Review article

Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis

Ioana A. Cristea, Robin N. Kok and Pim Cuijpers

Background
Cognitive bias modification (CBM) interventions are strongly advocated in research and clinical practice.

Aims
To examine the efficiency of CBM for clinically relevant outcomes, along with study quality, publication bias and potential moderators.

Method
We included randomised controlled trials (RCTs) of CBM interventions that reported clinically relevant outcomes assessed with standardised instruments.

Results
We identified 49 trials and grouped outcomes into anxiety and depression. Effect sizes were small considering all the samples, and mostly non-significant for patient samples. Effect sizes became non-significant when outliers were excluded and after adjustment for publication bias. The quality of the RCTs was suboptimal.

Conclusions
CBM may have small effects on mental health problems, but it is also very well possible that there are no significant clinically relevant effects. Research in this field is hampered by small and low-quality trials, and by risk of publication bias. Many positive outcomes are driven by extreme outliers.

Declaration of interest
None.

Cognitive bias modification (CBM) has been defined as the 'direct manipulation of a target cognitive bias, by extended exposure to task contingencies that favor predetermined patterns of processing selectivity'.[1] A recent review remarks that there has been 'exponential growth of research employing these cognitive bias modification (CBM) procedures',[1] especially in recent years. Research has focused mainly on two types of interventions[1]: attention bias modification (ABM) and interpretative bias modification (CBM-I). Another narrative review[2] concludes that both ABM and CBM-I can reliably have an impact on clinically relevant symptoms, with greatest confidence for anxiety symptoms and depression. The principle of ABM involves teaching participants to avoid the negative, 'threat' stimuli (usually pictures or words) by directing their attention, without their knowledge, to neutral or positive stimuli ('avoid threat'). The principle of CBM-I[1,3] is similar, but it uses more complex stimuli, such as ambiguous paragraphs or sentences. Participants are given a task that consistently disambiguates the valence of the sentence or paragraph (for example a word completion task), towards a positive (sometimes neutral) or a negative interpretation. Other interventions are also included under CBM, such as concreteness training (CNT)[4] or alcohol approach and avoidance training (A-AAT).[5]

Three previous meta-analyses examined the efficiency of various CBM approaches for psychological problems. The first[6] looked at studies of ABM for anxiety and found an effect size of 0.61. The second one[7] investigated both ABM and CBM-I for anxiety and depression and found a small, but significant post-intervention effect for anxiety and depression taken together ($g = 0.13$). Finally, the third meta-analysis[8] found non-significant effects for subjective experience following ABM. The present meta-analysis was prompted by a number of aspects that still remain unclear regarding the efficacy of CBM interventions. It is not clear if the previous meta-analyses were carried out on randomised controlled trials (RCTs). For instance, one of these[6] mentioned randomisation as an inclusion criterion, but out of 11 studies, two were experiments described in a narrative review, for which no independent, peer-reviewed publication as separate

RCTs existed.[3] Another[7] did not list randomisation as an inclusion criterion. Second, the quality of the studies included was not considered in any of the previous meta-analyses, even though there is ample evidence that the quality of the RCTs included in a meta-analysis can significantly bias outcomes,[9–13] especially for smaller trials.[12] Some CBM researchers themselves[14] have suggested that the field is overly reliant on small studies, the vast majority of which do not follow the Consolidated Standards of Reporting trials (CONSORT) guidelines.[15] Third, moderators of treatment effects that were considered were generally restricted to the purported mechanisms of action of CBM interventions (change in targeted bias), and to specific procedural details of the interventions (such as stimulus type, stimulus duration, stimulus orientation, number of trials). Finally, publication bias was analysed in earlier meta-analyses, but with quite different results. Our goal was to present an updated meta-analysis including all CBM interventions tested in RCTs for clinically relevant outcomes. A number of RCTs examining CBM interventions have been conducted since the three previous meta-analysis were published (we counted 18 new trials). We aimed to examine whether there is robust empirical evidence of strong methodological quality (i.e. reduced risk of bias) for the clinical efficacy of CBM interventions. CBM is ultimately advocated as a therapy, with the purpose of significantly reducing symptoms and distress, so its efficacy has to be clearly established in clinical samples. Finally, we wanted to analyse more general moderators of treatment response, relevant to most psychotherapies.

Method

Identification and selection of studies

We conducted a comprehensive literature search (see online supplement DS1) for the complete search string) in PubMed, PsycInfo, the Cochrane library and EMBASE through May 2013 using the following key words: "cognitive bias modification", "attention* bias modification", "attention" bias "training", "bias training", "interpret* bias modification". We checked the reference

sections of the three previous meta-analyses. We also periodically checked for newly published trials during the preparation of the meta-analysis until the end of September 2013.

We included studies in which: (a) participants were randomised and (b) the effect of a CBM intervention, alone or in combination with another treatment, was (c) compared to the effects of a control group, another active treatment or a combination of treatments, (d) in adults, (e) for clinically relevant outcomes, (f) assessed on established, standardised symptom or distress measures and (g) published in peer-reviewed scientific journals in English. We included studies that combined a CBM intervention with another active treatment (psychological or pharmacological), provided that there was a control condition for the CBM intervention (i.e. a group that received no CBM intervention, a no-contingency intervention, an intervention supposed to increase bias), whether alone or combined with another active treatment.

We considered active CBM interventions the ones designed to decrease bias, regardless of its type, and consequently that improved symptoms and mood. As relevant outcome measures, we excluded outcomes not related to clinical symptoms or distress, as well as outcomes that were not measured on established, standardised, instruments (for example, reaction times, biological data, data measured on Likert or visual analogue scales devised ad hoc). Dissertations and conference abstracts were not included.

Quality assessment and data extraction

The methodological quality of the included studies was assessed with five criteria of the risk of bias assessment tool, developed by the Cochrane Collaboration[16] to assess sources of bias in RCTs.

(a) Criterion 1: adequate generation of allocation sequence.

(b) Criterion 2: concealment of allocation to conditions.

(c) Criterion 3: prevention of knowledge of the allocated intervention to assessors of outcome.

(d) Criterion 4: prevention of knowledge of the allocated intervention to participants.

(e) Criterion 5: dealing with incomplete data.

Criterion 4 (masking of participants) was included because unlike other psychotherapy studies, in which it is impossible for participants to remain unaware of the allocated intervention, most CBM interventions are carried out without making participants aware of the contingency they are exposed to, the purpose of the intervention or, in some cases, even the fact they are being subjected to an intervention. In fact, lack of participant awareness of the training has been regarded as evidence of the implicit mechanism of action of these interventions, through modifying low-level biases, which are inaccessible to awareness.[1,14] Criterion 5 (dealing with incomplete data) was rated as positive if there were no missing data or if data were analysed in an intent-to-treat approach (meaning all randomised participants were included in the analysis). Risk of bias was rated by two independent researchers (I.C. and R.K.). Disagreements were discussed and if they remained unresolved, the senior author was consulted (P.C.).

We also coded several aspects of the included studies, as potential moderators.

(a) Type of sample: clinical – diagnosed using a structured clinical interview (such as Structured Clinical Interview for DSM Disorders (SCID)); subclinical/analogue – selected for high values on a scale for clinical symptoms or distress; unselected participants.

(b) Recruitment type: patient samples; community volunteers; university students.

(c) Delivery type: exclusively laboratory-based; including a home-based component.

(d) Participant compensation: yes (money, course credit, both); no (no type of compensation).

(e) Number of sessions.

(f) Publication year.

(g) Type of bias intervention: attentional (ABM); interpretational (CBM-I); other (CNT, A-AAT, combinations).

(h) Impact factor (from Web of Science) of the journal in which the study was published (at the time of publication).

This last moderator was chosen for exploratory analysis, since we noticed both CBM trials and CBM reviews have been published in top-tier journals in the field of clinical psychology. Other researchers have also noted the tremendous surge of interest in CBM interventions.[7,17] Top journals have dedicated special issues (for example *Journal of Abnormal Psychology*, February 2009; *Cognitive Therapy & Research*, April 2014) to CBM interventions. Although we acknowledge there are intrinsic problems with journal metrics such as the impact factor, it still can be seen as an indicator for the best articles in a given field. We wanted to examine whether there might be a trend of positive CBM results getting published in higher impact factor journals.

Meta-analysis

For each comparison between a CBM intervention (alone or in combination) and a comparison group, the effect size indicating the differences between the two groups at post-test was calculated (Cohen's *d* or standardised mean difference). The effect size were calculated by subtracting, at post-test, the mean score of the CBM group from the mean score of the comparison group, and dividing the result by the pooled standard deviation the two groups. According to Cohen,[18] effect sizes of 0.2 are considered small, whereas effect sizes of 0.5 are moderate and effect sizes of 0.8 are deemed large. Because a substantial proportion of studies had small sample sizes, we corrected the effect size for small sample bias, as recommended by Hedges & Olkin,[19] and reported the corrected indicator Hedges' *g*

We calculated and pooled the individual effect sizes with Comprehensive Meta-analysis (CMA; version 2.2.064 for Windows). As there was a lot of variability in the symptom outcomes considered and the instruments used to measure them, we grouped outcome measures into the following categories: anxiety – all (all anxiety outcomes, whether measured by disorder-specific or general anxiety instruments); general anxiety (all outcomes relating to general anxiety, such as the State-Trait Anxiety Inventory, excluding outcomes specific to various anxiety disorders); social anxiety; depression. If a study used more than one outcome from the same category or if the same outcome was measured by more than one instrument, an average effect size was computed. If means and standard deviations were not reported for the symptom outcomes in a study, we used the procedures recommended by CMA[20] to calculate the standardised mean difference from dichotomous data or from other statistics such as *t*-values or exact *P*-values. If the effect size could not be calculated, the study was excluded. For the studies that had more than one control group (i.e. sham/no contingency training, attend to threat or waitlist), we used only one control group to calculate effect size. We chose the group most similar to a placebo group (i.e. sham training), as this was the most common control condition in CBM studies. Also, this decision was made to reduce

effect size inflation attributable to non-specific effects of the intervention or to the fact that the control condition was supposed to achieve an opposite effect from the intervention.

We only reported effect sizes for post-test. Neither follow-up, nor post-challenge (i.e. after the confrontation with a stressor) data were considered because of considerable variability among studies. For follow-up, there was considerable variability regarding duration, and there was hardly any control as to whether participants underwent some other treatment in that period. For post-challenge data, the types of stressors used were extremely diverse (for example public speaking situation, contamination, a test), as were the outcome measures, which in most cases were not standardised (for example reaction times, *ad hoc* Likert scales).

We report results from two meta-analyses: one containing all participant samples, the other one including only studies in which participants were diagnosed with a clinical condition, by use of a standardised diagnostic interview (such as SCID, Anxiety Disorders Interview Schedule (ADIS)). To facilitate clinical interpretation, we also transformed the standardised mean difference into number needed to treat (NNT), using the formulae of Kraemer & Kupfer.[21] The NNT represents the number of patients that would have to be treated to generate one additional positive outcome.[22]

We expected considerable heterogeneity among studies and consequently decided to calculate mean effect sizes using a random effects model. To test for the homogeneity of effect sizes, we calculated the I^2 statistic, which indicates heterogeneity in percentages. A value of 0% indicates no observed heterogeneity, whereas as values over 0% refer to increasing heterogeneity, with 25% low, 50% moderate and 75% or above indicative of high heterogeneity.[23] We calculated 95% confidence intervals around I^2,[24] using the non-central χ^2-based approach with the heterogi module for Stata MP 13.1 for Mac.[25] We also calculated the Q-statistic, but only report whether it is significant. Outliers were defined as studies in which the 95% confidence interval was outside the 95% confidence interval of the pooled studies (on both sides of the confidence interval). Subgroup and meta-regression analyses were conducted with outliers removed.

Subgroup analyses were conducted using a mixed-effects model, in which studies within subgroups are pooled using the random-effects model, but tests for significant differences between subgroups are carried out using a fixed-effects model. Subgroups with fewer than three studies were not reported. For continuous moderator variables, we used meta-regression analyses to test whether there was a significant relationship between each of these variables and the effect size, and reported a Z-value and an associated P. We also conducted multivariate meta-regression analysis using the Stata program, both by including all moderators simultaneously, and by using a back-step procedure, removing the predictor with the highest P-value at each step.

Publication bias was assessed by visually inspecting the funnel plot for the main outcome categories and employing the trim and fill procedure of Duval & Tweedie[26] (as implemented in CMA, version 2.2.064), to obtain an estimate of the effect size after the publication bias has been taken into account. We also conducted Egger's test of the intercept to test the symmetry of the funnel plot.

Results

Selection and inclusion of studies

We examined a total of 738 records (356 after duplicates were removed) and excluded 264 based on inspection of the abstract. We retrieved the full text of the remaining 92 articles, totalling 97 trials. Figure 1 presents the flowchart of the inclusion process

and details the reasons for the exclusion of trials, following the PRISMA statement.[27] This process resulted in 44 published articles, with a total of 49 RCTs that met our inclusion criteria and were included in the meta-analysis.

Characteristics of included studies

We conducted the analyses only for post-test data, which totalled 52 comparisons from 49 RCTs (see online supplement DS2 for the complete list). Two RCTs[5,28] only reported follow-up symptom data. These studies were rated for descriptive characteristics, but did not contribute to post-test outcomes.

The number of intervention sessions ranged from 1 to 15, but 21 RCTs included only one session. Fifteen RCTs used unselected participants, 16 subclinical or analogue samples, 16 clinical mental-problem samples and 2 samples with a physical problem (pain). In total, 24 studies recruited participants from student samples, 15 from community volunteers, 9 patient samples and 1 was from general medical patients. Thirty-six comparisons were based on interventions carried out exclusively in the laboratory, while 13 also included a home component. Participants received compensation for participation (money, course credit or both) in 27 RCTs, and did not receive any compensation in 22 studies. There were 22 comparisons based on ABM interventions, whereas
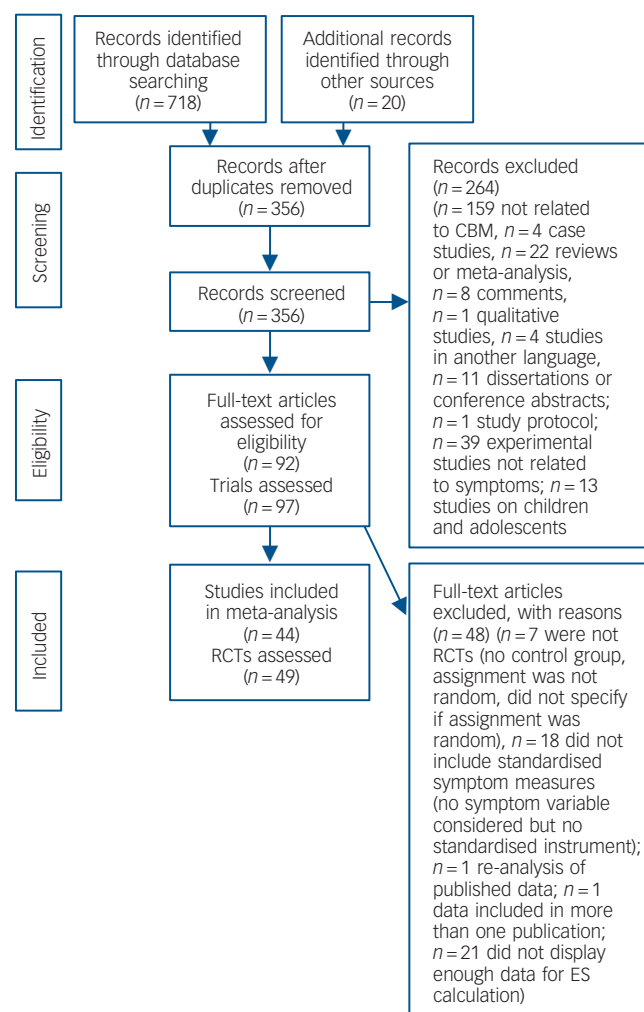

Fig. 1 Flow chart of selection and inclusion process, following the PRISMA statement.

CBM, cognitive bias modification; RCT, randomised controlled trials.

23 were based on CBM-I and 4 on other types of bias interventions (2 for CNT and 2 for A-AAT). Six RCTs were published in journals with an impact factor under 2, 24 in journals with an impact factor between 2 and 4, and 13 over 4 (impact factors for the other studies at the time of publication could not be retrieved). Online Table DS1 presents selected characteristics of the included RCTs.

Quality of the included studies

Overall, the quality of the included RCTs was not optimal. More than two-thirds of the studies (33/49) met fewer than three of the five quality criteria considered. One-fifth of the included RCTs (10/49) did not meet any quality criteria, and only 4% (5/49) met all criteria. Figure 2 presents the percentage of studies with a low, unclear (i.e. not enough information) and high risk of bias, for each of the quality criteria. It is worth mentioning that for all criteria except handling incomplete outcome data, a sizeable proportion of the RCTs (71% for sequence generation, 84% for allocation concealment, 57% for masking of assessors and 39% for masking of participants) did not provide the information necessary for assessing whether the criteria were met.

CBM compared with a control condition: all samples

Main effect sizes for all outcome categories

Anxiety (all measures). Figure 3 displays the forest plot of the standardised effect sizes of CBM interventions. The mean effect size was a g of 0.37 (95% CI 0.20–0.54) for 38 RCTs (41 comparisons). Heterogeneity was high ($I^2 = 72.66\%$) and highly significant (Table 1 and see online Table DS2 for a more detailed version of this table). Three studies were included in which two CBM interventions were compared with the same control group, meaning that multiple comparisons from these studies were not independent from each other. Their use in the same analyses could have affected the pooled effect size by leading to an artificial reduction of heterogeneity. We conducted sensitivity analyses including only one effect size per study to examine these possible effects, by first including only the comparisons with the largest effect sizes from each study and then only the ones with the smallest effect sizes. The resulting effect size and heterogeneity were very close to the ones found in the overall analysis. Three studies (four comparisons) were identified as outliers. With their removal, the effect size decreased to a g of 0.23 (95% CI 0.14–0.32) and heterogeneity became non-significant.

Anxiety (general). The mean effect size was a g of 0.38 (95% CI 0.17–0.59) for 31 RCTs (34 comparisons). Heterogeneity was high ($I^2 = 76.98\%$) and highly significant. Sensitivity analysis were conducted in the same way as for anxiety (all measures), with results remaining similar to the overall analysis. With the

exclusion of four outliers, the effect size decreased to 0.18 (95% CI 0.08–0.28) and heterogeneity became non-significant.

Social anxiety. Ten RCTs with ten comparisons resulted in a g of 0.40 (95% CI 0.06–0.74). Heterogeneity was high ($I^2 = 74.62\%$) and highly significant. Removal of one outlier reduced the effect size to a non-significant g of 0.23 (95% CI −0.001 to 0.46). Heterogeneity was no longer significant ($P > 0.05$).

Depression. We identified 17 RCTs with 17 comparisons, leading to a g of 0.43 (95% CI 0.16–0.71). Heterogeneity was high ($I^2 = 74.24\%$) and highly significant. With the removal of two outliers, the effect size decreased to a g of 0.33 (95% CI 0.16–0.50) and heterogeneity became non-significant.

Publication bias

We found significant publication bias for general anxiety, social anxiety and depression. For general anxiety, visual inspection and Egger's test indicated an asymmetric funnel plot (intercept 2.95, 95% CI 0.27–5.62, $P = 0.031$). The Duval & Tweedie trim and fill procedure did not impute any studies. For social anxiety inspection of the funnel plot (Fig. 4(b)) and the Duval & Tweedie trim and fill procedure indicated significant publication bias. After adjustment for missing studies ($n = 3$), the effect size decreased from a g of 0.40 to a non-significant g of 0.14 (95% CI −0.22 to 0.51). Egger's test was not significant (intercept 4.04, 95% CI −0.39 to 8.48, $P = 0.068$). For depression inspection of the funnel plot (Fig. 4(c)) and the Duval & Tweedie trim and fill procedure indicated significant publication bias. After adjustment for missing studies ($n = 7$), the effect size decreased from a g of 0.43 to a non-significant g of 0.09 (95% CI −0.21 to 0.39). Egger's test was not significant (intercept 3.25, 95% CI −0.61 to 7.12, $P = 0.09$).

Subgroup analyses

The results of subgroup analysis are shown in online Table DS2. For participant compensation, higher effect sizes were obtained if participants were compensated, than if they were not, for anxiety – all measures ($P = 0.028$), general anxiety ($P = 0.054$) and social anxiety ($P = 0.001$). For sample type, significantly higher effect sizes were obtained for subclinical or analogue samples than for clinical ones for social anxiety ($P = 0.025$). For recruitment, no significant differences were found. For delivery type, we found significantly higher effect sizes if the interventions were delivered exclusively in the laboratory, as opposed to them including a home-based component for anxiety – all measures ($P = 0.030$), general anxiety ($P = 0.006$) and social anxiety ($P = 0.001$). For type of bias, we found significantly higher effect
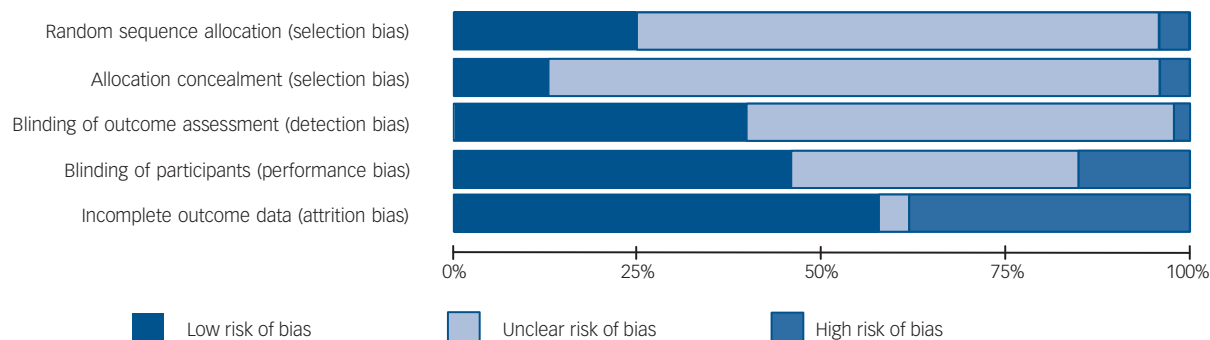

Fig. 2 Risk of bias graph: authors' judgements about each risk of bias item presented as percentages across all included studies.

Statistics for each study — Hedges' *g* and 95% CI

Study name	Hedges' *g*	Standard error	Lower limit	Upper limit	Z-Value	P-Value
Amir et al (2008)[29]	0.168	0.203	−0.230	0.566	0.829	0.407
Amir et al (2009)[30]	0.662	0.396	−0.113	1.438	1.674	0.094
Amir et al (2010)[31]	−0.076	0.261	−0.589	0.436	−0.292	0.771
Amir et al (2011)[32]	0.606	0.192	0.230	0.983	3.157	0.002
Baert et al (2010)[33] S1	−0.218	0.285	−0.777	0.341	−0.764	0.445
Baert et al (2010)[33] S2	0.246	0.336	−0.411	0.904	0.734	0.463
Beard & Amir (2008)[34]	0.656	0.385	−0.098	1.411	1.705	0.088
Boettcher et al (2013)[35]	0.008	0.219	−0.421	0.436	0.035	0.972
Bowler et al (2012)[36]	0.553	0.309	−0.053	1.159	1.787	0.074
Carlbring et al (2012)[37]	−0.060	0.223	−0.497	0.377	−0.271	0.786
Harris & Menzies (1998)[38]	−0.096	0.296	−0.677	0.484	−0.326	0.745
Hazen et al (2009)[39]	0.607	0.412	−0.201	1.415	1.472	0.141
Heeren et al (2012)[40]	0.060	0.318	−0.563	0.684	0.190	0.850
Hirsch et al (2007)[41]	0.535	0.402	−0.253	1.322	1.331	0.183
Holmes & Mathews (2005)[42] S2, C1	0.338	0.405	−0.457	1.133	0.834	0.404
Holmes & Mathews (2005)[42] S2, C2	−0.197	0.432	−1.043	0.649	−0.456	0.649
Hoppitt et al (2010)[43] C1	0.236	0.288	−0.329	0.801	0.818	0.413
Hoppitt et al (2010)[43] C2	0.091	0.282	−0.461	0.643	0.323	0.746
Lang et al (2012)[44]	0.280	0.382	−0.469	1.028	0.733	0.464
Lester et al (2011)[45] S1	2.213	0.325	1.576	2.850	6.805	0.000
Lester et al (2011)[45] S2	2.404	0.312	1.793	3.015	7.712	0.000
Li et al (2008)[46]	0.530	0.422	−0.296	1.356	1.258	0.209
Mackintosh et al (2006)[47] S2	0.755	0.338	0.093	1.418	2.236	0.025
Mathews et al (2007)[48]	0.430	0.318	−0.192	1.053	1.356	0.175
Murphy et al (2007)[49] C1	0.644	0.304	0.049	1.240	2.120	0.034
Murphy et al (2007)[49] C2	0.331	0.298	−0.253	0.916	1.112	0.266
Najmi & Amir (2010)[50]	0.137	0.273	−0.399	0.673	0.499	0.617
Neubauer et al (2013)[51]	−0.009	0.257	−0.513	0.495	−0.034	0.973
Rapee et al (2013)[52]	−0.134	0.202	−0.530	0.262	−0.663	0.508
Salemink et al (2007)[53]	0.351	0.184	−0.011	0.712	1.901	0.057
Salemink et al (2009)[54]	0.333	0.338	−0.328	0.995	0.987	0.324
Salemink et al (2007)[55]	0.237	0.221	−0.197	0.671	1.072	0.284
Schmidt et al (2009)[56]	2.365	0.455	1.474	3.256	5.202	0.000
Schoorl et al (2013)[57]	0.142	0.197	−0.244	0.529	0.722	0.470
Sharpe et al (2012)[58] S2	−0.363	0.404	−1.154	0.428	−0.900	0.368
Steel et al (2010)[59]	−0.262	0.214	−0.681	0.157	−1.227	0.220
Steinman & Teachman (2010)[60]	0.164	0.280	−0.385	0.713	0.585	0.558
Watkins et al (2012)[61]	0.747	0.245	0.267	1.228	3.049	0.002
Yiend et al (2005)[62] S1	0.722	0.443	−0.146	1.591	1.630	0.103
Yiend et al (2005)[62] S2	0.061	0.394	−0.712	0.834	0.155	0.877
Yiend et al (2005)[62] S3	0.222	0.440	−0.641	1.085	0.504	0.614
	0.374	0.087	0.203	0.545	4.285	0.000

−3.00 −1.50 0.00 1.50 3.00
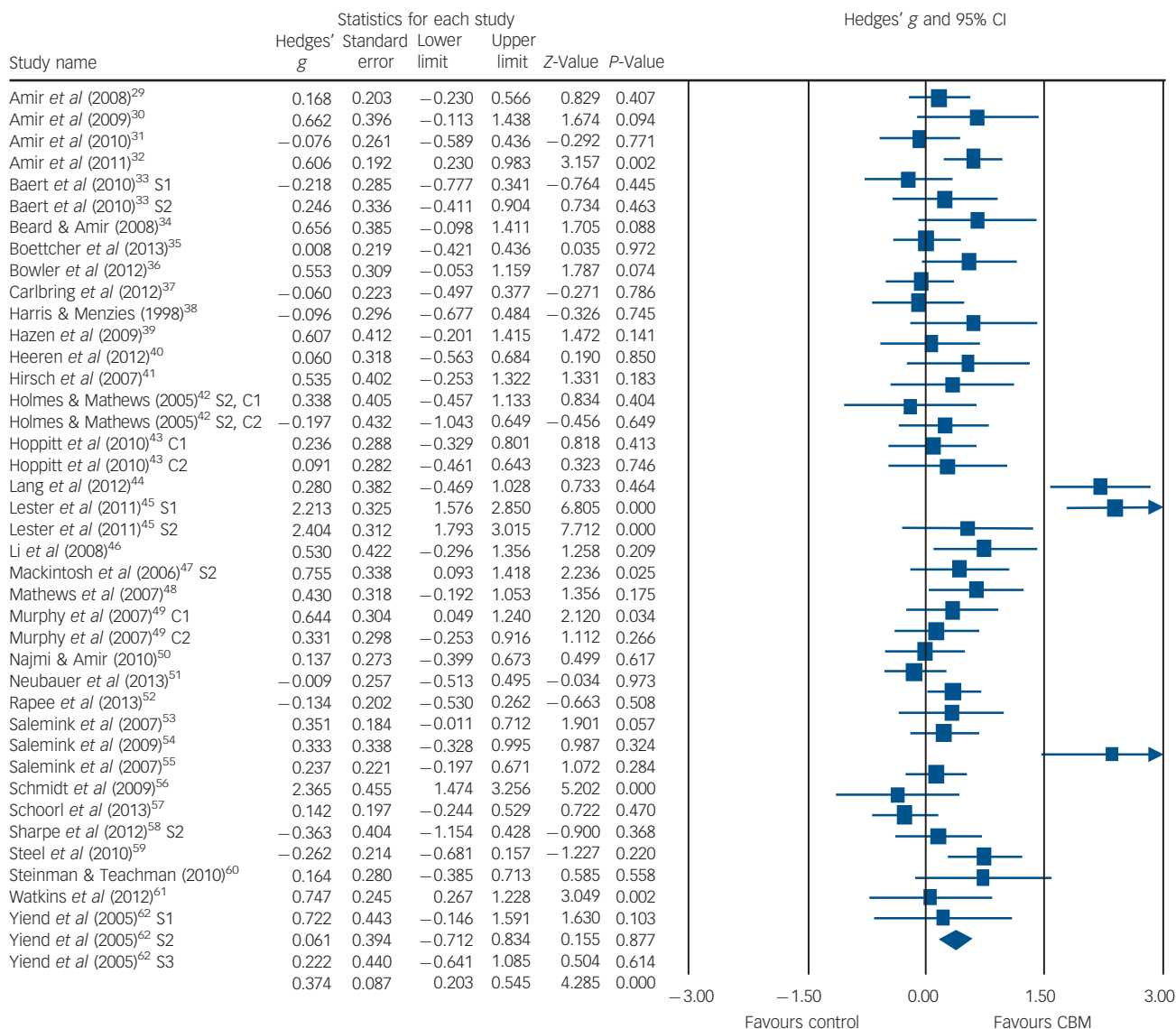Favours control — Favours CBM

Fig. 3 Standardised effect sizes of cognitive bias modification (CBM) interventions for all the samples for anxiety (all measures). S, study; C, comparison.

Table 1 Effects of cognitive bias modification interventions, compared with control, at post-test, for all samples and outcome categories[a]

Variable	n_comp	g (95% CI)	Z	I² (95% CI)[b]	Number needed to treat
Anxiety (all measures)	41	0.37 (0.20 to 0.54)	4.28	73* (63–80)	4.72
One effect size per study (only highest)	38	0.39 (0.21 to 0.57)	4.26	74* (65–81)	4.59
One effect size per study (only lowest)	38	0.37 (0.19 to 0.55)	4.00	75* (65–81)	4.72
Outliers removed[c]	37	0.23 (0.14 to 0.32)	4.96	2 (0–38)	7.69
General anxiety	34	0.38 (0.17 to 0.59)	3.58	77* (68–83)	4.59
One effect size per study (only highest)	31	0.41 (0.18 to 0.63)	3.56	78* (70–85)	4.27
One effect size per study (only lowest)	31	0.38 (0.15 to 0.61)	3.30	78* (71–85)	4.59
Outliers removed[c]	30	0.18 (0.08 to 0.28)	3.46	0 (0–41)	9.80
Social anxiety	10	0.40 (0.06 to 0.74)	2.34	75* (53–86)	4.39
Outlier removed[d]	9	0.23 (−0.001 to 0.46)	1.95	44 (0–75)	7.69
Generalised anxiety	3	0.68 (0.31 to 1.05)	3.64	0 (0–90)	2.63
Panic symptoms	4	0.02 (−0.43 to 0.44)	0.10	51 (0–84)	83.33
Depression	17	0.43 (0.16 to 0.71)	3.17	74* (59–84)	4.1
Outliers removed[e]	15	0.33 (0.16 to 0.50)	3.82	27 (0–61)	5.43

n_comp, number of comparisons.
a. All results are reported with Hedges g, using a random-effects model.
b. The P levels in this column indicate whether the Q-statistic is significant (the I² statistic does not include a test of significance).
c. Outliers were defined as studies in which the 95% CI was outside the 95% CI of the pooled studies. Below the 95% CI: Steel et al;[59] above the 95% CI: Lester et al[45] Study 1 and Study 2, Schmidt et al.[56]
d. Above the 95% CI: Schmidt et al.[56]
e. Below the 95% CI: Baert et al,[33] Study 1; above the 95% CI: Lester et al[45] Study 1 and Study 2.
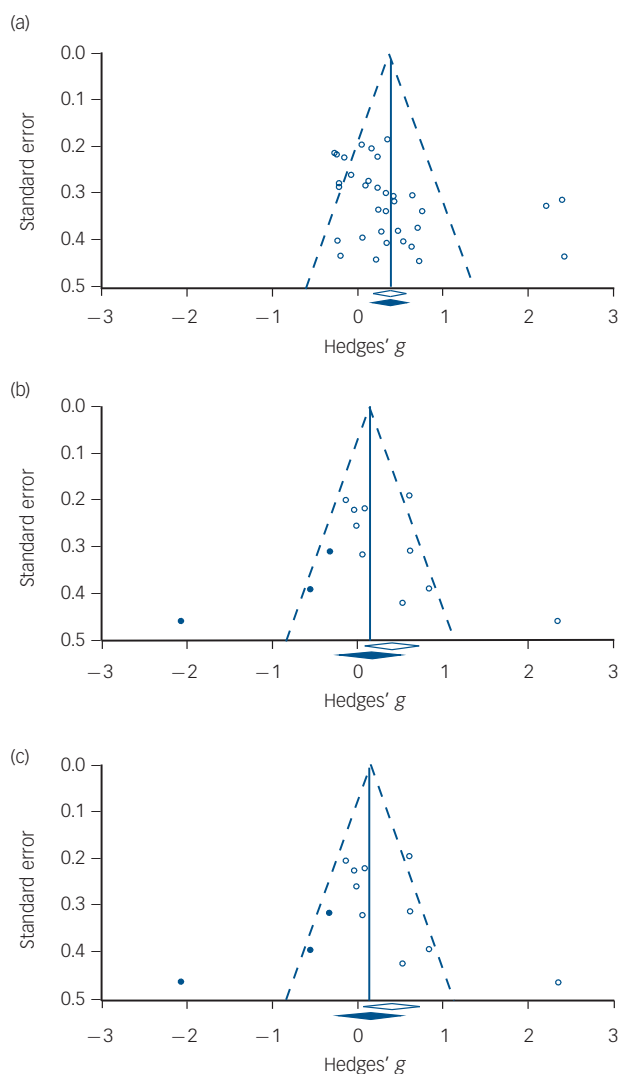*P<0.05; other results are not significant (P>0.05).

(a)

(b)

(c)

sizes for CBM-I than ABM for anxiety – all measures ($P = 0.048$), general anxiety ($P = 0.034$) and depression ($P = 0.014$).

For publication year, meta-regression indicated a significant, negative relationship between publication year and effect size for general anxiety (slope $b = -0.06$, 95% CI -0.10 to -0.02, $P = 0.003$), social anxiety (slope $b = -0.17$, 95% CI -0.29 to -0.05, $P = 0.003$), and depression, (slope $b = -0.10$, 95% CI -0.19 to -0.02, $P = 0.01$). For general anxiety, a multivariate meta-regression using a back-step procedure found publication year as significant (slope $b = -0.06$, 95% CI -0.10 to -0.01, $P = 0.006$). For number of sessions meta-regression indicated a significant, negative relationship between number of sessions and effect sizes for general anxiety (slope $b = -0.02$, 95% CI -0.05 to -0.0006, $P = 0.044$) and depression (slope $b = -0.05$, 95% CI -0.09 to -0.01, $P = 0.007$). In terms of quality (low risk of bias) we found a significant negative relationship between the quality score (number of criteria met) and effect sizes for general anxiety (slope $b = -0.07$, 95% CI -0.13 to -0.02, $P = 0.007$) and depression (slope $b = -0.10$, 95% CI -0.20 to -0.006, $P = 0.036$). We found a borderline significant relationship between

the journal impact factor and effect size (slope $b = 0.08$, 95% CI -0.001 to 0.16, $P = 0.053$) for anxiety (all measures).

CBM compared with a control condition: clinical samples only

Main effect sizes for all outcome categories

Anxiety (all measures). Thirteen RCTs, with 13 comparisons aggregated to a g of 0.28 (95% CI 0.01–0.55). Heterogeneity was high ($I^2 = 72.76\%$) and highly significant (Table 2). Removal of one outlier with an extremely high effect size ($g = 2.36$) reduced the effect size by approximately half, to a non-significant g of 0.16 (95% CI -0.03 to 0.35). Heterogeneity remained significant, but moderate ($I^2 = 47.44\%$).

General anxiety. Eight RCTs with eight comparisons led to a non-significant g of 0.29 (95% CI -0.14 to 0.74). Heterogeneity was high ($I^2 = 81.75\%$) and highly significant. Removal of one outlier let to a g of -0.01 (95% CI -0.22 to 0.20) and heterogeneity was no longer significant.

Social anxiety. Seven RCTs with seven comparisons aggregated into a non-significant g of 0.32 (95% CI –0.09 to 0.74), with high heterogeneity ($I^2 = 80.84\%$, Table 2). Removal of one outlier led to a g of 0.11 (95% CI -0.13 to 0.35) and heterogeneity was no longer significant.

Depression. Nine RCTs with nine comparisons aggregated into a g of 0.24 (95% CI 0.02 to 0.46), with non-significant heterogeneity.

Publication bias

We found evidence of publication bias for general anxiety and depression. In both cases inspection of the funnel plot and the Duval & Tweedie trim and fill procedure indicated significant publication bias. For general anxiety, after adjustment for missing studies ($n = 3$), the effect size decreased from a g of 0.29 to a non-significant g of -0.09 (95% CI -0.58 to 0.40). Egger's test was also significant (intercept 6.23, 95% CI 1.32–11.14, $P = 0.020$). For depression, after adjustment for missing studies ($n = 4$), the effect size decreased from a g of 0.24 to a non-significant g of 0.04 (95% CI -0.20 to 0.29). Egger's test was also significant (intercept 3.50, 95% CI 0.69–6.31, $P = 0.021$). Moreover, for depression a meta-regression analysis indicated a significant relationship between publication year and effect size (slope $b = -0.12$, 95% CI -0.24 to -0.008, $P = 0.036$).

Subgroup and mediation analysis

As there were few studies in each outcome category, we did not conduct subgroup analysis. We inspected which of the included studies in our meta-analysis had also conducted formal mediation analysis (i.e. not just correlations) to test whether the effects of the intervention on emotional outcomes were mediated by changes in bias. Only 11 out of the 49 RCTs had conducted formal tests of mediation. Only four studies[28,31,63,64] reported clear evidence of successful mediation. Three other studies found evidence of mediation for only one, but not for the other outcomes[30,40,50] and in two of these, the outcome was physiological[40] or behavioural[50] and not the emotional or symptom outcome our analysis focused on. Two studies[34,36] found evidence of mediation for change in one type of bias, but not in another, and in one of these there was mediation for both the CBM and the computerised CBT control group.[36] Finally, two studies found

Table 2 Effects of cognitive bias modification interventions, compared with control, at post-test, for clinical samples[a]

Variable	n_{comp}	g (95% CI)	Z	I^2 (95% CI)[b]	Number needed to treat
Anxiety (all measures)	13	0.28 (0.01 to 0.55)	2.06	73* (53–84)	6.41
Outlier removed[c]	12	0.16 (−0.03 to 0.35)	1.64	47* (0–73)	11.11
General anxiety	8	0.29 (−0.14 to 0.74)	1.32	82* (65–90)	5.95
Outlier removed[c]	7	−0.01 (−0.22 to 0.20)	−0.09	23 (0–66)	–
Social anxiety	7	0.32 (−0.09 to 0.74)	1.51	81*(61–91)	5.56
Outlier removed[c]	6	0.11 (−0.13 to 0.35)	0.89	43* (0–78)	16.13
Depression	9	0.24 (0.02 to 0.46)	2.19	39 (0–72)	7.46

n_{comp}, number of comparisons.
a. All results are reported with Hedges g, using a random-effects model.
b. The P levels in this column indicate whether the Q-statistic is significant (the I^2 statistic does not include a test of significance).
c. Outliers were defined as studies in which the 95% CI was outside the 95% CI of the pooled studies. Above the 95% CI: Schmidt et al.[56]
*P<0.05; other results are not significant (P>0.05).

no evidence of mediation of changes in symptom or emotional outcomes by change in bias.[5,65]

Discussion

Summary of main findings

We conducted an updated meta-analysis including all CBM interventions tested in RCTs for clinically relevant outcomes. We also evaluated, for the first time, the quality of the CBM studies by assessing the presence of risk of bias. CBM interventions are ultimately advocated as therapeutic, so we also separately examined their effects for clinical patients, not just for the blend of healthy, subclinical and clinical participants that were the focus of the three previous meta-analyses. Finally, we wanted to analyse more general moderators of treatment response, relevant to most psychotherapeutic interventions, and to examine more closely publication bias and its possible ramifications (for example a possible effect of time-lag bias).

Overall, for the meta-analysis including all the samples, the effects of CBM intervention in all outcome categories were small and showed a high degree of heterogeneity. Exclusion of outliers significantly reduced effect sizes across all outcome categories, in some cases by almost half. We note that for anxiety outcomes, three of the four outliers identified were not only outside the confidence intervals of the effect size, but had effect sizes almost ten times higher than the pooled mean effect size. One of these, Schmidt et al, 2009,[56] was identified as an outlier in two of the three previous meta-analysis.[6,7] Ironically, this was also one of the first articles that marked the ABM research and practice boom. Adjustment for publication bias also reduced effect sizes considerably, and for some outcomes rendered them non-significant.

The only other meta-analysis[7] that looked at both ABM and CBM-I interventions found even smaller, yet significant, results for both anxiety and depression, respectively, but no evidence of heterogeneity. We note that many new studies have been published over the past 3 years, since the search conducted for their meta-analysis. Whereas they included any controlled clinical trials, we restricted our meta-analysis to RCTs examining any CBM intervention, which used an established, standardised measure of symptom or distress-related outcomes.

For clinical samples, the effects of CBM interventions on anxiety and depression outcomes were small and in most cases non-significant; in the cases where they were significant, such as for depression, it seems to have been as a result of the presence of outliers and/or publication bias. The three previous meta-analysis[6–8] reported either non-significant differences or a trend towards higher effects for clinical or high symptomatology samples as compared with unselected ones. Including only RCTs and looking exclusively at participants with a clinical diagnosis, our data portray a different picture.

The effects of CBM interventions on types of bias was not specifically approached in this meta-analysis, as it was analysed in all three previous meta-analyses. In the studies that measured bias, this was generally considered just another outcome measure, alongside emotional and symptom outcomes. Even if mediation analysis was conducted, it could not reliably indicate that changes in bias were causally related to changes in symptoms, because they were both measured at the same time point. Measures of bias were heterogeneous, employing many different indices and procedure, not standardised and consequently difficult to interpret. Moreover, in many cases the task used to measure bias was the same one used in the intervention (for example the dot probe task), making it susceptible to the biasing effect of demand characteristics. Our examination of whether the studies included examined if the effects of the CBM interventions of emotional and symptom outcomes was mediated by changes in biases provided results at best mixed, reflecting the difficulties we highlighted above. Thus, it is evident we cannot even be sure which are the mechanisms of change in CBM.

Implications

Given that CBM interventions are primarily intended as cost-effective therapeutic alternatives,[66] alleged to have an impact on clinically relevant symptoms[1], we believe our results cast serious doubts on the majority of them having strong clinical utility. In contrast, effect sizes for other forms of psychotherapy have proven much more robust and of a greater magnitude, even when compared with a placebo condition. For instance, for adult depression a meta-analysis[67] found a g of 0.51 (corresponding to an NNT of 3.55) for cognitive–behavioural therapy as compared with placebo.

Leading CBM researchers[1] have argued that demand characteristics are an implausible explanation for CBM findings. Yet the results of our meta-analysis point in a different direction. Across anxiety outcome categories, we found that effect sizes were higher if participants received compensation for participation than if they did not, and if the intervention was delivered exclusively in the laboratory as opposed to also including a home-based component. Moreover, contrary to previous meta-analyses, our results showed that the effect sizes for general anxiety and depression were negatively linearly related with the number of sessions. Although the slope for this relationship is quite small, albeit significant, and we noted that a considerable proportion of RCTs included only one session, we can at least conclude that

the lack of reliable effects of CBM is not because of insufficient exposure to the intervention.

Put together, these results seem to indicate that it is not unlikely that many positive CBM findings may have been influenced by a variant of the 'experimenter effect'[67] or other experimental artefacts, unrelated to the scope and purported mechanisms of action of these interventions. In support of these conjectures comes the effect, present in many CBM studies, that the placebo control group ('no contingency') also shows improvement.[33–35,37,44,51,68] Surprisingly, it has been recently argued that at least for ABM in the case of social anxiety, the opposite intervention, assumed to increase bias and used as a control condition in previous studies, would be more beneficial.[35,69]

As importantly, we found strong evidence of publication bias, both for the all samples and the clinical samples meta-analyses. For the former, there was considerable publication bias for general anxiety, social anxiety and depression. In fact, for social anxiety and depression, adjustment for publication bias rendered effect sizes no longer significant. The same pattern emerged for the clinical samples meta-analysis. There was strong and consistent evidence of publication bias for general anxiety and depression.

We also found a strong and consistent negative linear relationship between publication year and effect size across most outcome categories: older studies showed significantly higher effect sizes, whereas more recent studies obtained effect sizes that were non-significant, very close to zero and, in some cases, even negative. This is a common phenomenon in intervention research, when a new intervention is proposed and tested, with the first studies showing large effects because of methodological considerations (i.e. the use of pilot, low powered studies where only large effects can overcome the significance threshold) and a strong publication bias for positive findings. In fact, promising new interventions like CBM are very susceptible to a particular type of publication bias –time lag bias – the phenomenon in which studies with positive results get to be published first and dominate the field, until the negative, but equally important, studies are published[70,71] – if they are published at all. Nonetheless, we suspect this phenomenon was aggravated for CBM by highly laudatory narrative reviews, comments and editorials, published before the efficiency of the new interventions had been established in well-powered, methodologically appropriate RCTs. In addition, if replication studies are based on high effect sizes found in early studies, there is a distinct possibility that these subsequent replication studies are powered to detect a large effect size and thus are severely underpowered.[72] Moreover, overtly positive pieces about CBM have been almost exclusively published in top-tier journals, thus contributing to indirectly enforcing the notion that we were witnessing the development of a powerful new therapy – 'a new clinical weapon'.[73]

Interestingly, CBM-I seems to have better results than ABM. Although this result seems to suggest differential potential, we need to be very cautious in its interpretation. Studies on CBM-I are a newer development and we run the risk of having caught them exactly in the phase where mostly positive results are getting published (i.e. time-lag bias). Also, by their nature, CBM-I interventions are more susceptible to demand characteristics because it is easier for participants to catch on to a particular interpretative pattern being favoured.

The view that the praise for CBM is much ahead of its corresponding empirical evidence is further sustained by results regarding the quality of the studies. Using a widely accepted and recommended instrument – the Cochrane Collaboration's Risk of bias assessment tool – we showed that the quality of the studies investigating CBM is substandard. In fact, a significant proportion of the included studies satisfied no quality criteria at all and about two-thirds satisfied fewer than three quality criteria. Moreover, it is even more problematic that for three of the five risk of bias criteria analysed (sequence generation, allocation concealment, masking of assessors), a significant proportion of the studies (ranging from 58 to 83%) was unclear or simply did not contain information to permit assessment. Even for the masking of participants, a characteristic asserted as a strong point of these studies,[14] 40% were rated as unclear.

Risk of bias can also be associated with artificial inflation of effect sizes. Indeed, our results showed that for both general anxiety and depression, the quality of the included RCTs was negatively related to effect size. The effect size decreased by 0.07 with every quality criterion that was satisfied for general anxiety and by 0.10 for depression.

Limitations

Given that CBM interventions include a number of different approaches, under different names and with many task variations, we might have missed some studies that would have been eligible. We tried to be inclusive with our search criteria and given that many CBM studies present themselves as experimental studies and not RCTs, we conducted our search without restricting it to RCTs. Statistical power might have also represented a problem, especially for some subgroup analysis and it might have prevented us from finding significant differences. There was a very high degree of heterogeneity for all the outcome categories considered and one might even wonder if it makes sense to combine these studies at all.[24] However, we underscore that the exclusion of outliers significantly reduced heterogeneity to non-significant values for all outcome categories considered. Nonetheless, confidence intervals around I^2 remained large, indicating heterogeneity was most likely still present. This is probably also as a result of the fact that CBM studies used a wide range of outcome measurements, which we were only able to group loosely into categories. Few studies declared a primary outcome measure.

Future directions

In conclusion, our updated meta-analysis of all types of CBM interventions examined in RCTs showed small effects for anxiety and depression outcomes and significant heterogeneity, when all participant samples were considered. The effects were marked by significant outliers and publication bias and were largely reduced when these factors were adjusted. For clinical patients, we showed small and mostly non-significant effects for anxiety and depression outcomes. Adjustment for publication bias rendered results for all outcome categories non-significant for clinical patients. The quality of the included RCTs was suboptimal and higher-quality studies obtained smaller, closer to zero, effect sizes. More sessions were associated with smaller effect sizes, as were the absence of participant compensation and the non-exclusively laboratory-based delivery of the intervention. Along with strong evidence of publication bias, publication year was a robust predictor of effect size, across almost all outcome categories: the more recent the study, the smaller and closer to zero its corresponding effect size. Our results highlight considerable problems with CBM interventions, underscoring their lack of clinically relevant effects for patients, as well as the genuine possibility of substantial artificial effect size inflation because of aspects unrelated to the interventions themselves, but to demand characteristics and publication bias.

Even if, at least recently, CBM researchers have started acknowledging the ever more frequent negative results and their significance for CBM interventions, the search for answers

continues to be confined by the strict boundaries of the paradigm, with researchers constantly 'trying out' new variations of CBM designs, tasks, instructions, doses (i.e. number of sessions) or moderating variables to attempt to validate a theoretical framework. We argue that this approach is detrimental, as it hinders the development of well-conducted, independently run RCTs. Unlike other forms of psychotherapy, not only is there no established protocol for CBM interventions, but it becomes difficult to choose from the wide variety of task variations and dosages proposed.

We believe that the only way of creating a more robust CBM is by replacing the myriad of experimental variations with proper clinical research (i.e. on clinical patients with a randomised trial design, a pre-specified, reproducible protocol, sufficient power) on the most promising CBM approaches (for instance AAT for alcohol problems). One CBM approach that has done that is CNT,[74,75] which now requires independent validation from other investigator groups in order to be considered an evidence-based treatment for depression. Only if such studies are conducted and result in clinically relevant outcomes, is there a future for CBM.

Ioana A. Cristea, PhD, Department of Clinical Psychology and Psychotherapy, Babeş-Bolyai University, Cluj-Napoca, Romania and Clinical Psychology Branch, Department of Surgical, Medical, Molecular and Critical Pathology, University of Pisa, Pisa, Italy; **Robin N. Kok**, MSc, Department of Clinical Psychology and the EMGO institute for Health and Care Research, Faculty of Psychology and Education, VU University Amsterdam, Amsterdam, The Netherlands and Centre for Mental Health Research, The Australian National University, Acton, Australia; **Pim Cuijpers**, PhD, Department of Clinical Psychology, VU University and EMGO Institute for Health and Care Research, VU University and VU University Medical Centre, Amsterdam, The Netherlands and Leuphana University, Lüneburg, Germany.

Correspondence: Ioana A. Cristea, Department of Clinical Psychology and Psychotherapy, Babeş-Bolyai University, Republicii Street 37, 400015, Cluj-Napoca, Romania. Email: ioana.cristea@ubbcluj.ro

First received 18 Feb 2014, final revision 28 Apr 2014, accepted 6 Jun 2014

References

1 MacLeod C, Mathews A. Cognitive bias modification approaches to anxiety. *Annu Rev Clin Psychol* 2012; **8**: 189–217.

2 Macleod C. Cognitive bias modification procedures in the management of mental disorders. *Curr Opin Psychiatry* 2012; **25**: 114–20.

3 Mathews A, MacLeod C. Induced processing biases have causal effects on anxiety. *Cogn Emot* 2002; **16**: 331–54.

4 Watkins ER, Baeyens CB, Read R. Concreteness training reduces dysphoria: proof-of-principle for repeated cognitive bias modification in depression. *J Abnorm Psychol* 2009; **118**: 55–64.

5 Wiers RW, Eberl C, Rinck M, Becker ES, Lindenmeyer J. Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol Sci* 2011; **22**: 490–7.

6 Hakamata Y, Lissek S, Bar-Haim Y, Britton JC, Fox NA, Leibenluft E, et al. Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. *Biol Psychiatry* 2010; **68**: 982–90.

7 Hallion LS, Ruscio AM. A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychol Bull* 2011; **137**: 940–58.

8 Beard C, Sawyer AT, Hofmann SG. Efficacy of attention bias modification using threat and appetitive stimuli: a meta-analytic review. *Behav Ther* 2012; **43**: 724–40.

9 Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008; **336**: 601–5.

10 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273**: 408–12.

11 Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998; **352**: 609–13.

12 Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2001; **135**: 982–9.

13 Cuijpers P, Smit F, Bohlmeijer E, Hollon SD, Andersson G. Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. *Br J Psychiatry* 2010; **196**: 173–8.

14 Beard C. Cognitive bias modification for anxiety: current evidence and future directions. *Expert Rev Neurother* 2011; **11**: 299–311.

15 Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010; **8**: 18.

16 Higgins JPT, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**: d5928.

17 Emmelkamp PMG. Attention bias modification: the Emperor's new suit? *BMC Med* 2012; **10**: 63.

18 Cohen J. *Statistical Power Analysis for the Behavioral Sciences (2nd edn).* Lawrence Erlbaum Associates, 1988.

19 Hedges LV, Olkin I. *Statistical Methods for Meta-Analysis.* Academic Press, 1985.

20 Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis.* Wiley, 2009.

21 Kraemer HC, Kupfer DJ. Size of treatment effects and their importance to clinical research and practice. *Biol Psychiatry* 2006; **59**: 990–6.

22 Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988; **318**: 1728–33.

23 Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557–60.

24 Ioannidis JPA, Patsopoulos NA, Evangelou E. Uncertainty in heterogeneity estimates in meta-analyses. *BMJ* 2007; **335**: 914–6.

25 Orsini N, Bottai M, Higgins J, Buchan I. *HETEROGI: Stata module to quantify heterogeneity in a meta-analysis.* EconPapers, 2006 (http://econpapers.repec.org/software/bocbocode/s449201.htm).

26 Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2000; **56**: 455–63.

27 Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med* 2009; **6**: e1000097.

28 Eberl C, Wiers RW, Pawelczack S, Rinck M, Becker ES, Lindenmeyer J. Approach bias modification in alcohol dependence: do clinical effects replicate and for whom does it work best? *Dev Cogn Neurosci* 2013; **4**: 38–51.

29 Amir N, Weber G, Beard C, Bomyea J, Taylor CT. The effect of a single-session attention modification program on response to a public-speaking challenge in socially anxious individuals. *J Abnorm Psychol* 2008; **117**: 860–8.

30 Amir N, Beard C, Burns M, Bomyea J. Attention modification program in individuals with generalized anxiety disorder. *J Abnorm Psychol* 2009; **118**: 28–33.

31 Amir N, Bomyea J, Beard C. The effect of single-session interpretation modification on attention bias in socially anxious individuals. *J Anxiety Disord* 2010; **24**: 178.

32 Amir N, Taylor CT, Donohue MC. Predictors of response to an attention modification program in generalized social phobia. *J Consult Clin Psychol* 2011; **79**: 533–41.

33 Baert S, De Raedt R, Schacht R, Koster EHW. Attentional bias training in depression: therapeutic effects depend on depression severity. *J Behav Ther Exp Psychiatry* 2010; **41**: 265–74.

34 Beard C, Amir N. A multi-session interpretation modification program: changes in interpretation and social anxiety symptoms. *Behav Res Ther* 2008; **46**: 1135–41.

35 Boettcher J, Leek L, Matson L, Holmes EA, Browning M, Macleod C, et al. Internet-Based attention bias modification for social anxiety: a randomised controlled comparison of training towards negative and training towards positive cues. *PloS One* 2013; **8**: e71760.

36 Bowler JO, Mackintosh B, Dunn BD, Mathews A, Dalgleish T, Hoppitt L. A comparison of cognitive bias modification for interpretation and computerized cognitive behavior therapy: effects on anxiety, depression, attentional control, and interpretive bias. *J Consult Clin Psychol* 2012; **80**: 1021–33.

37 Carlbring P, Apelstrand M, Sehlin H, Amir N, Rousseau A, Hofmann SG, et al. Internet-delivered attention bias modification training in individuals with

social anxiety disorder–a double blind randomized controlled trial. *BMC Psychiatry* 2012; **12**: 66.

38 Harris LM, Menzies RG. Changing attentional bias: can it effect self-reported anxiety? *Anxiety Stress Coping* 1998; **11**: 167–79.

39 Hazen RA, Vasey MW, Schmidt NB. Attentional retraining: a randomized clinical trial for pathological worry. *J Psychiatr Res* 2009; **43**: 627–633.

40 Heeren A, Reese HE, McNally RJ, Philippot P. Attention training toward and away from threat in social phobia: effects on subjective, behavioral, and physiological measures of anxiety. *Behav Res Ther* 2012; **50**: 30–9.

41 Hirsch CR, Mathews A, Clark DM. Inducing an interpretation bias changes self-imagery: a preliminary investigation. *Behav Res Ther* 2007; **45**: 2173–81.

42 Holmes EA, Mathews A. Mental imagery and emotion: a special relationship? *Emotion* 2005; **5**: 489–97.

43 Hoppitt L, Mathews A, Yiend J, Mackintosh B. Cognitive mechanisms underlying the emotional effects of bias modification. *Appl Cog Psychol* 2010; **24**: 312–25.

44 Lang TJ, Blackwell SE, Harmer CJ, Davison P, Holmes EA. Cognitive bias modification using mental imagery for depression: developing a novel computerized intervention to change negative thinking styles. *Eur J Pers* 2012; **26**: 145–57.

45 Lester KJ, Mathews A, Davison PS, Burgess JL, Yiend J. Modifying cognitive errors promotes cognitive well being: a new approach to bias modification. *J Behav Ther Exp Psychiatry* 2011; **42**: 298–308.

46 Li S, Tan J, Qian M, Liu X. Continual training of attentional bias in social anxiety. *Behav Res Ther* 2008; **46**: 905–12.

47 Mackintosh B, Mathews A, Yiend J, Ridgeway V, Cook E. Induced biases in emotional interpretation influence stress vulnerability and endure despite changes in context. *Behav Ther* 2006; **37**: 209–22.

48 Mathews A, Ridgeway V, Cook E, Yiend J. Inducing a benign interpretational bias reduces trait anxiety. *J Behav Ther Exp Psychiatry* 2007; **38**: 225–36.

49 Murphy R, Hirsch CR, Mathews A, Smith K, Clark DM. Facilitating a benign interpretation bias in a high socially anxious population. *Behav Res Therapy* 2007; **45**: 1517–29.

50 Najmi S, Amir N. The effect of attention training on a behavioral test of contamination fears in individuals with subclinical obsessive-compulsive symptoms. *J Abnorm Psychol* 2010; **119**: 136–42.

51 Neubauer K, von Auer M, Murray E, Petermann F, Helbig-Lang S, Gerlach AL. Internet-delivered attention modification training as a treatment for social phobia: a randomized controlled trial. *Behav Res Ther* 2013; **51**: 87–97.

52 Rapee, RM, MacLeod C, Carpenter L, Gaston JE, Frei J, Peters L, et al. Integrating cognitive bias modification into a standard cognitive behavioural treatment package for social phobia: a randomized controlled trial. *Behav Res Therapy* 2013; **51**: 207–215.

53 Salemink E, van den Hout M, Kindt M. Trained interpretive bias and anxiety. *Behav Res Ther* 2007; **45**: 329–40.

54 Salemink E, van den Hout M, Kindt M. Effects of positive interpretive bias modification in highly anxious individuals. *J Anxiety Dis* 2009; **23**: 676–83.

55 Salemink E, van den Hout M, Kindt M. Trained interpretive bias: validity and effects on anxiety. *J Behav Ther Exp Psychiatry* 2007; **38**: 212–24.

56 Schmidt NB, Richey JA, Buckner JD, Timpano KR. Attention training for generalized social anxiety disorder. *J Abnorm Psychol* 2009; **118**: 5–14.

57 Schoorl M, Putman P, Van Der Does W. Attentional bias modification in posttraumatic stress disorder: a randomized controlled trial. *Psychother Psychosom* 2013; **82**: 99–105.

58 Sharpe L, Ianiello M, Dear BF, Nicholson Perry K, Refshauge K, Nicholas MK. Is there a potential role for attention bias modification in pain patients? Results of 2 randomised, controlled trials. *Pain* 2012; **153**: 722–31.

59 Steel C, Wykes T, Ruddle A, Smith G, Shah DM, Holmes EA. Can we harness computerised cognitive bias modification to treat anxiety in schizophrenia? A first step highlighting the role of mental imagery. *Psychiatry Res* 2010; **178**: 451–5.

60 Steinman SA, Teachman BA. Modifying interpretations among individuals high in anxiety sensitivity. *J Anxiety Dis* 2010; **24**: 71–8.

61 Watkins ER, Taylor RS, Byng R, Baeyens C, Read R, Pearson K, et al. Guided self-help concreteness training as an intervention for major depression in primary care: a Phase II randomized controlled trial. *Psychol Med* 2012; **42**: 1359–71.

62 Yiend J, Mackintosh B, Mathews A. Enduring consequences of experimentally induced biases in interpretation. *Behav Res Therapy* 2005; **43**: 779–97.

63 See J, MacLeod C, Bridle R. The reduction of anxiety vulnerability through the modification of attentional bias: a real-world study using a home-based cognitive bias modification procedure. *J Abnorm Psychol* 2009; **118**: 65–75.

64 Wells TT, Beevers CG. Biased attention and dysphoria: manipulating selective attention reduces subsequent depressive symptoms. *Cogn Emot* 2010; **24**: 719–28.

65 Peters KD, Constans JI, Mathews A. Experimental modification of attribution processes. *J Abnorm Psychol* 2011; **120**: 168–73.

66 Van Bockstaele B, Verschuere B, Tibboel H, De Houwer J, Crombez G, Koster EHW. A review of current evidence for the causal impact of attentional bias on fear and anxiety. *Psychol Bull* 2014; **140**: 682–721.

67 Rosenthal R. *Experimenter Effects in Behavioral Research*. Appleton-Century-Crofts, 1966.

68 Amir N, Beard C, Taylor CT, Klumpp H, Elias J, Burns M, et al. Attention training in individuals with generalized social phobia: a randomized controlled trial. *J Consult Clin Psychol* 2009; **77**: 961–73.

69 Boettcher J, Hasselrot J, Sund E, Andersson G, Carlbring P. Combining attention training with internet-based cognitive-behavioural self-help for social anxiety: a randomised controlled trial. *Cogn Behav Ther* 2014; **43**: 34–48.

70 Higgins JPT, Green, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011)*. The Cochrane Collaboration, 2011 (www.cochrane-handbook.org.).

71 Ioannidis JPA. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA* 1998; **279**: 281–6.

72 Ioannidis JPA. Why most discovered true associations are inflated. *Epidemiology* 2008; **19**: 640–8.

73 Macleod C, Holmes EA. Cognitive bias modification: an intervention approach worth attending to. *Am J Psychiatry* 2012; **169**: 118–20.

74 Watkins ER, Baeyens CB, Read R. Concreteness training reduces dysphoria: proof-of-principle for repeated cognitive bias modification in depression. *J Abnorm Psychol* 2009; **118**: 55–64.

75 Watkins ER, Taylor RS, Byng R, Baeyens C, Read R, Pearson K, et al. Guided self-help concreteness training as an intervention for major depression in primary care: a Phase II randomized controlled trial. *Psychol Med* 2012; **42**: 1359–71.

EXTRA CONTENT ONLINE