	*df*	*p*	I^2
Control condition										
	Negative interpretations									
		Negative training	-0.80 (-1.12, -0.49)	5.00	<0.001	78	1.16	2	0.56	0
		Neutral training	-0.58 (-0.85, -0.31)	4.22	<0.001	59				
		No training	-0.66 (-0.99, -0.33)	3.95	<0.001	0				
	Positive interpretations									
		Negative training	-0.74 (-1.06, -0.42)	4.49	<0.001	66	3.53	2	0.17	43
		Neutral training	-0.36 (-0.59, -0.12)	2.99	0.003	40				
		No training	-0.42 (-1.00, 0.17)	1.39	0.16	79				
	Anxiety post-training									
		Negative training	-0.27 (-0.45, -0.09)	2.89	0.004	13	2.89	2	0.24	30
		Neutral training	-0.06 (-0.22, 0.10)	0.71	0.48	0				
		No training	-0.19 (-0.74, 0.37)	0.66	0.51	78				
Age group										
	Negative interpretations									
		Children	-0.56 (-0.75, -0.38)	5.89	<0.001	52	3.22	1	.07	69
		Adolescents	-0.90 (-1.22, -0.58)	5.57	<0.001	70				
	Positive interpretations									
		Children	-0.46 (-0.74, -0.19)	3.34	0.02	57	0.29	1	.59	0
		Adolescents	-0.58 (-0.87, -0.28)	3.84	<0.001	66				
	Anxiety post-training									
		Children	-0.24 (-0.46, -0.02)	2.17	0.02	55	1.85	1	.17	46
		Adolescents	-0.05 (-0.21, 0.11)	.65	.52	0				

Table S2: Results of meta-regression analyses

Moderator variable	Outcome variable	Regression coefficient (95% CIs)	*p*	I^2
No. training trials				
	Negative interpretations	-0.001 (-0.003, 0.001)	0.27	65.3
	Positive interpretations	-0.001 (-0.017, 0.15)	0.89	62.5
	Anxiety post-training	0.005 (-0.008, 0.018)	0.43	43.3
% males				
	Negative interpretations	0.000 (-0.013, 0.009)	0.99	66.0
	Positive interpretations	-0.001 (-0.026, 0.003)	0.73	62.1
	Anxiety post-training	0.005 (-0.003, 0.013)	0.22	41.6

Figure S1: Funnel plot of publication bias for negative interpretaions

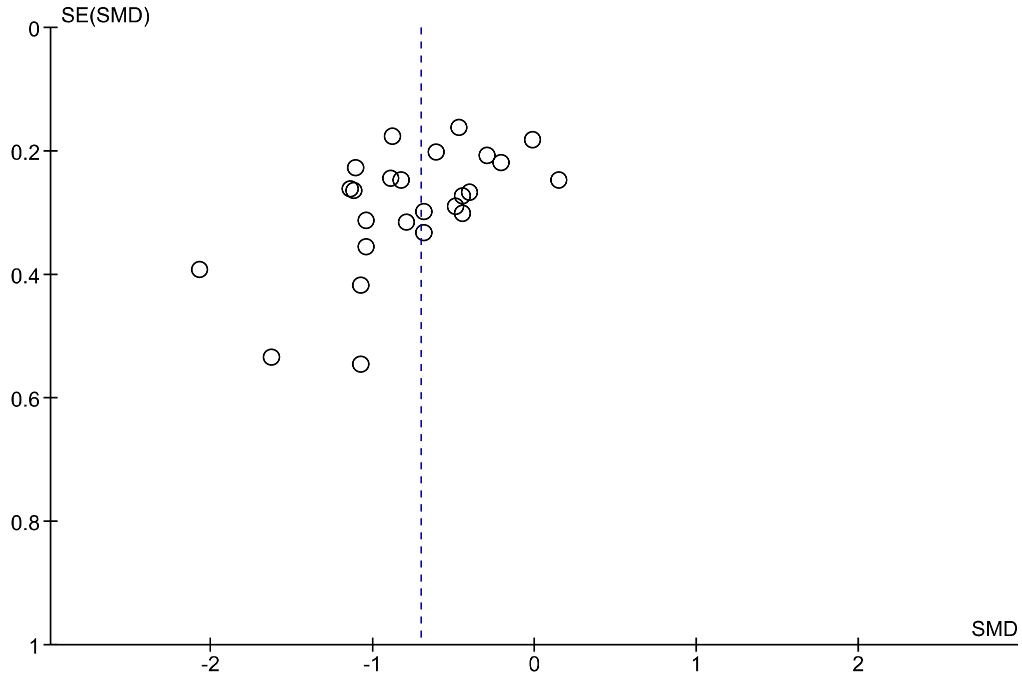

Figure S2: Funnel plot of publication bias for positive interpretations

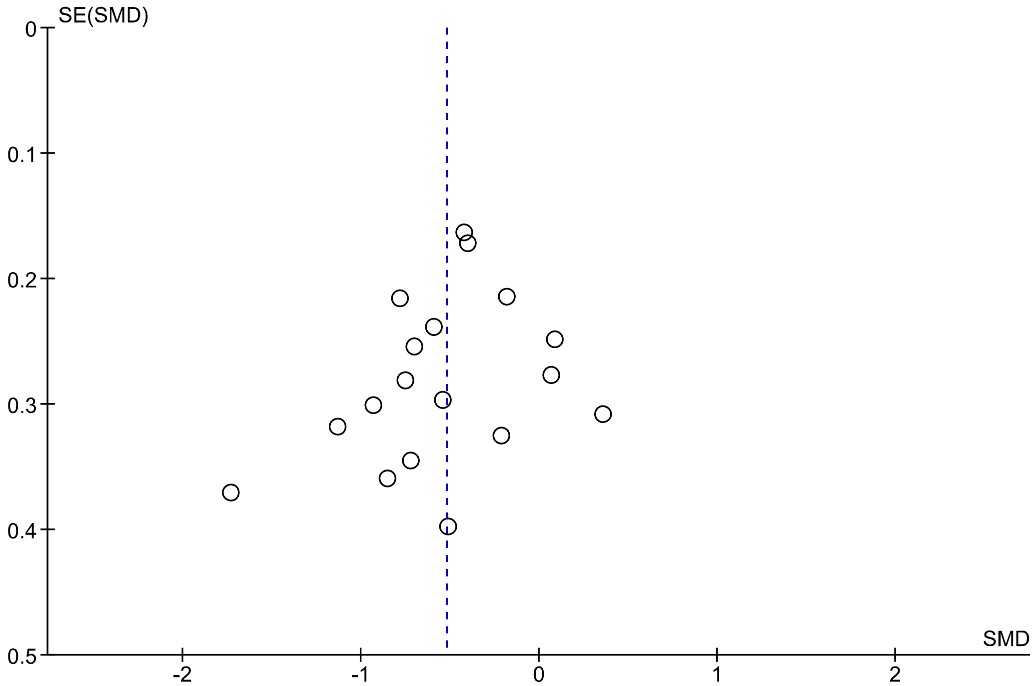

Figure S3: Funnel plot of publication bias for anxiety post-training

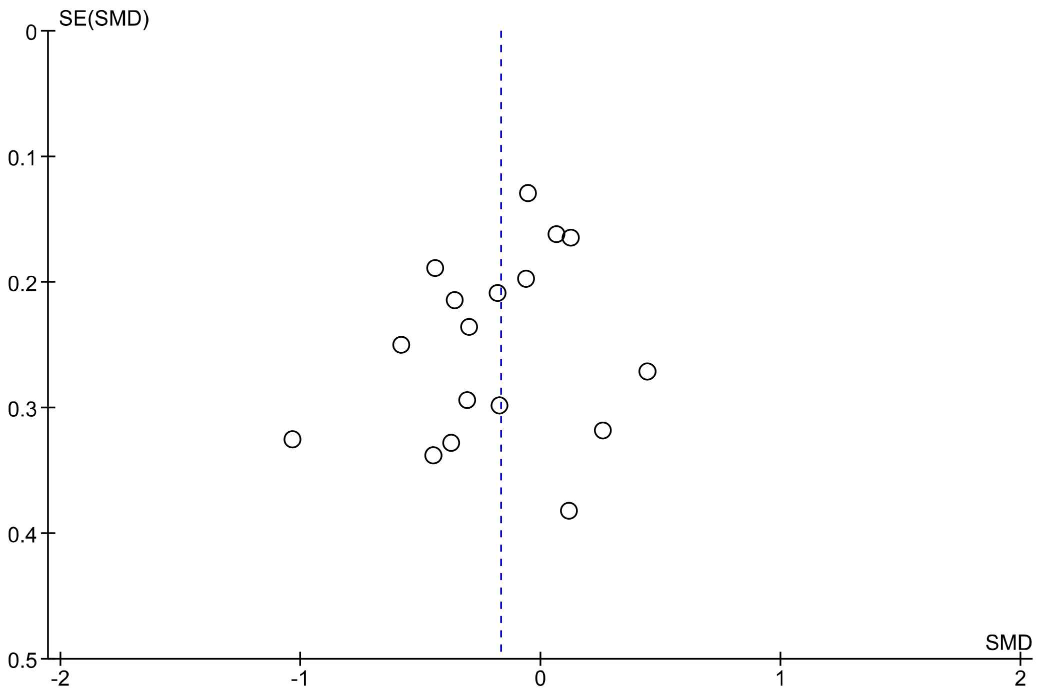

Figure S4: Funnel plot of publication bias for anxiety post-stressor

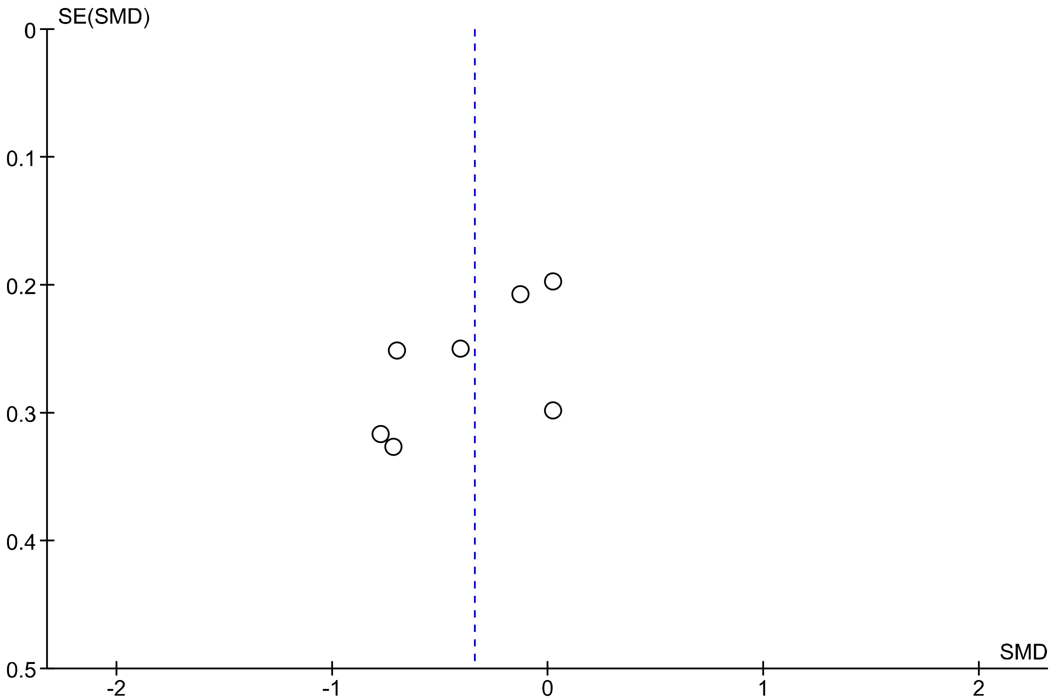

Figure S5: Forest plot of effect size of CBM-I versus control on positive interpretation bias

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Belli & Lau (2014)	-0.54	0.29704	24	23	5.2%	-0.54 [-1.12, 0.04]
Burnett Heyes et al (2017)	0.0649	0.2776	25	27	5.6%	0.06 [-0.48, 0.61]
Chan, Reynolds & Lau (2015)	0.0932	0.24824	33	32	6.1%	0.09 [-0.39, 0.58]
De Winter et al (2017)	-0.9307	0.300842	25	24	5.2%	-0.93 [-1.52, -0.34]
Fu, Du, Au & Lau (2013)	-0.512	0.39781	16	11	3.8%	-0.51 [-1.29, 0.27]
Fu, Du, Au & Lau (2015)	-0.5876	0.23908	37	36	6.2%	-0.59 [-1.06, -0.12]
Lau, Belli & Chopra (2013)	-1.7203	0.37012	20	20	4.2%	-1.72 [-2.45, -0.99]
Lau, Molyneaux et al (2011)	-0.7178	0.34439	17	19	4.5%	-0.72 [-1.39, -0.04]
Lothmann et al (2011)	-0.6995	0.2537	32	34	6.0%	-0.70 [-1.20, -0.20]
Orchard et al (2017)	-0.748	0.281	28	27	5.5%	-0.75 [-1.30, -0.20]
Salemink & Wiers (2011)	-0.4	0.17153	73	66	7.6%	-0.40 [-0.74, -0.06]
Telman et al (2013)	-1.1294	0.3175	23	23	4.9%	-1.13 [-1.75, -0.51]
Vassilopoulos & Brouzos (2016)	-0.2092	0.325765	20	18	4.8%	-0.21 [-0.85, 0.43]
Vassilopoulos et al (2009)	0.3635	0.30758	22	21	5.1%	0.36 [-0.24, 0.97]
Vassilopoulos, Blackwell et al (2014)	-0.7785	0.21559	53	41	6.7%	-0.78 [-1.20, -0.36]
Vassilopoulos, Brouzos & Andreau (2014)	-0.843	0.35847	16	18	4.3%	-0.84 [-1.55, -0.14]
Vassilopoulos, Moberly & Lau (2015)	-0.1796	0.21406	39	50	6.7%	-0.18 [-0.60, 0.24]
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4197	0.16347	77	76	7.7%	-0.42 [-0.74, -0.10]
Total (95% CI)			**580**	**566**	**100.0%**	**-0.52 [-0.72, -0.32]**

Heterogeneity: Tau² = 0.10; Chi² = 42.69, df = 17 (P = 0.0005); I² = 60%
Test for overall effect: Z = 5.22 (P < 0.00001)

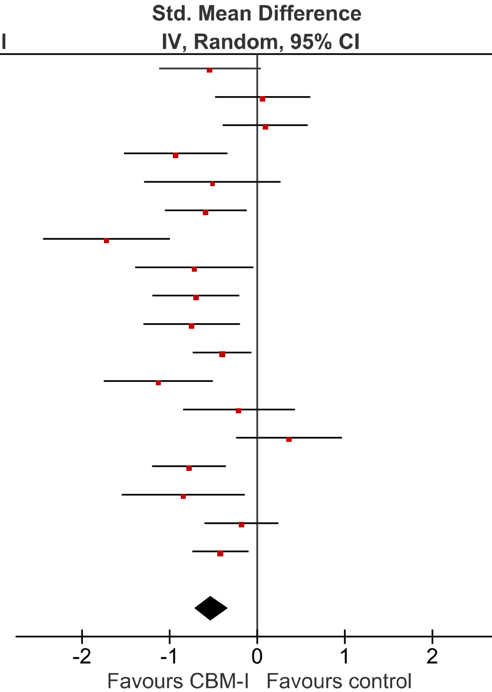

Figure S6: Forest plot of effect size of CBM-I versus control on anxiety post-stressor

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
Chan, Reynolds & Lau (2015)	-0.404	0.25061	33	32	14.8%	-0.40 [-0.90, 0.09]
Lau, Belli & Chopra (2013)	-0.7143	0.32617	20	20	10.8%	-0.71 [-1.35, -0.08]
Lester et al (2011a)	0.0291	0.19707	51	52	18.5%	0.03 [-0.36, 0.42]
Lester et al (2011b)	-0.7001	0.25173	34	33	14.7%	-0.70 [-1.19, -0.21]
Vassilopoulos et al (2009)	-0.7724	0.31622	22	21	11.3%	-0.77 [-1.39, -0.15]
Vassilopoulos, Blackwell et al (2014)	-0.1229	0.20819	53	41	17.7%	-0.12 [-0.53, 0.29]
White et al (2016)	0.0282	0.29824	23	22	12.1%	0.03 [-0.56, 0.61]
Total (95% CI)			**236**	**221**	**100.0%**	**-0.34 [-0.60, -0.08]**

Heterogeneity: $Tau^2 = 0.06$; $Chi^2 = 11.27$, df = 6 (P = 0.08); $I^2 = 47\%$
Test for overall effect: Z = 2.56 (P = 0.01)

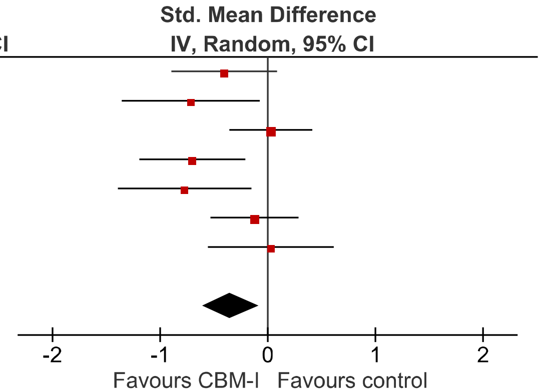

Figure S7: Forest plot of control group comparison for negative interpretations

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.1.1 Negative training control						
Belli & Lau (2014)	-0.6775	0.30003	24	23		Not estimable
Burnett Heyes et al (2017)	-0.3953	0.269	28	29		Not estimable
Chan, Reynolds & Lau (2015)	0.1488	0.24844	33	32		Not estimable
De Winter et al (2017)	-0.485	0.289923	25	24		Not estimable
Fu, Du, Au & Lau (2013)	-1.0739	0.41804	16	11		Not estimable
Fu, Du, Au & Lau (2015)	-0.8873	0.24533	37	36		Not estimable
Klein et al (2015)	-0.2035	0.22023	40	43	4.7%	-0.20 [-0.64, 0.23]
Lau, Belli & Chopra (2013)	-2.0586	0.3911	20	20	3.0%	-2.06 [-2.83, -1.29]
Lau, Molyneaux et al (2011)	-1.0349	0.35543	17	19	3.3%	-1.03 [-1.73, -0.34]
Lester et al (2011a)	-0.6085	0.20158	51	52	4.9%	-0.61 [-1.00, -0.21]
Lester et al (2011b)	-1.1295	0.26311	34	33	4.2%	-1.13 [-1.65, -0.61]
Lothmann et al (2011)	-1.1182	0.26482	32	34	4.2%	-1.12 [-1.64, -0.60]
Muris et al (2008)	-0.8196	0.24898	36	34	4.4%	-0.82 [-1.31, -0.33]
Muris et al (2009)	-0.0076	0.18281	63	57	5.2%	-0.01 [-0.37, 0.35]
Orchard et al (2017)	-0.4463	0.273061	28	27		Not estimable
Salemink & Wiers (2011)	-0.8785	0.17783	73	66		Not estimable
Stoddard et al (2016): Experiment 2	-1.6179	0.533648	8	11		Not estimable
Telman et al (2013)	-1.038	0.31411	23	23	3.7%	-1.04 [-1.65, -0.42]
Vassilopoulos & Brouzos (2016)	-0.6843	0.334235	20	18		Not estimable
Vassilopoulos et al (2009)	-0.7939	0.31686	22	21		Not estimable
Vassilopoulos, Blackwell et al (2014)	-0.2865	0.20903	53	41	4.9%	-0.29 [-0.70, 0.12]
Vassilopoulos, Brouzos & Andreau (2014)	-1.067	0.54439	16	18		Not estimable
Vassilopoulos, Moberly & Lau (2015)	-1.1073	0.22918	39	50	4.6%	-1.11 [-1.56, -0.66]
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4648	0.16385	77	76		Not estimable
White et al (2016)	-0.4396	0.301786	23	22		Not estimable
Subtotal (95% CI)			**408**	**406**	**47.2%**	**-0.80 [-1.12, -0.49]**
Heterogeneity: Tau² = 0.22; Chi² = 46.02, df = 10 (P < 0.00001); I² = 78%						
Test for overall effect: Z = 5.00 (P < 0.00001)						
4.1.2 Neutral training control						
Belli & Lau (2014)	-0.6775	0.30003	24	23	3.9%	-0.68 [-1.27, -0.09]
Burnett Heyes et al (2017)	-0.3953	0.269	28	29	4.2%	-0.40 [-0.92, 0.13]
Chan, Reynolds & Lau (2015)	0.1488	0.24844	33	32	4.4%	0.15 [-0.34, 0.64]
De Winter et al (2017)	-0.485	0.289923	25	24	4.0%	-0.48 [-1.05, 0.08]
Fu, Du, Au & Lau (2013)	-1.0739	0.41804	16	11	2.8%	-1.07 [-1.89, -0.25]
Fu, Du, Au & Lau (2015)	-0.8873	0.24533	37	36	4.4%	-0.89 [-1.37, -0.41]
Klein et al (2015)	-0.2035	0.22023	40	43	4.7%	-0.20 [-0.64, 0.23]
Lau, Belli & Chopra (2013)	-2.0586	0.3911	20	20		Not estimable
Lau, Molyneaux et al (2011)	-1.0349	0.35543	17	19		Not estimable
Lester et al (2011a)	-0.6085	0.20158	51	52		Not estimable
Lester et al (2011b)	-1.1295	0.26311	34	33		Not estimable
Lothmann et al (2011)	-1.1182	0.26482	32	34		Not estimable
Muris et al (2008)	-0.8196	0.24898	36	34		Not estimable
Muris et al (2009)	-0.0076	0.18281	63	57		Not estimable
Orchard et al (2017)	-0.4463	0.273061	28	27		Not estimable
Salemink & Wiers (2011)	-0.8785	0.17783	73	66	5.2%	-0.88 [-1.23, -0.53]
Stoddard et al (2016): Experiment 2	-1.6179	0.533648	8	11	2.1%	-1.62 [-2.66, -0.57]
Telman et al (2013)	-1.038	0.31411	23	23		Not estimable
Vassilopoulos & Brouzos (2016)	-0.6843	0.334235	20	18		Not estimable
Vassilopoulos et al (2009)	-0.7939	0.31686	22	21		Not estimable
Vassilopoulos, Blackwell et al (2014)	-0.2865	0.20903	53	41		Not estimable
Vassilopoulos, Brouzos & Andreau (2014)	-1.067	0.54439	16	18		Not estimable
Vassilopoulos, Moberly & Lau (2015)	-1.1073	0.22918	39	50		Not estimable
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4648	0.16385	77	76		Not estimable
White et al (2016)	-0.4396	0.301786	23	22	3.8%	-0.44 [-1.03, 0.15]
Subtotal (95% CI)			**307**	**297**	**39.5%**	**-0.58 [-0.85, -0.31]**
Heterogeneity: Tau² = 0.10; Chi² = 21.94, df = 9 (P = 0.009); I² = 59%						
Test for overall effect: Z = 4.22 (P < 0.0001)						
4.1.3 No training control						
Belli & Lau (2014)	-0.6775	0.30003	24	23		Not estimable
Burnett Heyes et al (2017)	-0.3953	0.269	28	29		Not estimable
Chan, Reynolds & Lau (2015)	0.1488	0.24844	33	32		Not estimable
De Winter et al (2017)	-0.485	0.289923	25	24		Not estimable
Fu, Du, Au & Lau (2013)	-1.0739	0.41804	16	11		Not estimable
Fu, Du, Au & Lau (2015)	-0.8873	0.24533	37	36		Not estimable
Klein et al (2015)	-0.2035	0.22023	40	43		Not estimable
Lau, Belli & Chopra (2013)	-2.0586	0.3911	20	20		Not estimable
Lau, Molyneaux et al (2011)	-1.0349	0.35543	17	19		Not estimable
Lester et al (2011a)	-1.1295	0.26311	34	33		Not estimable
Lester et al (2011b)	-0.6085	0.20158	51	52		Not estimable
Lothmann et al (2011)	-1.1182	0.26482	32	34		Not estimable
Muris et al (2008)	-0.8196	0.24898	36	34		Not estimable
Muris et al (2009)	-0.0076	0.18281	63	57		Not estimable
Orchard et al (2017)	-0.4463	0.273061	28	27	4.1%	-0.45 [-0.98, 0.09]
Salemink & Wiers (2011)	-0.8785	0.17783	73	66		Not estimable
Stoddard et al (2016): Experiment 2	-1.6179	0.533648	8	11		Not estimable
Telman et al (2013)	-1.038	0.31411	23	23		Not estimable
Vassilopoulos & Brouzos (2016)	-0.6843	0.334235	20	18	3.5%	-0.68 [-1.34, -0.03]
Vassilopoulos et al (2009)	-0.7939	0.31686	22	21	3.7%	-0.79 [-1.41, -0.17]
Vassilopoulos, Blackwell et al (2014)	-0.2865	0.20903	53	41		Not estimable
Vassilopoulos, Brouzos & Andreau (2014)	-1.067	0.54439	16	18	2.0%	-1.07 [-2.13, -0.00]
Vassilopoulos, Moberly & Lau (2015)	-1.1073	0.22918	39	50		Not estimable
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4648	0.16385	77	76		Not estimable
White et al (2016)	-0.4396	0.301786	23	22		Not estimable
Subtotal (95% CI)			**86**	**84**	**13.3%**	**-0.66 [-0.99, -0.33]**
Heterogeneity: Tau² = 0.00; Chi² = 1.36, df = 3 (P = 0.72); I² = 0%						
Test for overall effect: Z = 3.95 (P < 0.0001)						
Total (95% CI)			**801**	**787**	**100.0%**	**-0.69 [-0.88, -0.51]**
Heterogeneity: Tau² = 0.13; Chi² = 70.31, df = 24 (P < 0.00001); I² = 66%						
Test for overall effect: Z = 7.48 (P < 0.00001)						
Test for subgroup differences: Chi² = 1.16, df = 2 (P = 0.56), I² = 0%						

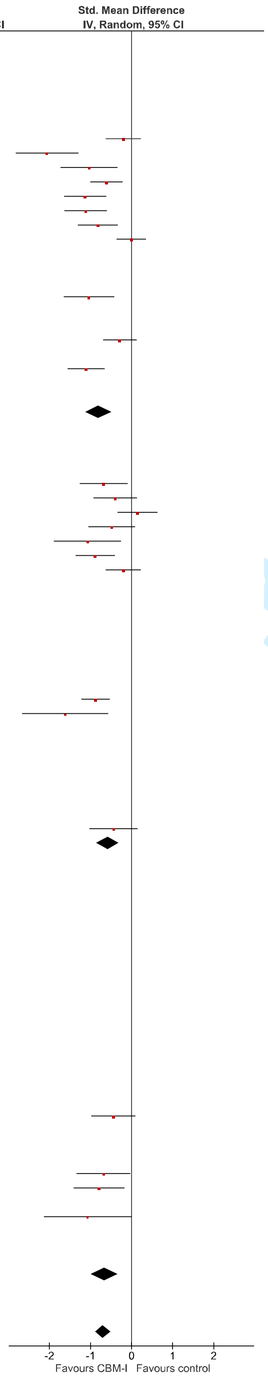

Figure S8: Forest plot of control group comparison for positive interpretations

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.2.1 Negative training control						
Belli & Lau (2014)	-0.54	0.29704	24	23	0.0%	-0.54 [-1.12, 0.04]
Burnett Heyes et al (2017)	0.0649	0.2776	25	27	0.0%	0.06 [-0.48, 0.61]
Chan, Reynolds & Lau (2015)	0.0932	0.24824	33	32	0.0%	0.09 [-0.39, 0.58]
De Winter et al (2017)	-0.9307	0.300842	25	24	0.0%	-0.93 [-1.52, -0.34]
Fu, Du, Au & Lau (2013)	-0.512	0.39781	16	11	0.0%	-0.51 [-1.29, 0.27]
Fu, Du, Au & Lau (2015)	-0.5876	0.23908	37	36	0.0%	-0.59 [-1.06, -0.12]
Lau, Belli & Chopra (2013)	-1.7203	0.37012	20	20	4.0%	-1.72 [-2.45, -0.99]
Lau, Molyneaux et al (2011)	-0.7178	0.34439	17	19	4.4%	-0.72 [-1.39, -0.04]
Lothmann et al (2011)	-0.6995	0.2537	32	34	5.7%	-0.70 [-1.20, -0.20]
Orchard et al (2017)	-0.748	0.281	28	27	0.0%	-0.75 [-1.30, -0.20]
Salemink & Wiers (2011)	-0.4	0.17153	73	66	0.0%	-0.40 [-0.74, -0.06]
Telman et al (2013)	-1.1294	0.3175	23	23	4.7%	-1.13 [-1.75, -0.51]
Vassilopoulos & Brouzos (2016)	-0.2092	0.325765	20	18	0.0%	-0.21 [-0.85, 0.43]
Vassilopoulos et al (2009)	0.3635	0.30758	22	21	0.0%	0.36 [-0.24, 0.97]
Vassilopoulos, Blackwell et al (2014)	-0.7785	0.21559	53	41	6.4%	-0.78 [-1.20, -0.36]
Vassilopoulos, Brouzos & Andreau (2014)	-0.843	0.35847	16	18	0.0%	-0.84 [-1.55, -0.14]
Vassilopoulos, Moberly & Lau (2015)	-0.1796	0.21406	39	50	6.4%	-0.18 [-0.60, 0.24]
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4197	0.16347	77	76	7.4%	-0.42 [-0.74, -0.10]
Subtotal (95% CI)			**261**	**263**	**39.1%**	**-0.74 [-1.06, -0.42]**
Heterogeneity: Tau² = 0.12; Chi² = 17.85, df = 6 (P = 0.007); I² = 66%						
Test for overall effect: Z = 4.49 (P < 0.00001)						
4.2.2 Neutral training control						
Belli & Lau (2014)	-0.54	0.29704	24	23	5.0%	-0.54 [-1.12, 0.04]
Burnett Heyes et al (2017)	0.0649	0.2776	25	27	5.3%	0.06 [-0.48, 0.61]
Chan, Reynolds & Lau (2015)	0.0932	0.24824	33	32	5.8%	0.09 [-0.39, 0.58]
Fu, Du, Au & Lau (2013)	-0.9307	0.300842	25	24	5.0%	-0.93 [-1.52, -0.34]
Fu, Du, Au & Lau (2015)	-0.512	0.39781	16	11	3.7%	-0.51 [-1.29, 0.27]
Lau, Belli & Chopra (2013)	-0.5876	0.23908	37	36	0.0%	-0.59 [-1.06, -0.12]
Lau, Molyneaux et al (2011)	-1.7203	0.37012	20	20	0.0%	-1.72 [-2.45, -0.99]
Lester et al (2011a)	-0.7178	0.34439	17	19	0.0%	-0.72 [-1.39, -0.04]
Lester et al (2011b)	-0.6995	0.2537	32	34	0.0%	-0.70 [-1.20, -0.20]
Orchard et al (2017)	-0.748	0.281	28	27	0.0%	-0.75 [-1.30, -0.20]
Salemink & Wiers (2011)	-0.4	0.17153	73	66	7.2%	-0.40 [-0.74, -0.06]
Vassilopoulos & Brouzos (2016)	-1.1294	0.3175	23	23	0.0%	-1.13 [-1.75, -0.51]
Vassilopoulos & Moberly (2013)	-0.2092	0.325765	20	18	0.0%	-0.21 [-0.85, 0.43]
Vassilopoulos et al (2009)	0.3635	0.30758	22	21	0.0%	0.36 [-0.24, 0.97]
Vassilopoulos, Blackwell et al (2014)	-0.7785	0.21559	53	41	0.0%	-0.78 [-1.20, -0.36]
Vassilopoulos, Moberly & Lau (2015)	-0.843	0.35847	16	18	0.0%	-0.84 [-1.55, -0.14]
Vassilopoulos, Moberly & Zisimatou (2013)	-0.1796	0.21406	39	50	0.0%	-0.18 [-0.60, 0.24]
White et al (2016)	-0.4197	0.16347	77	76	7.4%	-0.42 [-0.74, -0.10]
Subtotal (95% CI)			**273**	**259**	**39.5%**	**-0.36 [-0.59, -0.12]**
Heterogeneity: Tau² = 0.04; Chi² = 9.98, df = 6 (P = 0.13); I² = 40%						
Test for overall effect: Z = 2.99 (P = 0.003)						
4.2.3 No training control						
Belli & Lau (2014)	-0.54	0.29704	24	23	0.0%	-0.54 [-1.12, 0.04]
Burnett Heyes et al (2017)	0.0649	0.2776	25	27	0.0%	0.06 [-0.48, 0.61]
Chan, Reynolds & Lau (2015)	0.0932	0.24824	33	32	0.0%	0.09 [-0.39, 0.58]
Fu, Du, Au & Lau (2013)	-0.9307	0.300842	25	24	0.0%	-0.93 [-1.52, -0.34]
Fu, Du, Au & Lau (2015)	-0.512	0.39781	16	11	0.0%	-0.51 [-1.29, 0.27]
Lau, Belli & Chopra (2013)	-0.5876	0.23908	37	36	0.0%	-0.59 [-1.06, -0.12]
Lau, Molyneaux et al (2011)	-1.7203	0.37012	20	20	0.0%	-1.72 [-2.45, -0.99]
Lester et al (2011a)	-0.7178	0.34439	17	19	0.0%	-0.72 [-1.39, -0.04]
Lester et al (2011b)	-0.6995	0.2537	32	34	0.0%	-0.70 [-1.20, -0.20]
Orchard et al (2017)	-0.748	0.281	28	27	5.3%	-0.75 [-1.30, -0.20]
Salemink & Wiers (2011)	-0.4	0.17153	73	66	0.0%	-0.40 [-0.74, -0.06]
Vassilopoulos & Brouzos (2016)	-1.1294	0.3175	23	23	4.7%	-1.13 [-1.75, -0.51]
Vassilopoulos & Moberly (2013)	-0.2092	0.325765	20	18	0.0%	-0.21 [-0.85, 0.43]
Vassilopoulos et al (2009)	0.3635	0.30758	22	21	4.9%	0.36 [-0.24, 0.97]
Vassilopoulos, Blackwell et al (2014)	-0.7785	0.21559	53	41	0.0%	-0.78 [-1.20, -0.36]
Vassilopoulos, Moberly & Lau (2015)	-0.843	0.35847	16	18	0.0%	-0.84 [-1.55, -0.14]
Vassilopoulos, Moberly & Zisimatou (2013)	-0.1796	0.21406	39	50	6.4%	-0.18 [-0.60, 0.24]
White et al (2016)	-0.4197	0.16347	77	76	0.0%	-0.42 [-0.74, -0.10]
Subtotal (95% CI)			**112**	**121**	**21.4%**	**-0.42 [-1.00, 0.17]**
Heterogeneity: Tau² = 0.28; Chi² = 14.01, df = 3 (P = 0.003); I² = 79%						
Test for overall effect: Z = 1.39 (P = 0.16)						
Total (95% CI)			**646**	**643**	**100.0%**	**-0.52 [-0.71, -0.32]**
Heterogeneity: Tau² = 0.11; Chi² = 47.05, df = 17 (P = 0.0001); I² = 64%						
Test for overall effect: Z = 5.23 (P < 0.00001)						
Test for subgroup differences: Chi² = 3.53, df = 2 (P = 0.17), I² = 43.3%						

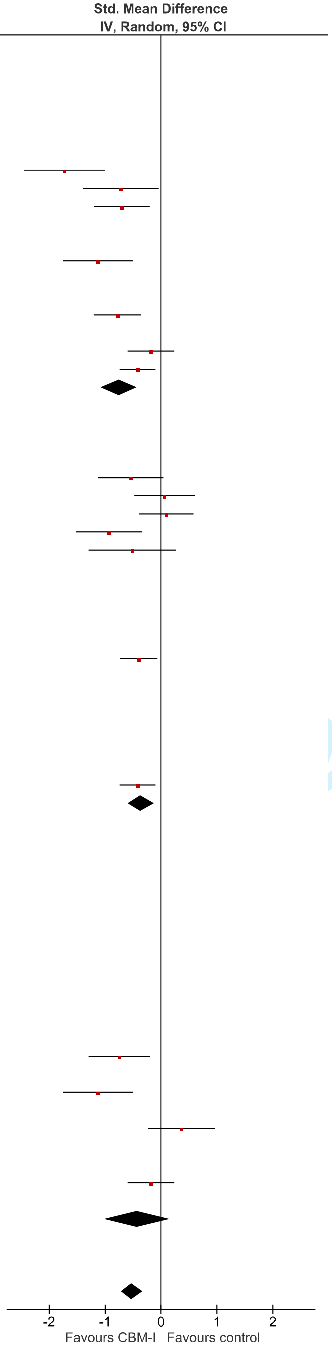

Figure S9: Forest plot of control group comparison for anxiety post-training

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.7.1 Positive training control						
Belli & Lau (2014)	-0.3061	0.29349	24	23	0.0%	-0.31 [-0.88, 0.27]
Burnett Heyes et al (2017)	-0.0545	0.13023	29	30	0.0%	-0.05 [-0.31, 0.20]
Fu, Du, Au & Lau (2013)	0.1196	0.3822	16	12	0.0%	0.12 [-0.63, 0.87]
Fu, Du, Au & Lau (2015)	-0.298	0.23541	37	36	0.0%	-0.30 [-0.76, 0.16]
Lau, Belli & Chopra (2013)	0.2562	0.31753	20	20	4.0%	0.26 [-0.37, 0.88]
Lau, Molyneaux et al (2011)	-0.4415	0.33786	17	19	3.7%	-0.44 [-1.10, 0.22]
Lester et al (2011a)	-0.0565	0.197117	51	52	7.3%	-0.06 [-0.44, 0.33]
Lester et al (2011b)	-0.5796	0.2494	34	33	5.6%	-0.58 [-1.07, -0.09]
Orchard et al (2017)	0.4449	0.271	29	27	0.0%	0.44 [-0.09, 0.98]
Salemink & Wiers (2011)	0.1264	0.16457	73	75	0.0%	0.13 [-0.20, 0.45]
Vassilopoulos & Brouzos (2016)	-0.3688	0.327628	20	18	0.0%	-0.37 [-1.01, 0.27]
Vassilopoulos & Moberly (2013)	-0.4386	0.18908	61	54	7.6%	-0.44 [-0.81, -0.07]
Vassilopoulos et al (2009)	-1.0288	0.32462	22	21	0.0%	-1.03 [-1.67, -0.39]
Vassilopoulos, Blackwell et al (2014)	-0.1773	0.20839	53	41	6.9%	-0.18 [-0.59, 0.23]
Vassilopoulos, Moberly & Lau (2015)	-0.3549	0.21528	39	50	6.6%	-0.35 [-0.78, 0.07]
Vassilopoulos, Moberly & Zisimatou (2013)	0.0667	0.16173	77	76	0.0%	0.07 [-0.25, 0.38]
White et al (2016)	-0.1693	0.29875	23	22	0.0%	-0.17 [-0.75, 0.42]
Subtotal (95% CI)			**275**	**269**	**41.7%**	**-0.27 [-0.45, -0.09]**
Heterogeneity: Tau² = 0.01; Chi² = 6.87, df = 6 (P = 0.33); I² = 13%						
Test for overall effect: Z = 2.89 (P = 0.004)						
4.7.2 Neutral training control						
Belli & Lau (2014)	-0.3061	0.29349	24	23	4.5%	-0.31 [-0.88, 0.27]
Burnett Heyes et al (2017)	-0.0545	0.13023	29	30	10.3%	-0.05 [-0.31, 0.20]
Fu, Du, Au & Lau (2013)	0.1196	0.3822	16	12	3.0%	0.12 [-0.63, 0.87]
Fu, Du, Au & Lau (2015)	-0.298	0.23541	37	36	6.0%	-0.30 [-0.76, 0.16]
Lau, Belli & Chopra (2013)	0.2562	0.31753	20	20	0.0%	0.26 [-0.37, 0.88]
Lau, Molyneaux et al (2011)	-0.4415	0.33786	17	19	0.0%	-0.44 [-1.10, 0.22]
Lester et al (2011a)	-0.0565	0.197117	51	52	0.0%	-0.06 [-0.44, 0.33]
Lester et al (2011b)	-0.5796	0.2494	34	33	0.0%	-0.58 [-1.07, -0.09]
Orchard et al (2017)	0.4449	0.271	29	27	0.0%	0.44 [-0.09, 0.98]
Salemink & Wiers (2011)	0.1264	0.16457	73	75	8.6%	0.13 [-0.20, 0.45]
Vassilopoulos & Brouzos (2016)	-0.3688	0.327628	20	18	0.0%	-0.37 [-1.01, 0.27]
Vassilopoulos & Moberly (2013)	-0.4386	0.18908	61	54	0.0%	-0.44 [-0.81, -0.07]
Vassilopoulos et al (2009)	-1.0288	0.32462	22	21	0.0%	-1.03 [-1.67, -0.39]
Vassilopoulos, Blackwell et al (2014)	-0.1773	0.20839	53	41	0.0%	-0.18 [-0.59, 0.23]
Vassilopoulos, Moberly & Lau (2015)	-0.3549	0.21528	39	50	0.0%	-0.35 [-0.78, 0.07]
Vassilopoulos, Moberly & Zisimatou (2013)	0.0667	0.16173	77	76	0.0%	0.07 [-0.25, 0.38]
White et al (2016)	-0.1693	0.29875	23	22	4.4%	-0.17 [-0.75, 0.42]
Subtotal (95% CI)			**202**	**198**	**36.8%**	**-0.06 [-0.22, 0.10]**
Heterogeneity: Tau² = 0.00; Chi² = 3.36, df = 5 (P = 0.64); I² = 0%						
Test for overall effect: Z = 0.71 (P = 0.48)						
4.7.3 No training control						
Belli & Lau (2014)	-0.3061	0.29349	24	23	0.0%	-0.31 [-0.88, 0.27]
Burnett Heyes et al (2017)	-0.0545	0.13023	29	30	0.0%	-0.05 [-0.31, 0.20]
Fu, Du, Au & Lau (2013)	0.1196	0.3822	16	12	0.0%	0.12 [-0.63, 0.87]
Fu, Du, Au & Lau (2015)	-0.298	0.23541	37	36	0.0%	-0.30 [-0.76, 0.16]
Lau, Belli & Chopra (2013)	0.2562	0.31753	20	20	0.0%	0.26 [-0.37, 0.88]
Lau, Molyneaux et al (2011)	-0.4415	0.33786	17	19	0.0%	-0.44 [-1.10, 0.22]
Lester et al (2011a)	-0.0565	0.197117	51	52	0.0%	-0.06 [-0.44, 0.33]
Lester et al (2011b)	-0.5796	0.2494	34	33	0.0%	-0.58 [-1.07, -0.09]
Orchard et al (2017)	0.4449	0.271	29	27	5.0%	0.44 [-0.09, 0.98]
Salemink & Wiers (2011)	0.1264	0.16457	73	75	0.0%	0.13 [-0.20, 0.45]
Vassilopoulos & Brouzos (2016)	-0.3688	0.327628	20	18	3.8%	-0.37 [-1.01, 0.27]
Vassilopoulos & Moberly (2013)	-0.4386	0.18908	61	54	0.0%	-0.44 [-0.81, -0.07]
Vassilopoulos et al (2009)	-1.0288	0.32462	22	21	3.9%	-1.03 [-1.67, -0.39]
Vassilopoulos, Blackwell et al (2014)	-0.1773	0.20839	53	41	0.0%	-0.18 [-0.59, 0.23]
Vassilopoulos, Moberly & Lau (2015)	-0.3549	0.21528	39	50	0.0%	-0.35 [-0.78, 0.07]
Vassilopoulos, Moberly & Zisimatou (2013)	0.0667	0.16173	77	76	8.8%	0.07 [-0.25, 0.38]
White et al (2016)	-0.1693	0.29875	23	22	0.0%	-0.17 [-0.75, 0.42]
Subtotal (95% CI)			**148**	**142**	**21.5%**	**-0.19 [-0.74, 0.37]**
Heterogeneity: Tau² = 0.24; Chi² = 13.88, df = 3 (P = 0.003); I² = 78%						
Test for overall effect: Z = 0.66 (P = 0.51)						
Total (95% CI)			**625**	**609**	**100.0%**	**-0.17 [-0.31, -0.02]**
Heterogeneity: Tau² = 0.04; Chi² = 27.73, df = 16 (P = 0.03); I² = 42%						
Test for overall effect: Z = 2.25 (P = 0.02)						
Test for subgroup differences: Chi² = 2.84, df = 2 (P = 0.24), I² = 29.5%						

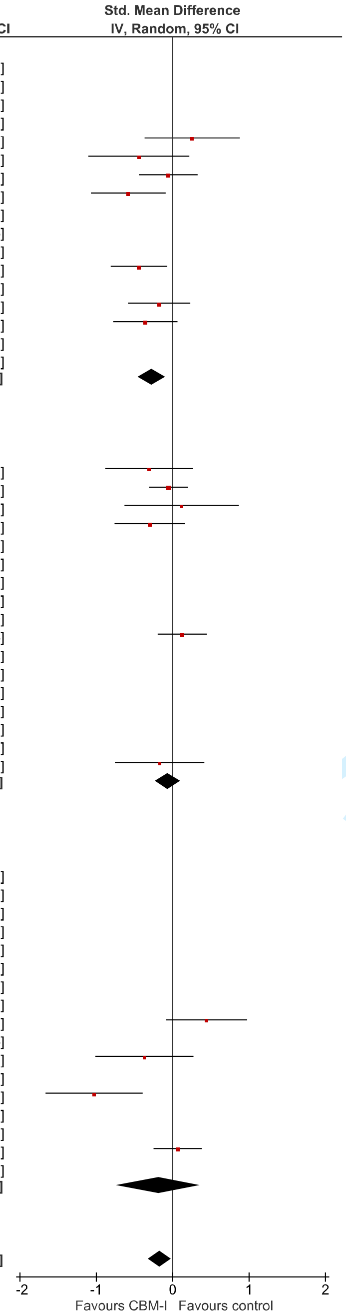

Figure S10: Forest plot of age group comparison for negative interpretations

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.1.1 Children						
Belli & Lau (2014)	-0.6775	0.30003	24	23		Not estimable
Burnett Heyes et al (2017)	-0.3953	0.269	28	29		Not estimable
Chan, Reynolds & Lau (2015)	0.1488	0.24844	33	32		Not estimable
De Winter et al (2017)	-0.485	0.289923	25	24	3.9%	-0.48 [-1.05, 0.08]
Fu, Du, Au & Lau (2013)	-1.0739	0.41804	16	11		Not estimable
Fu, Du, Au & Lau (2015)	-0.8873	0.24533	37	36		Not estimable
Klein et al (2015)	-0.2035	0.22023	40	43	4.7%	-0.20 [-0.64, 0.23]
Lau, Belli & Chopra (2013)	-2.0586	0.3911	20	20		Not estimable
Lau, Molyneaux et al (2011)	-1.0349	0.35543	17	19		Not estimable
Lester et al (2011a)	-0.6085	0.20158	51	52	5.0%	-0.61 [-1.00, -0.21]
Lester et al (2011b)	-1.1295	0.26311	34	33	4.2%	-1.13 [-1.65, -0.61]
Lothmann et al (2011)	-1.1182	0.26482	32	34		Not estimable
Muris et al (2008)	-0.8196	0.24898	36	34	4.4%	-0.82 [-1.31, -0.33]
Muris et al (2009)	-0.0076	0.18281	63	57	5.2%	-0.01 [-0.37, 0.35]
Orchard et al (2017)	-0.4463	0.273061	28	27	4.1%	-0.45 [-0.98, 0.09]
Salemink & Wiers (2011)	-0.8785	0.17783	73	66		Not estimable
Stoddard et al (2016): Experiment 2	-1.6179	0.533648	8	11		Not estimable
Telman et al (2013)	-1.038	0.31411	23	23		Not estimable
Vassilopoulos & Brouzos (2016)	-0.6843	0.334235	20	18	3.4%	-0.68 [-1.34, -0.03]
Vassilopoulos et al (2009)	-0.7939	0.31686	22	21	3.6%	-0.79 [-1.41, -0.17]
Vassilopoulos, Blackwell et al (2014)	-0.2865	0.20903	53	41	4.9%	-0.29 [-0.70, 0.12]
Vassilopoulos, Brouzos & Andreau (2014)	-1.067	0.54439	16	18	1.9%	-1.07 [-2.13, -0.00]
Vassilopoulos, Moberly & Lau (2015)	-1.1073	0.22918	39	50	4.6%	-1.11 [-1.56, -0.66]
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4648	0.16385	77	76	5.4%	-0.46 [-0.79, -0.14]
White et al (2016)	-0.4396	0.301786	23	22	3.8%	-0.44 [-1.03, 0.15]
Subtotal (95% CI)			**527**	**516**	**59.2%**	**-0.56 [-0.75, -0.38]**
Heterogeneity: Tau² = 0.06; Chi² = 27.00, df = 13 (P = 0.01); I² = 52%						
Test for overall effect: Z = 5.89 (P < 0.00001)						
4.1.2 Adolescents						
Belli & Lau (2014)	-0.6775	0.30003	24	23	3.8%	-0.68 [-1.27, -0.09]
Burnett Heyes et al (2017)	-0.3953	0.269	28	29	4.2%	-0.40 [-0.92, 0.13]
Chan, Reynolds & Lau (2015)	0.1488	0.24844	33	32	4.4%	0.15 [-0.34, 0.64]
De Winter et al (2017)	-0.485	0.289923	25	24		Not estimable
Fu, Du, Au & Lau (2013)	-1.0739	0.41804	16	11	2.7%	-1.07 [-1.89, -0.25]
Fu, Du, Au & Lau (2015)	-0.8873	0.24533	37	36	4.4%	-0.89 [-1.37, -0.41]
Klein et al (2015)	-0.2035	0.22023	40	43		Not estimable
Lau, Belli & Chopra (2013)	-2.0586	0.3911	20	20	2.9%	-2.06 [-2.83, -1.29]
Lau, Molyneaux et al (2011)	-1.0349	0.35543	17	19	3.2%	-1.03 [-1.73, -0.34]
Lester et al (2011a)	-0.6085	0.20158	51	52		Not estimable
Lester et al (2011b)	-1.1295	0.26311	34	33		Not estimable
Lothmann et al (2011)	-1.1182	0.26482	32	34	4.2%	-1.12 [-1.64, -0.60]
Muris et al (2008)	-0.8196	0.24898	36	34		Not estimable
Muris et al (2009)	-0.0076	0.18281	63	57		Not estimable
Orchard et al (2017)	-0.4463	0.273061	28	27		Not estimable
Salemink & Wiers (2011)	-0.8785	0.17783	73	66	5.3%	-0.88 [-1.23, -0.53]
Stoddard et al (2016): Experiment 2	-1.6179	0.533648	8	11	2.0%	-1.62 [-2.66, -0.57]
Telman et al (2013)	-1.038	0.31411	23	23	3.7%	-1.04 [-1.65, -0.42]
Vassilopoulos & Brouzos (2016)	-0.6843	0.334235	20	18		Not estimable
Vassilopoulos et al (2009)	-0.7939	0.31686	22	21		Not estimable
Vassilopoulos, Blackwell et al (2014)	-0.2865	0.20903	53	41		Not estimable
Vassilopoulos, Brouzos & Andreau (2014)	-1.067	0.54439	16	18		Not estimable
Vassilopoulos, Moberly & Lau (2015)	-1.1073	0.22918	39	50		Not estimable
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4648	0.16385	77	76		Not estimable
White et al (2016)	-0.4396	0.301786	23	22		Not estimable
Subtotal (95% CI)			**311**	**304**	**40.8%**	**-0.90 [-1.22, -0.58]**
Heterogeneity: Tau² = 0.19; Chi² = 32.86, df = 10 (P = 0.0003); I² = 70%						
Test for overall effect: Z = 5.57 (P < 0.00001)						
Total (95% CI)			**838**	**820**	**100.0%**	**-0.70 [-0.88, -0.53]**
Heterogeneity: Tau² = 0.12; Chi² = 67.57, df = 24 (P < 0.00001); I² = 64%						
Test for overall effect: Z = 7.84 (P < 0.00001)						
Test for subgroup differences: Chi² = 3.22, df = 1 (P = 0.07), I² = 68.9%						

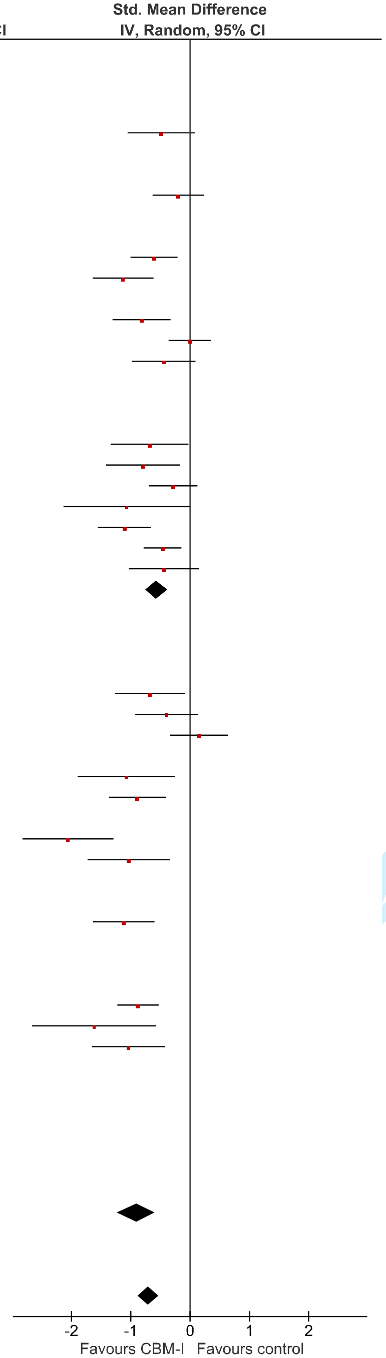

Figure S11: Forest plot of age group comparison for positive interpretations

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.2.1 Children						
Belli & Lau (2014)	-0.54	0.29704	24	23		Not estimable
Burnett Heyes et al (2017)	0.0649	0.2776	25	27		Not estimable
Chan, Reynolds & Lau (2015)	0.0932	0.24824	33	32		Not estimable
De Winter et al (2017)	-0.9307	0.300842	25	24	5.2%	-0.93 [-1.52, -0.34]
Fu, Du, Au & Lau (2013)	-0.512	0.39781	16	11		Not estimable
Fu, Du, Au & Lau (2015)	-0.5876	0.23908	37	36		Not estimable
Lau, Belli & Chopra (2013)	-1.7203	0.37012	20	20		Not estimable
Lau, Molyneaux et al (2011)	-0.7178	0.34439	17	19		Not estimable
Lothmann et al (2011)	-0.6995	0.2537	32	34		Not estimable
Orchard et al (2017)	-0.748	0.281	28	27	5.5%	-0.75 [-1.30, -0.20]
Salemink & Wiers (2011)	-0.4	0.17153	73	66		Not estimable
Telman et al (2013)	-1.1294	0.3175	23	23		Not estimable
Vassilopoulos & Brouzos (2016)	-0.2092	0.325765	20	18	4.8%	-0.21 [-0.85, 0.43]
Vassilopoulos et al (2009)	0.3635	0.30758	22	21	5.1%	0.36 [-0.24, 0.97]
Vassilopoulos, Blackwell et al (2014)	-0.7785	0.21559	53	41	6.7%	-0.78 [-1.20, -0.36]
Vassilopoulos, Brouzos & Andreau (2014)	-0.843	0.35847	16	18	4.3%	-0.84 [-1.55, -0.14]
Vassilopoulos, Moberly & Lau (2015)	-0.1796	0.21406	39	50	6.7%	-0.18 [-0.60, 0.24]
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4197	0.16347	77	76	7.7%	-0.42 [-0.74, -0.10]
Subtotal (95% CI)			**280**	**275**	**46.0%**	**-0.46 [-0.74, -0.19]**
Heterogeneity: Tau² = 0.08; Chi² = 16.36, df = 7 (P = 0.02); I² = 57%						
Test for overall effect: Z = 3.34 (P = 0.0008)						
4.2.2 Adolescents						
Belli & Lau (2014)	-0.54	0.29704	24	23	5.2%	-0.54 [-1.12, 0.04]
Burnett Heyes et al (2017)	0.0649	0.2776	25	27	5.6%	0.06 [-0.48, 0.61]
Chan, Reynolds & Lau (2015)	0.0932	0.24824	33	32	6.1%	0.09 [-0.39, 0.58]
De Winter et al (2017)	-0.9307	0.300842	25	24		Not estimable
Fu, Du, Au & Lau (2013)	-0.512	0.39781	16	11	3.8%	-0.51 [-1.29, 0.27]
Fu, Du, Au & Lau (2015)	-0.5876	0.23908	37	36	6.2%	-0.59 [-1.06, -0.12]
Lau, Belli & Chopra (2013)	-1.7203	0.37012	20	20	4.2%	-1.72 [-2.45, -0.99]
Lau, Molyneaux et al (2011)	-0.7178	0.34439	17	19	4.5%	-0.72 [-1.39, -0.04]
Lothmann et al (2011)	-0.6995	0.2537	32	34	6.0%	-0.70 [-1.20, -0.20]
Orchard et al (2017)	-0.748	0.281	28	27		Not estimable
Salemink & Wiers (2011)	-0.4	0.17153	73	66	7.6%	-0.40 [-0.74, -0.06]
Telman et al (2013)	-1.1294	0.3175	23	23	4.9%	-1.13 [-1.75, -0.51]
Vassilopoulos & Brouzos (2016)	-0.2092	0.325765	20	18		Not estimable
Vassilopoulos et al (2009)	0.3635	0.30758	22	21		Not estimable
Vassilopoulos, Blackwell et al (2014)	-0.7785	0.21559	53	41		Not estimable
Vassilopoulos, Brouzos & Andreau (2014)	-0.843	0.35847	16	18		Not estimable
Vassilopoulos, Moberly & Lau (2015)	-0.1796	0.21406	39	50		Not estimable
Vassilopoulos, Moberly & Zisimatou (2013)	-0.4197	0.16347	77	76		Not estimable
Subtotal (95% CI)			**300**	**291**	**54.0%**	**-0.58 [-0.87, -0.28]**
Heterogeneity: Tau² = 0.14; Chi² = 26.14, df = 9 (P = 0.002); I² = 66%						
Test for overall effect: Z = 3.84 (P = 0.0001)						
Total (95% CI)			**580**	**566**	**100.0%**	**-0.52 [-0.72, -0.32]**
Heterogeneity: Tau² = 0.10; Chi² = 42.69, df = 17 (P = 0.0005); I² = 60%						
Test for overall effect: Z = 5.22 (P < 0.00001)						
Test for subgroup differences: Chi² = 0.29, df = 1 (P = 0.59), I² = 0%						

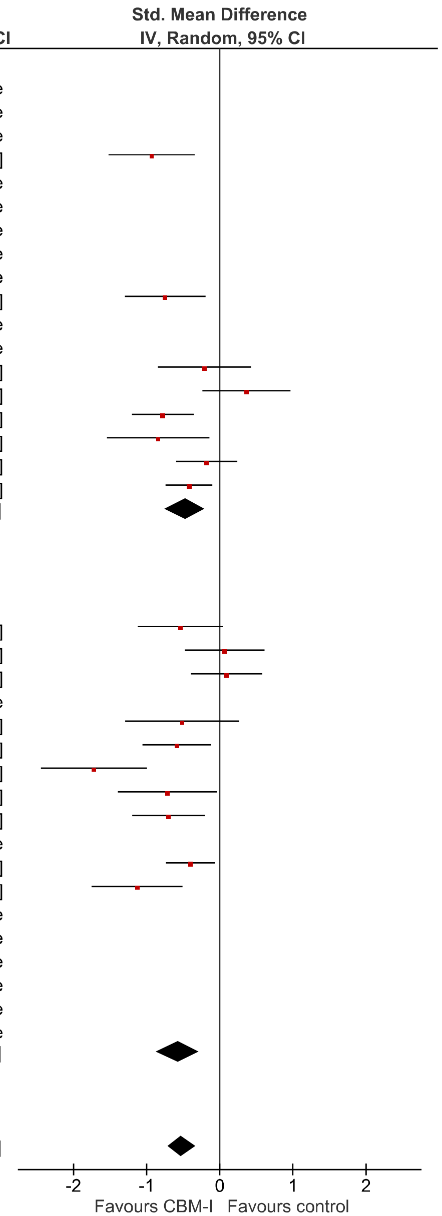

Figure S12: Forest plot of age group comparison for anxiety post-training

Study or Subgroup	Std. Mean Difference	SE	CBM-I Total	Control Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.7.1 Children						
Belli & Lau (2014)	-0.3061	0.29349	24	23		Not estimable
Burnett Heyes et al (2017)	-0.0545	0.13023	29	30		Not estimable
Fu, Du, Au & Lau (2013)	0.1196	0.3822	16	12		Not estimable
Fu, Du, Au & Lau (2015)	-0.298	0.23541	37	36		Not estimable
Lau, Belli & Chopra (2013)	0.2562	0.31753	20	20		Not estimable
Lau, Molyneaux et al (2011)	-0.4415	0.33786	17	19		Not estimable
Lester et al (2011a)	-0.0565	0.197117	51	52	7.3%	-0.06 [-0.44, 0.33]
Lester et al (2011b)	-0.5796	0.2494	34	33	5.6%	-0.58 [-1.07, -0.09]
Orchard et al (2017)	0.4449	0.271	29	27	5.0%	0.44 [-0.09, 0.98]
Salemink & Wiers (2011)	0.1264	0.16457	73	75		Not estimable
Vassilopoulos & Brouzos (2016)	-0.3688	0.327628	20	18	3.8%	-0.37 [-1.01, 0.27]
Vassilopoulos & Moberly (2013)	-0.4386	0.18908	61	54	7.6%	-0.44 [-0.81, -0.07]
Vassilopoulos et al (2009)	-1.0288	0.32462	22	21	3.9%	-1.03 [-1.67, -0.39]
Vassilopoulos, Blackwell et al (2014)	-0.1773	0.20839	53	41	6.9%	-0.18 [-0.59, 0.23]
Vassilopoulos, Moberly & Lau (2015)	-0.3549	0.21528	39	50	6.6%	-0.35 [-0.78, 0.07]
Vassilopoulos, Moberly & Zisimatou (2013)	0.0667	0.16173	77	76	8.8%	0.07 [-0.25, 0.38]
White et al (2016)	-0.1693	0.29875	23	22	4.4%	-0.17 [-0.75, 0.42]
Subtotal (95% CI)			**409**	**394**	**59.9%**	**-0.24 [-0.46, -0.02]**
Heterogeneity: Tau² = 0.06; Chi² = 20.12, df = 9 (P = 0.02); I² = 55%						
Test for overall effect: Z = 2.17 (P = 0.03)						
4.7.2 Adolescents						
Belli & Lau (2014)	-0.3061	0.29349	24	23	4.5%	-0.31 [-0.88, 0.27]
Burnett Heyes et al (2017)	-0.0545	0.13023	29	30	10.3%	-0.05 [-0.31, 0.20]
Fu, Du, Au & Lau (2013)	0.1196	0.3822	16	12	3.0%	0.12 [-0.63, 0.87]
Fu, Du, Au & Lau (2015)	-0.298	0.23541	37	36	6.0%	-0.30 [-0.76, 0.16]
Lau, Belli & Chopra (2013)	0.2562	0.31753	20	20	4.0%	0.26 [-0.37, 0.88]
Lau, Molyneaux et al (2011)	-0.4415	0.33786	17	19	3.7%	-0.44 [-1.10, 0.22]
Lester et al (2011a)	-0.0565	0.197117	51	52		Not estimable
Lester et al (2011b)	-0.5796	0.2494	34	33		Not estimable
Orchard et al (2017)	0.4449	0.271	29	27		Not estimable
Salemink & Wiers (2011)	0.1264	0.16457	73	75	8.6%	0.13 [-0.20, 0.45]
Vassilopoulos & Brouzos (2016)	-0.3688	0.327628	20	18		Not estimable
Vassilopoulos & Moberly (2013)	-0.4386	0.18908	61	54		Not estimable
Vassilopoulos et al (2009)	-1.0288	0.32462	22	21		Not estimable
Vassilopoulos, Blackwell et al (2014)	-0.1773	0.20839	53	41		Not estimable
Vassilopoulos, Moberly & Lau (2015)	-0.3549	0.21528	39	50		Not estimable
Vassilopoulos, Moberly & Zisimatou (2013)	0.0667	0.16173	77	76		Not estimable
White et al (2016)	-0.1693	0.29875	23	22		Not estimable
Subtotal (95% CI)			**216**	**215**	**40.1%**	**-0.05 [-0.21, 0.11]**
Heterogeneity: Tau² = 0.00; Chi² = 5.49, df = 6 (P = 0.48); I² = 0%						
Test for overall effect: Z = 0.65 (P = 0.52)						
Total (95% CI)			**625**	**609**	**100.0%**	**-0.17 [-0.31, -0.02]**
Heterogeneity: Tau² = 0.04; Chi² = 27.73, df = 16 (P = 0.03); I² = 42%						
Test for overall effect: Z = 2.25 (P = 0.02)						
Test for subgroup differences: Chi² = 1.85, df = 1 (P = 0.17), I² = 45.9%