Clinical Psychology Review 52 (2017) 92–107

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

Autobiographical episodic memory-based training for the treatment of mood, anxiety and stress-related disorders: A systematic review and meta-analysis

Caitlin Hitchcock [a,*], Aliza Werner-Seidler [b], Simon E. Blackwell [a,c], Tim Dalgleish [a,d]

[a] Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK
[b] The Black Dog Institute, Sydney, Australia
[c] Ruhr-Universität Bochum, Bochum, Germany
[d] Cambridgeshire and Peterborough NHS Foundation Trust, UK

HIGHLIGHTS

- Autobiographical episodic memory training (AET) improves depression ($d = 0.32$).
- There was limited research on AET for diagnosed anxiety and stress-related disorders.
- AET is a promising option for low-intensity treatment of affective disturbance.

ARTICLE INFO

Article history:
Received 23 May 2016
Received in revised form 28 September 2016
Accepted 15 December 2016
Available online 21 December 2016

Keywords:
Autobiographical memory
Episodic memory
Memory training
Depression
Anxiety
Posttraumatic stress

ABSTRACT

We review evidence for training programmes that manipulate autobiographical processing in order to treat mood, anxiety, and stress-related disorders, using the GRADE criteria to judge evidence quality. We also position the current status of this research within the UK Medical Research Council's (2000, 2008) framework for the development of novel interventions. A literature search according to PRISMA guidelines identified 15 studies that compared an autobiographical episodic memory-based training (AET) programme to a control condition, in samples with a clinician-derived diagnosis. Identified AET programmes included Memory Specificity Training (Raes, Williams, & Hermans, 2009), concreteness training (Watkins, Baeyens, & Read, 2009), Competitive Memory Training (Korrelboom, van der Weele, Gjaltema, & Hoogstraten, 2009), imagery-based training of future autobiographical episodes (Blackwell & Holmes, 2010), and life review/reminiscence therapy (Arean et al., 1993). Cohen's d was calculated for between-group differences in symptom change from pre- to post-intervention and to follow-up. We also completed meta-analyses for programmes evaluated across multiple studies, and for the overall effect of AET as a treatment approach. Results demonstrated promising evidence for AET in the treatment of depression ($d = 0.32$), however effect sizes varied substantially (from −0.18 to 1.91) across the different training protocols. Currently, research on AET for the treatment of anxiety and stress-related disorders is not yet at a stage to draw firm conclusions regarding efficacy as there were only a very small number of studies which met inclusion criteria. AET offers a potential avenue through which low-intensity treatment for affective disturbance might be offered.

Contents

* Corresponding author.
E-mail address: Caitlin.hitchcock@mrc-cbu.cam.ac.uk (C. Hitchcock).

http://dx.doi.org/10.1016/j.cpr.2016.12.003

International guidelines for the development of novel psychological interventions emphasize the need for such complex interventions to be derived from theory and experimental science (Medical Research Council, 2000, 2008). Autobiographical episodic memory-based training (AET) paradigms are a prototypical example of interventions developed from basic science showing a robust relationship between psychological symptoms and biases in the processing of autobiographical memory. Such chronic and maladaptive patterns in remembering are common across mood, anxiety, and trauma- and stress-related disorders. Although dysfunction may be experienced in multiple memory systems (e.g., recognition memory, Brand, Jolles, & Gispen-de Wied, 1992; working memory, Christopher & MacDonald, 2005), the biased processing of personal (autobiographical) memories, particularly those of an emotional nature, is a core feature of mood, anxiety, and stress-related disorders (Beck, 1967; Coles & Heimberg, 2002). Disturbance in the processing of emotional personal memories plays a key role in maintaining current symptoms (Morgan, 2010; Sumner, Griffith, & Mineka, 2010), and can also persist during remission, increasing the risk for relapse (Gotlib & Joormann, 2010). As such, targeting maladaptive autobiographical memory processing patterns may reduce current symptoms, as well as reduce the risk of relapse.

1. Autobiographical processing in mood, anxiety, and stress-related disorders

Autobiographical memory comprises a personal store of life experiences that is critical in shaping our sense of self. According to dominant theoretical models (e.g., Conway & Pleydell-Pearce, 2000), the process of recalling personal memories involves a search of autobiographical knowledge hierarchically arranged from broad, general, summarized autobiographical information down to event-specific detail. This allows an individual to retrieve memories at different levels of abstraction ranging from generalizations (e.g., enjoying picnics on summer days), which provide broad summaries of the self and world, to specific, single episode memories that are rich in detail (e.g., the day we had a picnic in last summer's heat wave). In affective disorders, the search process becomes disrupted when attempting to retrieve personal memories. Several different mechanisms have been proposed to cause such disruption. These include the selective recall of information that is consistent with current mood (i.e., mood-congruent retrieval; Matt, Vázquez, & Campbell, 1992), and disruptions in the mechanisms governing the retrieval process (e.g., executive control; Dalgleish et al., 2007). For a theoretical account of these mechanisms, see Williams et al. (2007). Disruption in autobiographical retrieval creates a bias toward overgeneral processing of autobiographical information, along with bias in the emotional valence of retrieved information, and we consider each of these in turn.

1.1. Overgeneralization

Mood and stress-related disorders are associated with the systematic retrieval of autobiographical information in an abstract and overgeneral manner (Sumner et al., 2010; Watkins et al., 2012; Williams et al., 2007). When the hierarchical search of autobiographical information is interrupted prior to reaching event-specific information, general summaries will dominate retrieval, such that the recall of individual autobiographical episodes is compromised.

Such impoverished recollection which lacks specific, detailed emotional information may directly impact mood, but can also perpetuate symptoms through intermediate processes that influence symptoms. For example, the recall of emotional memories in an overgeneral and

abstract manner can promote abstract rumination, which can worsen symptoms of anxiety and depression (Nolen-Hoeksema, 2000; Starr & Davila, 2012). Overgeneral recall also appears to have a negative impact on social functioning, as reduced specificity can impair social problem solving (Jing, Madore, & Schacter, 2016), and the sharing of specific personal experiences is a key influence on feelings of social closeness (Beike, Brandon, & Cole, 2016). A predisposition toward broad and general summary memories is also associated with generalised views of the self and world (Watkins et al., 2009), which anchor disorder-related schemata. For example, repeated recall of summaries of threatening events may perpetuate the belief that the world is an unsafe place (Bryant, Sutherland, & Guthrie, 2007; Sumner et al., 2010). Finally, overgeneralization may not only affect past autobiographical memories, but also future projections, as the skills that underlie recall of past autobiographical episodes also underlie the ability to imagine future events (Addis, Hach, & Tippett, 2016; Schacter & Addis, 2007; Williams et al., 2007). This is particularly evident in posttraumatic stress disorder (PTSD; Kleim, Graham, Fihosy, Stott, & Ehlers, 2014) and complicated grief (Robinaugh & McNally, 2013).

1.2. Reduced positivity

Mood, anxiety, and stress-related disorders are not only related to disruptions in the specificity of memory recall, but also to biases in the valence of retrieved information, particularly in the form of relatively reduced access to positive autobiographical memories. For example, studies have shown that negative material is more easily activated in depressive disorders (Gotlib & Joormann, 2010), while bias toward threat-relevant meaning is characteristic of anxiety and stress-related disorders (Coles & Heimberg, 2002; Morgan, 2010). Reduced memory positivity is not only reflected by reduced salience of positive meanings, but also by diminution of the positive affect that can be derived from the memory. For example, positive autobiographical episodes are reported as less vivid in both mood (Werner-Seidler & Moulds, 2011) and anxiety disorders (Moscovitch, Gavric, Merrifield, Bielak, & Moscovitch, 2011), and this extends to future episodes (Holmes, Blackwell, Burnett Heyes, Renner, & Raes, 2016). Recollection of positive episodes can also have little or limited beneficial impact on affect for those experiencing psychological disorders (Joormann, Siemer, & Gotlib, 2007; Werner-Seidler & Moulds, 2012). As recall of positive memories is a widely-used mood regulation technique (Joormann et al., 2007), this reduced salience of positive autobiographical episodes can impair efforts to regulate negative moods and thus maintain symptoms of affective and stress-related disorders.

2. Rationale for autobiographical episodic training

As the nature of these autobiographical processing biases is now well established (for reviews see Dalgleish & Werner-Seidler, 2014 and Coles & Heimberg, 2002), novel cognitive interventions have been developed to specifically target and modify these biases. We define such autobiographical episodic training (AET) as any training protocol that targets either retrieval of past autobiographical episodes or projection of future autobiographical episodes, with the aim of modifying processing biases (e.g., overgeneralization, reduced salience of positive material). While traditional cognitive therapies target the content of autobiographical material to reduce symptoms, AET seeks to modify the *processing* of autobiographical information. In this way, AET interventions do not focus on restructuring the specific appraisals attached to an event (as in cognitive restructuring), or on reducing distress caused by the memory of that event (e.g., through exposure), but rather on modifying the processing disturbances of biases that lead an individual to recall these memories in the first place.

The development of AET paradigms represents a prototypical example of translation of basic science into novel clinical interventions. Recommendations (Medical Research Council, 2000, 2008) for the development and dissemination of complex interventions encourage an iterative phase-based approach in which experimental findings are translated into clinical techniques, which are then evaluated in pilot trials, before progressing to efficacy trials in comparison to existing gold-standard interventions, and thereafter to larger definitive and pragmatic trials of treatment effectiveness (see Fig. 1; Medical Research Council, 2000).

In terms of this evolutionary trajectory, drawing on experimental evidence that processing of autobiographical episodes is malleable to manipulation, researchers developed clinical techniques to improve retrieval of specific autobiographical episodes (Debeer, Raes, Williams, & Hermans, 2011) and vivid, positive memories (Dalgleish et al., 2013). These techniques were subsequently translated into clinical interventions that promote repeated retrieval or projection of specific autobiographical episodes, or used repeated engagement in positive emotion attached to autobiographical episodes to strengthen positive representations in memory (e.g., Raes et al., 2009). Small, uncontrolled proof-of-principle trials of these interventions provided initial evidence that AET was able to influence symptoms, and served as a foundation for subsequent controlled pilot and efficacy trials that sought to further evaluate the promise of AET as a treatment approach (Medical Research Council, 2000, 2008).

In this review, as well as synthesising the currently available controlled trial data for AET we will also reflect on the current status of each of the key interventions with respect to this translational trajectory of complex intervention development (Medical Research Council, 2000, 2008). More formally, we will evaluate the controlled trials evidence for each AET intervention against the set of established GRADE criteria (Balshem, Helfand, Schünemann, Oxman, Kunz, Brozek, et al., 2011), to provide an indication of whether further, later-stage trials of treatment efficacy and effectiveness are warranted.

Our review summarizes the literature on AET programmes that target the processing of either past or future autobiographical episodes, with the aim of treating symptoms of mood, anxiety, and stress-related disorders. There are previous opinion pieces and reviews of the nature of autobiographical deficits (Dalgleish & Werner-Seidler, 2014) and of memory training for depression (Becker, Vanderhasselt, & Vrijsen, 2015). However, these articles are narrative summaries of the principles behind training, are restricted to depression, and do not provide a systematic synthesis of the data.

There are currently a number of studies that have explored the effect of AET on psychological symptoms, but not yet enough to allow for comprehensive quantitative meta-analysis of each separate AET paradigm (i.e., only one study has been identified for several of the AET paradigms). We therefore only present meta-analytic data for those paradigms that have been examined in two or more controlled studies. Cochrane guidance suggests that quantitative meta-analysis may produce misleading results when there is a large degree of heterogeneity between studies. Our review of the literature indicated that observed heterogeneity between studies was not sufficient to preclude a cross-paradigm meta-analysis of the efficacy of AET. However, because of this issue, and as Cochrane guidance advises transparency in the decision to complete meta-analysis, we emphasize the diversity between the analysed studies, and present the overall analysis only to inform the estimate of the likely effect size of AET as a treatment approach. In circumstances of inconsistency in effect size between studies, separate subgroup analysis should be considered (Guyatt et al., 2011), and in discussing and evaluating the evidence, we refer to the effect sizes for the individual studies.

3. Method

3.1. Protocol and registration

This review was conducted in accordance with PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & The, 2009), and the review protocol

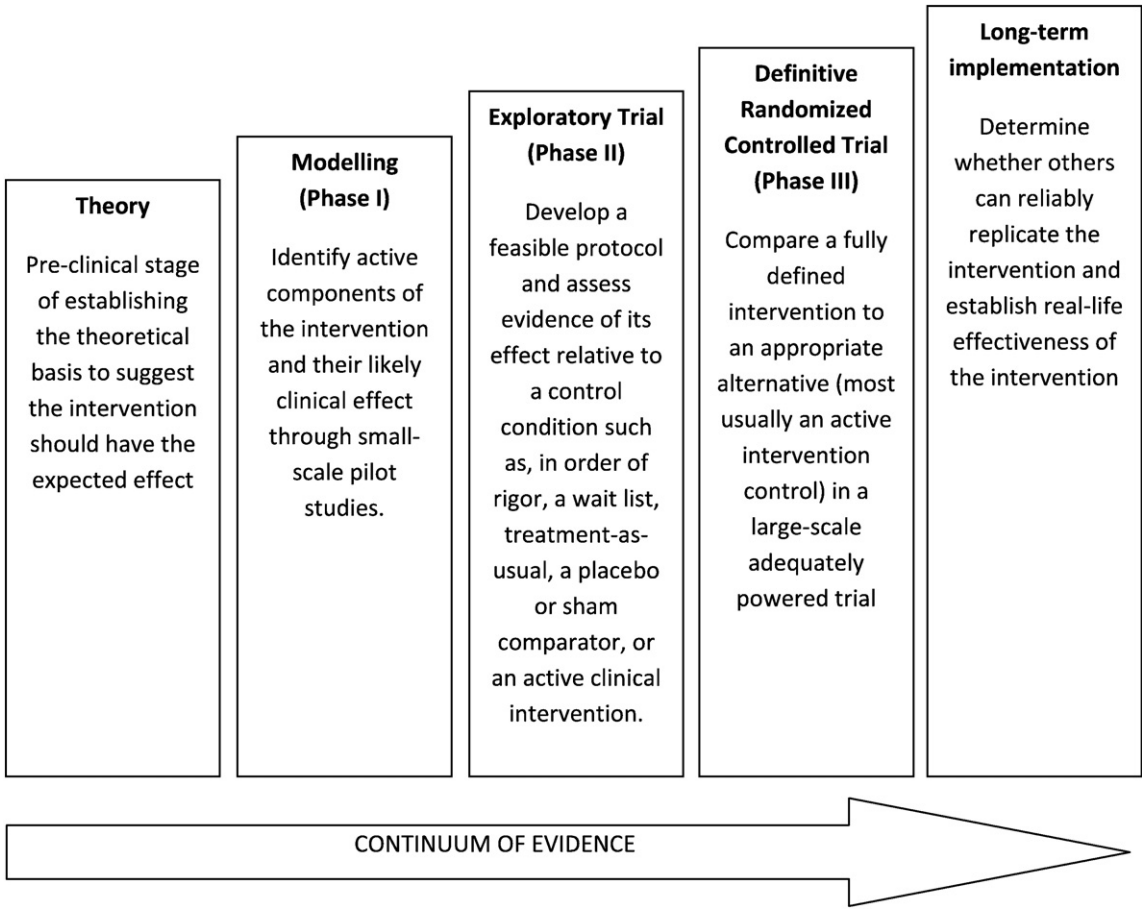

Fig. 1. Medical Research Council (MRC) framework for development and evaluation of complex interventions. Adapted from MRC (2000).

was registered with PROSPERO (http://www.crd.york.ac.uk/PROSPERO/index.asp), registration number CRD42015027445.

3.2. Eligibility criteria

We include studies that have examined the effect of an AET protocol on a primary outcome of clinical symptoms of diagnosed mood, anxiety, or stress-related psychiatric disorders, in samples with a clinician-made diagnosis. Due to our emphasis on the treatment efficacy of training, we do not include studies that only include samples with subclinical symptoms. While studies using subclinical samples are valuable in establishing proof-of-principle prior to trials in clinical populations, and treating subclinical symptoms is a valuable aim in itself, our aim was to investigate treatment efficacy in clinical populations rather than investigate broader questions about the efficacy of procedures in targeting mechanisms of interest. Further, there is considerable variation across subclinical studies in terms of their aims and the kinds of outcome measures used (e.g., ranging from low level symptoms of anxiety to changes on bespoke visual analogue scales) such that attempting to compile effect sizes across clinical and subclinical samples would not provide a clear picture. Consequently we chose to restrict our analyses to clinical samples to provide a clearer indication of the current treatment efficacy of these programmes.

Training may have been directed at processing of previous experiences, or at future-oriented autobiographical episodes. We elected to also include those programmes that explicitly target the projection of future episodes as there is good evidence that the same processing skills are used in the recall and projection of autobiographical episodes, and their adaptive functions are closely linked (e.g. Schacter, Addis, & Buckner, 2007). As we were interested in the training of autobiographical processing skills, we excluded studies that focused on the recall of only one (or a discrete set of) prior episode(s), as is common in exposure-based treatments for trauma (e.g., trauma-focused Cognitive Behavioral Therapy; CBT).

Studies also needed to include a randomisation procedure and a control condition. Exclusion criteria were therefore uncontrolled studies/trials, samples without a primary diagnosis of a mood, anxiety or stress-related disorder, and primary outcome measures that were not symptoms of a mood, anxiety, or stress-related disorder. Studies were required to be published in the English language in peer-reviewed journals.

3.3. Study identification and selection

Articles were identified through an iterative series of PsychInfo and PubMed searches up to 18 March 2016, independently completed by both the first and second authors. Studies that could be excluded on the basis of title and abstract were immediately rejected. Full text articles of the remaining publications where then independently reviewed by the first and second author to determine inclusion. Any discrepancy between the authors was resolved via discussion with the last author. The PRISMA diagram for study screening is presented in Fig. 2.

We began by completing Boolean keyword searches in PsychInfo using the terms; 1) 'memory training AND depression', 2) 'memory training AND anxiety disorders'. We elected to use the broad term 'memory training' to maximize search sensitivity. The relevant papers identified in these two initial searches reported four different types of AET: memory specificity training; competitive memory training; positive memory training; and life review therapy, which is also called reminiscence therapy. Reference lists of identified articles were examined to

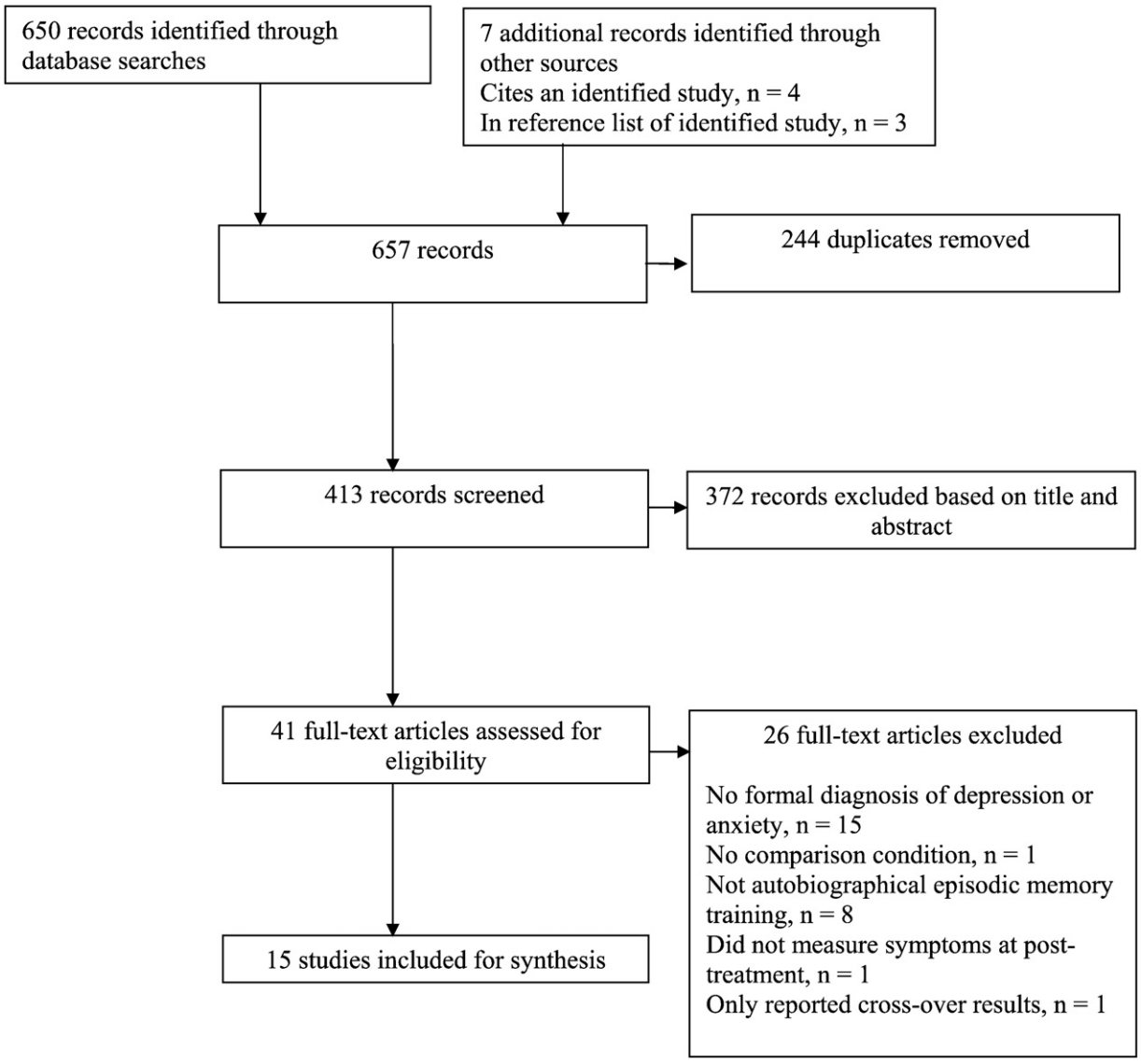

Fig. 2. PRISMA flow chart.

source further articles, as were papers that cited the included articles. This identified two further training types that met inclusion criteria. These were concreteness training, and two separate cognitive bias modification (CBM) protocols which focussed on improving the projection of future autobiographical episodes; positive imagery based CBM (Blackwell & Holmes, 2010) and a version of CBM for interpretation which included an imagery component (Yiend et al., 2014).

We next completed a second series of searches, for each permutation of identified training type (e.g., memory specificity training) and disorder (e.g., obsessive-compulsive disorder), for all of the mood, anxiety, and stress-related disorders listed in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013). All of the above searches were then repeated using text word searches in PubMed (example search strategy; 'memory training[Text Word] AND depression[Text Word]').

3.4. Primary outcomes

Our primary outcomes were effect sizes (measured using Cohen's *d*) for the difference between active and control groups in change on a continuous self-report measure of symptoms from a) pre- to post-intervention, and b) change from pre-intervention to follow-up, when follow-up assessments were completed. Follow-up assessments were only considered when design of the study did not include a crossover or half-crossover.

3.5. Data extraction

Data needed to calculate effect sizes from each identified study were extracted into a Microsoft Excel spreadsheet by the first author, and independently checked by the second author. Information extracted included: Author names, year of publication, diagnosis, name of the main symptom measure, name of the training paradigm and control condition, number of participants in each condition, and mean and standard deviation for pre-intervention, post-intervention and follow-up (where possible) for training and control conditions. Where this information was not included in the publication, the authors were emailed to request the required data. We received a 100% response rate from emailed authors.

3.6. Risk of bias

We used the Cochrane Risk of Bias Tool to assess study quality and risk of bias (Higgins & Green, 2008). The tool allowed us to determine

risk of bias in: 1) allocation sequence (i.e., method used for random allocation of participants to groups); 2) allocation concealment (i.e., whether intervention allocations were appropriately concealed during enrolment); 3) blinding of outcome assessors and study personnel; 4) incomplete outcome data (i.e., appropriate reporting and consideration of missing data due to attrition and exclusions, for example, through intent-to-treat analysis); and 5) selective reporting (i.e., degree of reporting of all pre-specified outcomes). Evaluations of risk of bias (reported in Table 2) were completed independently by the first and second authors, and discrepancies were resolved through discussion.[1]

3.7. Evaluation of the quality of the body of evidence

In addition to rating the risk of bias in each study, we also considered the status of intervention development in the context of the MRC's guidelines for complex intervention development (Medical Research Council, 2000, 2008) as well as the overall quality of the evidence, across studies of each training paradigm. A systematic approach to making judgements about the quality of a body of evidence ensures accurate recommendations, can help prevent errors, and aids clear communication of information (Atkins et al., 2004). We therefore elected to use the GRADE guidelines (Balshem, Helfand, Schünemann, Oxman, Kunz, Brozek, et al., 2011) to evaluate evidence quality, because this approach is used to determine whether a treatment approach is recommended by the National Institute for Health and Care Excellence (NICE, 2012). A body of evidence that is formed by randomised trials (as is the focus of this review) begins with a rating of high, which is then reduced by the presence of undesirable factors, but may also be increased by the presence of desirable factors (see Table 1). Undesirable factors comprise: risk of bias; inconsistency across studies in the estimate of the effect; indirectness (i.e., participants and outcome measures are dissimilar to those of interest); imprecision in data analysis and account of missing data; and publication bias. Desirable factors comprise the presence of a large effect size across studies, evidence of a dose response gradient, and careful consideration of possible confounds. We critically appraised evidence for each training paradigm, and AET overall, against these criteria.

4. Statistical analyses

4.1. Calculation of individual effect sizes

Cohen's *d* and the associated 95% confidence intervals were calculated using intent-to-treat data for the difference between AET and control groups in change on a continuous self-report measure of symptoms from pre- to post-intervention, and where possible, pre-intervention to follow-up. In our effect size calculation, we used pooled variance, calculated using the pre- and post- (or pre- and follow-up) standard deviations on the symptom measure for AET and control groups.[2] Although our focus was on treatment outcomes, we did also use this procedure to calculate the between-group effect sizes (Cohen's *d*s) for pre- to post-intervention change in specificity and/or positivity (see Table 2) to provide an indication of whether the reviewed interventions did impact the targeted underlying processes. Effect sizes of 0.20, 0.50 and 0.80 refer to small, medium, and large effects, respectively (Cohen, 1992), and are in favor of the treatment condition, except where noted.

4.2. Calculation of pooled effect sizes

We also calculated a pooled Cohen's *d* (and associated 95% confidence intervals) for the overall effect of AET on self-reported symptoms, along with subgroup analysis for those training paradigms that had been examined in two or more studies, as meta-analysis of only two studies can still be helpful when seeking to determine the true effect (Cumming, 2013). The pooled Cohen's *d* was weighted (see supplementary materials for forest plot) by sample size of the individual studies, and was calculated using a random effects model in Comprehensive Meta-analysis (Version 3, Biostat, Inc.). We elected to use a random effects model due to the expectation of significant heterogeneity amongst interventions. The random effects model assumes that the size of the true effect varies between studies, and therefore allows a summary estimate of the mean effect of AET on symptoms, while still considering that the size of the effect may vary between training paradigms.

4.3. Testing heterogeneity

When calculation of a pooled effect was completed, we also calculated the I^2 statistic to provide an indication of heterogeneity across effect sizes. A value of 0 indicates no heterogeneity, and 0.25, 0.50, and 0.75 indicate low, moderate, and high heterogeneity, respectively (Higgins & Thompson, 2002).

4.4. Testing for potential publication bias

We also inspected the funnel plot of the estimate of effect (see supplementary materials) to test for potential publication bias. To account for any estimated publication bias, we conducted the Duval and Tweedie Trim and Fill procedure (Duval & Tweedie, 2000) to determine an adjusted effect size. The procedure takes into account estimated publication bias observed within the funnel plot, and corrects the estimate of the effect. As the analysis of potential publication bias can only be calculated when there are three or more studies, we were unable to provide statistical indication of bias where the pooled effect was determined from only two studies.

5. Results

A total of 15 studies met inclusion criteria, 12 of which examined AET in mood disorders, two in anxiety disorders, and one in stress-related disorders. Table 2 presents study characteristics, risk of bias, and effect sizes for all reviewed studies. One key point of difference between the identified studies was the type of control group used. Use of no intervention, placebo/sham, or active comparison groups reflects the progression of phase-based evaluation of treatments advocated by the MRC (2000, 2008). As the size of the between-group effect is fundamentally impacted by the type of control group used, it is important to consider the type of control when evaluating treatment efficacy. We identified six types of control conditions (see Table 2). Below we summarize the evidence for each training paradigm, presented by disorder.

6. Mood disorders

The literature search identified 12 studies of the effect of an AET protocol on symptoms of Major Depressive Disorder (MDD). No studies were identified that examined the impact of training on symptoms of bipolar disorders, dysthymia, or depressive disorder not otherwise specified. Therefore we present the current state of the evidence in MDD samples.

[1] The authors initially agreed on 87% of ratings. After discussion of discrepancies, authors agreed on 100% of ratings.

[2] Although calculation of Cohen's *d* using an estimated correlation between pre- and post-intervention scores is standard in meta-analysis (where the coefficient may be taken from another study), imputation may be unreliable and is best avoided when sample sizes are small (Higgins, Deeks, & Altman, 2008). Further, estimated within-group correlations differ substantially between active treatments and non-active controls (Balk et al., 2012), and the reviewed studies differed in whether the comparison group received an established, alternate treatment (e.g., problem solving therapy), or no further contact. Use of pooled variance in calculations of Cohen's *d* therefore offered a more conservative estimate of effect size.

Table 1
GRADE criteria for rating the quality of a body of evidence.

Initial quality of a body of evidence for randomised trials	Rating is *reduced* by 1 point for a serious rating, and 2 points for a very serious rating of	Rating is *increased* by 1 point for a rating of large and 2 points for a rating of very large	Quality of a body of evidence
High + + + + (four points)	– Risk of bias	+ Effect size	High (four points) Further research is very unlikely to change our confidence in the estimate of the effect
	– Inconsistency in effect size between studies	+ Dose response is evident	Moderate (three points) Further research is likely to have an important impact on our confidence in the estimate of the effect, and may change the effect
	– Indirectness in measurement of the outcome variables	+ Consideration of all plausible confounds increases confidence in the estimate of effect	Low (two points) Further research is very likely to have an important impact on our confidence in the estimate of the effect, and may change the effect
	– Imprecision in data analysis and treatment of missing data		Very low (one point) Any estimate of effect is very uncertain
	– Publication bias		

Note. Adapted from Balshem et al. (2011).

6.1. Memory specificity training

One study examined of the impact of Memory Specificity Training (MEST; Raes et al., 2009) on symptoms of MDD. MEST seeks to reverse the relative difficulty in retrieving specific autobiographical memories that is characteristic of depression (Raes et al., 2009). MEST was first developed by Raes et al. (2009) as a four-session, group training programme in which individuals repeatedly practice recalling specific autobiographical events in response to cue words, which may be of positive, neutral, or negative emotional valence. Group sessions are led by a therapist who explains the role of memory biases in maintaining symptoms, provides education on the different types of autobiographical memories (i.e., specific episodes versus categorical summaries), and guides cued-recall exercises. Group members are also required to complete cued-recall as homework, to facilitate transfer to daily life. The first uncontrolled trial of MEST with depressed inpatients (Raes et al., 2009) established that memory specificity was amenable to change. These findings provided a foundation for a randomised, controlled trial (RCT) of MEST.

6.1.1. Synthesis of results

The efficacy of MEST in treating depression symptoms was compared to a no-intervention control group in a sample of 23 bereaved and depressed Afghani adolescent refugees living in Iran (Neshat-Doost et al., 2013). Although there was no difference in self-report of depressive symptoms between the MEST group and the control group at post-training, results from the two month follow-up indicated a large effect size favoring the MEST group (see Table 2). To account for this delayed effect, the authors speculate that the effect of MEST on processes implicated in memory specificity (e.g., rumination, problem-solving) needed to consolidate before effects transferred to symptoms (Neshat-Doost et al., 2013). Notably, a mediation analysis revealed that change in memory specificity mediated the effect of condition on depressive symptoms.

Applying the GRADE criteria, the evidence for the efficacy of MEST in treating depression is of moderate quality. The initial high quality rating was reduced one point by the possibility of publication bias (there has only been one, positive result published), and reduced a further point by the inability to assert consistency between research studies (as there has only been one trial completed). Further research is therefore likely to have an important impact on our confidence in the current estimate of the effect. The body of evidence then regained one point for the evidence of a large effect size, resulting in an overall rating of moderate quality. In sum, the current evidence suggests there may be merit in training the retrieval of specific memory episodes to alleviate depressive symptoms. A larger RCT comparing MEST to an active control group is now needed and is indeed underway (Dalgleish et al., 2014). This will provide much-needed data to determine the efficacy of MEST as a treatment option for depression.

6.2. Life review/reminiscence therapy

Life review, or reminiscence, therapy, (LRT) is similar to MEST in that it attempts to train specificity in a group-based format, however the method of training is different. Rather than using structured exercises to facilitate recall of specific memories, the therapist asks the participants a series of questions that are designed to prompt specific memories (e.g., 'What do you consider to be the most important thing that you have done in your life?'). Training aims to improve the salience of specific experiences within the life story, on the basis that failure to integrate one's life experiences will contribute to depression. The training requires individuals to revisit specific experiences in their life, and assists them to integrate these experiences into their life story.

6.2.1. Synthesis of results

Our search identified two randomised, controlled trials, which had examined the impact of LRT on symptoms of MDD. One RCT compared a LRT protocol to supportive therapy (Serrano et al., 2012). There were no between-group differences in change in depressive symptoms from pre-treatment to post-treatment, six-week follow-up, or six month follow-up. There is therefore no evidence to support the efficacy of this type of training over an active psychological programme in treating symptoms of depression. Another RCT (Arean et al., 1993) compared training against problem-solving therapy and against a wait-list control. Older adults diagnosed with MDD were randomised to receive twelve, 1.5 h weekly sessions of problem solving therapy or LRT, or were placed on a wait-list. At post-treatment, both problem solving therapy and LRT produced a reduction in depression symptoms, relative to wait-list control. However, problem-solving therapy was significantly more efficacious than LRT in reducing self-report symptoms at post-treatment. The training paradigm therefore failed to outperform an established treatment option for depression.

As there were two studies that had compared LRT against an active control condition, we calculated the pooled estimate of the effect of LRT on symptoms of depression. The analysis demonstrated a very small effect size in favor of the control condition ($d = -0.16$, 95% CI [−0.67, 0.07], $p = 0.53$). There was a low degree of heterogeneity between studies, $I^2 < 0.01$, $p = 0.92$, however this is likely influenced by the low number of studies examined.

Comparison of the literature against the GRADE criteria indicates a high quality of evidence. The high rating obtained by the completion

Table 2
Study characteristics and effect sizes for reviewed studies.

Authors	Disorder	Training	*n*	Comparison group (type)	*n*	Process measure (bias type)	Process *d* at post	95% CI for *d*	Outcome measure	Assessed at	Outcome *d*	95% CI for *d*	Risk of bias
Arean et al. (1993)	MDD	Reminiscence therapy	28	Problem solving therapy (active)	19	–	–	–	BDI	Post	−0.18	[−0.79, 0.43]	−?+−?
										3 m	0.13	[−0.48, 0.74]	
				Waitlist control	20					Post	0.71	[0.09, 1.32]	
Blackwell et al. (2015)	MDD	Imagery CBM	76	Control CBM (sham training)	74	Positive vividness on prospective Imagery test (P, S)	0.31	[−0.02, 0.64]	BDI-II	Post	−0.03	[−0.36, 0.29]	− − − − −
										1 m	0.08	[−0.25, 0.41]	
										3 m	0.05	[−0.28, 0.41]	
										6 m	−0.02	[−0.35, 0.31]	
						Negative responses on SST (N)	−0.05	[−0.39, 0.29]					
Ekkers et al. (2011)	MDD	COMET (worry and rumination)	53	TAU (specified)	40	Rumination on sadness Scale (N)	0.58	[0.14, 1.02]	Geriatric depression scale	Post	0.60	[0.16, 1.03]	− − − −?
						Ruminative response scale (N)	0.51	[0.07, 0.95]					
Korrelboom et al. (2012)	MDD	COMET (self-esteem)	31	TAU (specified)	30	The self-esteem rating scale- positive subscale (P)	0.44	[−0.08, 0.97]	BDI-II	Post	0.62	[0.08, 1.15]	− −? −?
						Rosenberg self-esteem scale (P)	1.01	[0.46, 1.57]					
						Rumination on sadness scale (N)	0.70	[0.16, 1.23]					
Korrelboom et al. (2014)	Panic	COMET (panic)	70	Applied relaxation (active)	73	–			Panic appraisal inventory- coping with panic	Post	−0.23	[−0.56, 0.11]	?? ? −?
Lang et al. (2012)	MDD	Imagery CBM	13	Control CBM (sham training)	13	Negative responses on SST (N)	0.85	[0.00, 1.71]	BDI-II	Post	0.66	[−0.18, 1.50]	−? ? + ?
										2w	0.38	[0.44, 1.20]	
						Response to intrusions questionnaire (N)	0.93	[0.07, 1.80]					
Moradi et al. (2014)	PTSD	MEST	12	No intervention control	12	Specific memories to positive cues (P, S)	2.28	[1.17, 3.39]	Impact of events scale- revised	Post	6.37	[4.19, 8.55]	− − − − −
										3 m	6.05	[3.96, 8.13]	
Neshat-Doost et al. (2013)	MDD	MEST	12	No intervention control	11	Specific memories (S)	1.36	[0.38, 2.34]	MFQ	Post	0.26	[−0.62, 1.14]	− − − −?
										2 m	0.80	[−0.11, 1.71]	
Schneider et al. (2015)	OCD	COMET (obsessions)	34	Waitlist control	31	–			Y-BOCS	Post	0.11	[−0.39, 0.62]	− − − −?
Serrano et al. (2012)	MDD	Life review therapy	9	Supportive therapy (active)	8	Specific memories (S)	0.46	[−0.72.1.64]	Geriatric depression scale	Post	−0.12	[−1.16, 0.91]	−? − −?
										6w	−0.60	[−1.66, 0.46]	
										6 m	−0.26	[−1.30, 0.78]	

(continued on next page)

Table 2 (*continued*)

Authors	Disorder	Training	*n*	Comparison group (type)	*n*	Process measure (bias type)	Process *d* at post	95% CI for *d*	Outcome measure	Assessed at	Outcome *d*	95% CI for *d*	Risk of bias
Torkan et al. (2014)	MDD	Imagery CBM	13	Control CBM (sham training)	13	Negative responses on SST (N)	1.44	[0.52, 2.36]	BDI-II	Post	0.91	[0.05, 1.77]	−? ? −?
						Ruminative Response Scale (N)	1.38	[0.47, 2.30]		2w	1.10	[0.22, 1.98]	
						VVIQ (S)	0.58	[−0.26, 1.41]					
				No intervention control	13	Negative responses on SST (N)	1.29	[0.38, 2.19]		Post	1.91	[0.59, 3.23]	
						Ruminative Response Scale (N)	1.58	[0.64, 2.53]					
						VVIQ (S)	0.82	[0.18, −1.68]					
Watkins et al. (2012)	MDD	Concreteness training	40	Relaxation training (active)	39	Observer rating of concreteness (S)	0.79	[0.31, 1.26]	BDI-II	Post	0.31	[−0.15, 0.77]	− − − − −
						Negative over-generalisation on ASQ (N, S)	0.73	[0.26, 1.20]		6 m	0.06	[−0.39, 0.52]	
						Ruminative response scale (N)	0.60	[0.13, 1.06]					
				TAU (non-specified)	42	Observer rating of concreteness (S)	0.57	[0.11, 1.03]		Post	0.96	[0.49, 1.43]	
						Negative over-generalisation on ASQ (N, S)	0.66	[0.20, 1.12]		6 m	0.61	[0.15, 1.07]	
						Ruminative response scale (N)	0.76	[0.30, 1.22]					
Williams et al. (2013)	MDD	Imagery CBM	38	Waitlist control	31	Positive responses on AST (P)	0.71	[0.18, 1.25]	BDI-II	Post	0.66	[0.14, 1.19]	−? − −?
						Negative responses on SST (N)	0.15	[−0.36, 0.67]					
						Repetitive thinking questionnaire (N)	0.89	[0.35, 1.43]					
Williams et al. (2015)	MDD	Imagery CBM	36	Control CBM (sham training)	39	Positive responses on AST (P)	0.04	[−0.42, 0.51]	BDI-II	Post	0.29	[−0.18, 0.76]	− − − − −
						Repetitive thinking questionnaire (N)	0.11	[−0.36, 0.58]					
Yiend et al. (2014)	MDD	CBM for interpretation (imagery of future episodes component)	17	Control CBM (sham training)	19	Positive responses on SST (P)	0.53	[−0.14, 1.19]	BDI-II	End of CBM session	−	−	− − − −?
						Negative responses on SST (N)	0.56	[−0.10, 1.22]		1 m	0.06	[−0.62, 0.75]	

Note. Process measure is the measure of specificity (S), positive bias (P), or negative bias (N) through which the intervention was proposed to influence symptoms. All Cohen's *ds* are for the intent-to treat, between-group difference in change from pre- to post-intervention (and from pre- to follow-up when completed), and is in favor of the treatment condition, except when value is negative.
Type of comparison group; no intervention control, waitlist control (in which participants expected to receive treatment at a later date), sham training/placebo control (where the control mimicked the experimental treatment but lacked the proposed active ingredients, thus aiming to control for non-specific treatment effects), non-specified treatment as usual (TAU; which was overseen by a GP and may or may not have involved active psychological and/or pharmacological treatment), specified TAU (active psychological and/or pharmacological treatment provided by a psychiatric service) and active control (in all cases a non-formulation driven psychological intervention, such as relaxation or problem solving therapy).
OCD = obsessive compulsive disorder; PTSD = posttraumatic stress disorder; MEST = memory specificity training; COMET = competitive memory training,
SST = the scrambled sentences task (Rude, Wenzlaff, Gibbs, Vane, & Whitney, 2002); AST = ambiguous sentences task (Berna, Lang, Goodwin, & Holmes, 2011); BDI = Beck depression inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961); BDI-II = Beck depression inventory- second edition (Beck, Steer, & Brown, 1996); ASQ = attributional style questionnaire (Peterson et al., 1982); VVIQ = vividness of visual imagery questionnaire (Marks, 1973); Y-BOCS = Yale-Brown obsessive-compulsive scale (Goodman et al., 1989); MFQ = mood and feeling questionnaire (Messer et al., 1995); w = week follow-up; m = month follow-up.
For risk of bias, − = low risk of bias, + = high risk of bias, and ? = unclear risk of bias on the following indices: random sequencing, allocation concealment, blinding of outcome assessment, incomplete outcome data, and selective reporting.

of randomised controlled trials was not lowered by presence of any of the reducing factors. When considering the consistency criteria, it is worth noting that the size of the effect is comparable between studies, and the two confidence intervals overlap considerably, increasing our confidence in the consistency of findings. In sum, there is no evidence of superior treatment effects of LRT relative to non-formulation driven psychological interventions (Arean et al., 1993; Serrano et al., 2012), which indicates that effects are likely to also be inferior to current treatment approaches for depression.

6.3. Concreteness training

The concreteness training paradigm (Watkins et al., 2009) aims to improve concreteness in recall of negative experiences through daily training exercises. The exercises require the participant to identify specific incidents that have been moderately upsetting, and (1) use imagery to focus on sensory detail of the event, (2) notice the process and sequence in which the event unfolded, and (3) specify the behaviors or steps required to move forward. Training begins with a therapist-guided session, followed by daily completion of a workbook for six weeks. Pilot work demonstrated that the programme reduced depression and rumination in a subclinical sample (Watkins et al., 2009), providing a foundation for the first RCT in a clinical group.

6.3.1. Synthesis of results

One study was identified as targeting overgeneralization using concreteness training. Watkins et al. (2012) randomised 121 participants with MDD to either concreteness training, relaxation training, or treatment as usual (TAU; operationalized as continuation of current care provided via the participant's primary care GP). Results demonstrated that concreteness training did significantly impact the targeted process measures, producing a decrease in both rumination and overgeneralization (i.e., rating of the cause of a negative event as globalised and stable) at post-treatment, relative to relaxation training. Although both relaxation and concreteness training produced a greater reduction in depression symptoms than TAU, no differences were observed between relaxation and concreteness training at post-treatment, three month follow-up, or six month follow-up.

Using the GRADE criteria, the current evidence for the efficacy of concreteness training is of moderate quality. Evidence has been obtained in a randomised controlled trial which showed minimal risk of bias or indirectness, however, the rating was reduced by the inability to discount inconsistency, as there has only been one RCT completed. At this point, current evidence suggests that concreteness training may modify maladaptive cognitive processes that perpetuate depression, but that the direct effect on symptoms is not superior to relaxation training.

6.4. Competitive memory training

While the above training programmes have focussed on reducing overgeneralization, competitive memory training (COMET; Ekkers et al., 2011) aims to modify the valence of activated memory representations. A range of COMET protocols have been developed to target several forms of psychopathology (e.g., depression and anxiety), and this brief training (usually seven or eight sessions) is delivered in small groups. COMET training protocols are based on theory of competitive memory retrieval hierarchies (Brewin, 2006), which proposes that the success of any effective psychological treatment (e.g., cognitive therapy) is attained by changes in the relative activation of positive and negative material such that positive material is strengthened. In line with this theoretical framework, the goal of COMET is to produce preferential recall of positive material from memory by: (1) identifying dysfunctional meanings that are high in the retrieval hierarchy; (2) identifying more functional alternatives that are low in the hierarchy; and (3) focusing on making the alternative, positive representation more easily retrievable by increased activation frequency, enhanced emotional salience, and stronger associations to the target (Korrelboom, Peeters, Blom, & Huijbrechts, 2014). Our search identified COMET protocols that aimed to alter representations associated with self-esteem, and worry and rumination. This was typically achieved through elaborating on past episodes in which personal positive characteristics have been present, aided by the use of imagery, positive self-verbalisation, and repeated retrieval (Korrelboom, Maarsingh, & Huijbrechts, 2012). An early uncontrolled pilot trial reported a decrease in depressive symptoms following the delivery of a COMET protocol for low-esteem (Korrelboom et al., 2009), and research has since progressed to feasibility trials.

6.4.1. Synthesis of results

Our search identified two randomised, controlled trials that had compared the addition of a COMET protocol to TAU, against TAU only (operationalised as continuation of current care provided via the outpatient clinic from which participants were recruited, while on a waiting list to receive COMET at a later time point).One RCT in a sample of 61 depressed individuals compared COMET plus specified TAU to specified TAU only (Korrelboom et al., 2012). COMET produced a significant improvement in positive aspects of self-esteem, demonstrating that the programme was able to improve positivity of information related to the self. Results also indicated a large effect on depressive symptoms in favor of the COMET group at post-intervention, although a mediating role of the proposed treatment mechanism was not analysed. A separate RCT (Ekkers et al., 2011) assessing COMET for worry and rumination in a sample of older adults with MDD also demonstrated that the addition of COMET to specified TAU produced a greater reduction in depressive symptoms than specified TAU alone, although the effect size was smaller. Again, an improvement was observed on the targeted mechanism of change (rumination on negative material), but a mediating role was not evaluated.

As there were two studies that had examined the effect of COMET in depression, we calculated an estimate of the pooled effect. The pooled effect indicated a medium effect of COMET, relative to specified TAU ($d = 0.62$, 95% CI [0.28, 0.94], $p < 0.001$). There was a low degree of heterogeneity between studies, $I^2 < 0.01$, $p = 0.95$, however this is again likely to have been impacted by the low number of studies examined.

The GRADE criteria indicate a moderate quality rating for the COMET body of evidence. The reviewed studies establish that rigorous RCT methodology has evaluated COMET against TAU control groups. There was only one factor that reduced the initial high quality rating; the risk of bias from therapist allegiance effects, as all current evidence has been obtained by the same research group. We thereby arrived at a rating of moderate quality. Based on current findings, COMET appears to be an empirically-based AET programme that shows clinical promise in producing a medium decrease in depressive symptoms, relative to TAU.

6.5. Future-oriented episodic autobiographical processing

As already noted, AET paradigms have also targeted the simulation of potential future episodes, and explored whether repeated practice in imagining hypothetical future-oriented positive autobiographical events may be beneficial in depression. The approach that has been most investigated in depressed samples, positive imagery cognitive bias modification ('imagery CBM'; Lang, Blackwell, Harmer, Davison, & Holmes, 2012), takes the form of simple computerized training. During imagery CBM participants practice imagining positive resolutions for ambiguous scenarios, with the aim of training a bias to automatically imagine positive outcomes for novel ambiguous situations in everyday life. In one training format, participants listen to brief descriptions of everyday situations, structured so that they begin as ambiguous but then always resolve positively. Participants are instructed to vividly imagine themselves in the scenarios as they unfold, as if actively involved in the

situations described. In another training format, participants are presented with an ambiguous photo of an everyday scene, with a positive caption of a few words, and are instructed to generate a mental image combining the photo with the positive words (based on Holmes, Mathews, Mackintosh, & Dalgleish, 2008). This imagery-based CBM is adapted from previous experimental work investigating the modification of interpretation bias (Mathews & Mackintosh, 2000), in which imagining oneself in training scenarios, generally presented as written text, was seen as a central component of the bias modification procedure. The imagery CBM investigated in depression was adapted in further experimental work to enhance the imagery component, for example via audio presentation of training scenarios (e.g. Holmes, Lang, & Shah, 2009), and its investigation in depression was further motivated by the observation that depression is associated with reduced vividness for positive future-oriented mental imagery (Morina, Deeprose, Pusowski, Schmid, & Holmes, 2011). Within an AET framework, imagery CBM can be understood as both reversing an existing valence bias (i.e., training access to positive rather than negative representations when confronted with ambiguity), and training in the generation of specific (future) episodes (via the generation of hundreds of positive mental images). Following successful translation of the paradigm to a depressed population via a case-series (Blackwell & Holmes, 2010), exploratory trials have now been completed.

6.5.1. Synthesis of results

The literature search identified six studies that have compared CBM protocols involving generation of autobiographical episodes (via mental imagery) to a control condition in a sample of depressed adults and measured symptom outcomes. Five of these used the positive imagery CBM described above. One further study (Yiend et al., 2014) determined whether depressive symptoms may be improved by a single session of a variant of CBM for interpretation designed to target depression-related cognitive errors (see Lester, Mathews, Davison, Burgess, & Yiend, 2011). This additionally included an exercise in which the individual had to project a series of specific events that may occur in the future. The study demonstrated that a single session of this CBM variant was able to produce an increase in positivity bias (see Table 2) by the end of the session, but that this did not impact depression symptoms one month later. However, it may be that a single session of an intervention is not enough to produce a significant change in symptoms one month later.

Of the five studies investigating the positive imagery CBM, most ($n = 4$) investigated a one-week training schedule of daily sessions completed from home, each involving 64 unique training scenarios. In two studies, the training was preceded by a face-to-face introductory session, after which the participants completed sessions on their home computer. Using a combination of photograph and auditory cues for training, Lang et al. (2012) compared positive imagery CBM to a 'sham training' control CBM condition in which half of the ambiguous training scenarios resolved positively and half resolved negatively, thus aiming to remove the training contingency to always imagine a positive resolution. The large effect size was in favor of positive imagery CBM, although the 95% confidence interval overlapped with zero. Using only auditory cues, Torkan et al. (2014) compared positive imagery CBM to both a sham CBM control condition in which participants were presented with identical training scenarios but without the instruction to imagine them, and to a no-task control. A large effect on depressive symptoms was observed relative to no-task control over the one week of training, and relative to sham CBM at both post-training and at two-week follow-up. Notably, the training produced a large decrease in negative bias in both studies (see Table 2). Correlations were found between bias change and symptom reduction, although formal mediation analyses were not conducted.

Other research has investigated the efficacy of imagery CBM delivered remotely via the internet (i.e., without the face-to-face introduction). In two of the identified studies, one week of training preceded a ten-week internet CBT (iCBT) intervention. The addition of the training programme to CBT is therefore distinct from other reviewed studies, which have only explored training as a supplement to TAU (e.g., evaluations of COMET). In a first RCT, Williams, Blackwell, Mackenzie, Holmes, and Andrews (2013) compared the imagery CBM + iCBT intervention to a waitlist control. In a second RCT (Williams et al., 2015), the combined intervention was compared to a sham training/control version of the CBM (half of the scenarios resolving positively, half resolving negatively, as in Lang et al., 2012), also followed by the ten-week iCBT course. The first study found a moderate between-group effect for an increase in positive bias and a reduction in symptoms of depression in favor of the active treatment compared to the waitlist after the one-week training, using intent-to-treat analysis. Further, a mediation analysis provided evidence for partial mediation of the symptom reduction via bias change. In the second study, the decrease in symptoms of depression over the one-week training for the combined iCBT + imagery CBM condition was not significantly greater than in the combined iCBT + sham CBM control condition in the intent-to-intent analysis, corresponding to a small between-group effect size. A negligible effect was found on positive bias. In a per-protocol analysis including those participants who completed every session of the imagery CBM or sham CBM, there was a significantly greater reduction in symptoms in the imagery CBM compared to the sham CBM condition, corresponding to a moderate ($d = 0.70$) effect size.

Finally, one RCT has investigated a longer four-week version of the training in MDD, delivered via the internet after an initial face-to-face introduction. Blackwell et al. (2015) compared imagery CBM to a sham CBM control condition (half of the scenarios resolving positively, half resolving negatively, and no instructions to use imagery), and followed up participants over six months post-intervention. There was no difference between conditions in vividness of positive episodes or in decrease in depression symptoms over the course of the study. Although increase in positive imagery vividness correlated with decrease in depression symptoms over the course of the intervention, the lack of between-group differences in process or symptom outcomes meant that mediation analysis was not warranted.

The pooled effect of CBM protocols including imagery-based generation of autobiographical episodes was of medium size ($d = 0.46$, 95% CI [0.19, 0.73], $p = 0.001$). There was no evidence of heterogeneity between studies ($I^2 = 0.01$, $p = 0.42$), and the estimate remained stable once corrected for estimated publication bias, $d = 0.41$, 95% CI [0.14, 0.68]. This pooled estimate included a study that compared imagery-based CBM to a waitlist control, however five of the six trials used the more rigorous comparison (Medical Research Council, 2000) of a comparator CBM condition that aimed to control for the non-specific effects of undergoing training and of engaging with the training stimuli – a more advanced phase of evaluation with respect to the trajectory of intervention development (Medical Research Council, 2000). To reflect this, we completed a follow-up analysis of the pooled estimate of imagery-based CBM versus a sham training control CBM on depressive symptoms. The analysis demonstrated a small, though non-significant, effect of imagery-based CBM, relative to control CBM ($d = 0.26$, 95% CI [−0.06, 0.58], $p = 0.11$). There was a low degree of heterogeneity between studies, $I^2 = 37.83$, $p = 0.17$. The estimate of the effect did substantially decrease once corrected for estimated publication bias, $d = 0.09$, 95% CI [−0.26, 0.44]. In interpreting this result, it is worth noting that although designated as 'sham' training, the control conditions are unlikely to be psychologically 'inert' as they generally retain many of the 'active' components of the training that may provide beneficial effects (e.g. repeated generation of mental imagery, exposure to emotionally ambiguous material and its resolution). We expand on this further in the Discussion section.

Applying the GRADE criteria, the body of evidence for imagery CBM is of moderate quality. The high rating awarded for completion of RCTs is decreased by estimation of publication bias and inconsistency in treatment effects between studies. Effect sizes also vary considerably

between intent-to-treat and per-protocol analyses within the individual studies, indicating that dose responsiveness may need further evaluation. Indeed, there is considerable variation in effect size between the one session, one week, and four week training protocols, although there are not yet enough studies of each training length to allow analysis of training length as a moderator of treatment efficacy. While the initial clinical studies produced promising results, the outcomes from more recent, rigorous RCTs have been more mixed.

6.6. Synthesis of results across training paradigms for depression

The pooled effect across AET paradigms assessed in MDD samples (i.e., all 12 studies) demonstrated a small effect of AET on depressive symptoms, $d = 0.32$, 95% CI [0.13, 0.52]. There was a low degree of heterogeneity between studies, $I^2 = 30.88$, $p = 0.14$. The estimate of the effect remained stable once corrected for estimated publication bias, $d = 0.32$, 95% CI [0.13, 0.52]. As this estimate includes comparisons to waitlist control groups, active control groups (e.g., support groups), continuation of TAU, and other established psychological treatments, we also calculated a pooled effect for studies with control conditions other than waitlist ($n = 10$) only. As noted above, evaluating treatment effects against active comparison groups provides a more rigorous test of treatment efficacy (Medical Research Council, 2000, 2008). The pooled effect remained consistent, $d = 0.29$, 95% CI [0.08, 0.50], as did heterogeneity between studies, $I^2 = 35.09$, $p = 0.13$. This remained the case when the effect was corrected for estimated publication bias, $d = 0.29$, 95% CI [0.08, 0.51]. Across analyses that control for estimated publication bias and non-specific treatment effects, there is substantial overlap in the confidence intervals around the pooled effect, and confidence intervals do not contain zero. As such, the true effect of AET on depressive symptoms is likely to be of small to moderate size.

Using GRADE criteria, the body of evidence for AET in treatment of depression is of moderate quality. The available evidence has been obtained in RCTs, however, there is inconsistency in treatment effects between studies, resulting in a one point reduction of quality rating. Inconsistency is evident in a large spread of confidence intervals in which the true effect of each training paradigm is thought to lie (see Table 2). This is likely to reflect not only the true efficacy of the training paradigms, but also the utilized comparison conditions. Specifically, ascending along the MRC's (2000) trajectory of intervention development, MEST has been evaluated in only one controlled study showing efficacy versus a waitlist. COMET has shown efficacy versus ongoing TAU in two studies. All but one imagery-based CBM study has used a sham CBM control condition, with early indications of potential efficacy but mixed findings from more recent higher-quality RCTs. LRT has been compared against alternative psychological therapies and not shown efficacy in comparison to them, and concreteness training has shown superiority over ongoing TAU but not over relaxation training.

7. Anxiety disorders

There were only two studies that examined AET as a treatment option for anxiety disorders, reflecting the early stage of research in this area. Both studies evaluated a COMET protocol.

7.1. COMET

Two COMET protocols have been developed for use with anxiety disorders – one for Obsessive Compulsive Disorder (OCD) and the other for Panic Disorder. The protocol for Panic Disorder follows much the same format as the protocols used for depression, although positive appraisals regarding self-control are the focus of training (i.e., increased activation of specific episodes where one has felt in control), rather the broad positive self-associations that are trained in the depression-focussed protocols. One RCT has evaluated the COMET protocol for panic (Korrelboom et al., 2014). The trial compared group-based COMET to group-based Applied Relaxation (which involved therapist-led relaxation exercises) in a sample of 143 participants experiencing Panic Disorder. Intent-to-treat analysis indicated a significant reduction in panic-related symptoms in both groups, but no between-group differences at post-training. Although not inferior to Applied Relaxation, COMET is therefore not currently well supported for the treatment of Panic Disorder.

For OCD, the protocol focusses on identifying and improving the accessibility of adaptive themes that counter the patient's obsessions (e.g., I have shaken hands before and not been infected). The OCD focussed protocol differs slightly from previous protocols in that it is self-guided. Training is presented in eight workbook-based sessions, which the individual completes at home. Following encouraging results from an uncontrolled pilot trial (Korrelboom, van der Gaag, Hendriks, Huijbrechts, & Berretty, 2008), an RCT assessed online delivery of the programme relative to a waitlist control group in a sample of 65 individuals with OCD (Schneider, Wittekind, Talhof, Korrelboom, & Moritz, 2015). Results indicated that change in symptoms of OCD did not differ between the COMET and control groups at post-treatment (Schneider et al., 2015).

7.1.1. Synthesis of results

As the identified studies examined symptoms of different disorders, we opted not to calculate a pooled estimate of the effect. The current standing of the literature indicates that further research is needed before conclusions can be drawn on the efficacy of AET programmes for treatment of anxiety. We were however able to consider the quality of evidence for COMET in treating anxiety disorders. Applying the GRADE approach, the quality of current evidence for COMET in anxiety disorders is high, as evidence has been obtained in randomised trials, treatment effects are consistent between studies, and Cochrane risk of bias ratings indicate that the likelihood of bias and imprecision is relatively low in the reviewed studies. The body of evidence for other training paradigms for anxiety disorders is non-existent. Further research is clearly needed to comprehensively assess the effect of AET on symptoms of anxiety disorders.

8. Stress-related disorders

Memory-based therapeutic techniques for stress- and trauma-related disorders have primarily focussed on incorporating new information into the discrete trauma memory (e.g., using memory rescripting in Cognitive Behavioral Therapy; Foa, Keane, Friedman, & Cohen, 2008), rather than on modifying autobiographical processing more generically. However, focus on training processing ability is increasing, both with respect to intrusive memories (outside the scope of this review, but see Holmes, James, Kilford, & Deeprose, 2010) and episodic autobiographical processing. One study – using MEST – with a sample experiencing PTSD met inclusion criteria for this review.

The tendency to retrieve overgeneral categories of events is a cognitive marker of PTSD and predicts poor outcomes for the disorder, over and above symptom level (e.g., Kleim & Ehlers, 2008). There has been one small RCT of MEST for individuals with PTSD (Moradi et al., 2014). In this study, 24 Iranian combat veterans with PTSD were randomised to either five sessions of MEST, or a no-intervention control group who only completed the assessments. Results established that, at post-training, individuals in the MEST group had a significantly larger improvement in the recall of specific memories, and a larger decrease in PTSD symptoms than controls. Both between-group effect sizes were sustained at a three-month follow-up.

According to the GRADE criteria, the quality of the current body of evidence is moderate. The presence of only one, positive published study indicates the possibility of publication bias, and the absence of other studies limits our ability to discount inconsistency. However, the quality rating is improved by a large effect size. Although the results from this study are encouraging and suggest that MEST is a promising avenue through which PTSD symptoms may be targeted, it is important

to note that the reviewed study was a relatively small efficacy trial with a waitlist control. A larger RCT with an active control group is now needed to determine treatment effects, and to also investigate whether change in specificity mediates the effect of training on PTSD symptoms, as has been observed in MEST for depression. The current evidence suggests that AET holds promise in the treatment of PTSD, and is worthy of further investigation.

9. Discussion

Guidance on the development of complex interventions outlines a trajectory of intervention development from initial discovery science through to fully implemented interventions, via a series of iterative phases comprising pilot, definitive and pragmatic clinical trials to determine a treatments efficacy, effectiveness and real-world tractability (MRC, 2000). AET programmes represent a prototypical example of a set of novel, complex interventions translated from basic science and positioned on this trajectory. Current evidence for AET is at the exploratory trial stage (see Fig. 1), as the active components of treatment have been identified, preliminary efficacy has been established in initial feasibility and pilot trials, and in some cases in exploratory trials against an active comparison intervention with relatively small numbers of participants, but in most cases fully-powered larger-scale trials have yet to be conducted. Within this, our review indicates that there is substantial variability in the status of the evidence for AET in mood, anxiety, and stress-related disorders and so here we make a number of recommendations for future research on the efficacy of AET as a psychological intervention, as a function of disorder type.

NICE endorsement as a treatment option requires a strong evidence base, as rated by the GRADE approach (Balshem, Helfand, Schünemann, Oxman, Kunz, Brozek, et al., 2011). Using the GRADE criteria, current evidence for AET for depression is of a moderate strength, and a number of steps need to be taken if AET is to be further developed, before it can be recommended in clinical guidelines. First, it needs to be determined that the training programmes are equivalent to or superior to established psychological interventions. Imagery-based CBM protocols have received the most empirical attention, and we found a medium pooled treatment effect. However, the pooled treatment effect relative to a sham CBM control was small, non-significant, and substantially reduced once estimated publication bias had been corrected. Given the closely-matched nature of the sham training programs used as comparators, the effect sizes relative to sham controls may be interpreted as isolating one (or two) aspects of the training within an additive design rather than efficacy estimates of the CBM protocols. The use of such closely-matched controls reflects the experimental work from which these paradigms originate and is in many ways a strength of this research, although in moving to clinical trials such control conditions may not be optimal for establishing estimates of efficacy, and this has been noted as a challenge for the field (e.g. Blackwell et al., 2015; Hirsch, Meeten, Krahé, & Reeder, 2016). However, current evidence does not allow the conclusion that the specific effects attributable to imagery-based CBM are significantly greater than zero. Post hoc analyses in the current exploratory trials demonstrate a significant effect of imagery CBM on anhedonia specifically, and that treatment effects are larger for those who attempt all training sessions (e.g., Blackwell et al., 2015). This may indicate that further consideration of the active components of imagery CBM (i.e., training length, auditory or pictorial cues, in-person introduction to the training) or the specific outcomes that may be impacted by the programme are needed prior to embarking on definitive trials.

Current, early phase evidence indicates that COMET is superior to continued TAU, and that MEST is superior to waitlist. The moderate quality rating for the body of evidence for each of these training types provides a solid foundation for progression to exploratory trials against existing active treatments. The results from future trials with relatively small numbers of participants (e.g. Dalgleish et al., 2014) will be needed prior to consideration of definitive trials. We recommend that future trials be carried out by diverse research groups (for example, the current evidence base for MEST and COMET in depression rests largely with specific research groups; Luborsky et al., 1999), and also examine moderating and mediating factors through the use of embedded process-outcome studies. Evaluation of the proposed mechanisms of treatment will be vital in moving this field forward. Our review reveals significant effect sizes for changes in the targeted process alongside significant improvement on symptom measures (cf Clarke, Notebaert, & MacLeod, 2014; MacLeod & Grafton, 2016), but with an absence of formal mediation analyses in most cases.

If these ongoing trials demonstrate support for the AET programmes, it will need to be determined whether AET is most efficacious as a stand-alone treatment, or whether it may add value to established interventions for depression, such as CBT. Research on imagery CBM programmes has begun to investigate the addition of training to CBT, but extension of this work to other training paradigms may clarify the optimal presentation of training programmes. As the processing biases targeted by training may contribute to dysfunctional cognitive beliefs, the completion of AET prior to completion of CBT may lay the foundation for cognitive restructuring. For example, increased ability to recall past positive events may help when challenging negative core beliefs. It is proposed that downstream cognitive processes may need time to change before effects are seen on psychological symptoms (Neshat-Doost et al., 2013; Williams et al., 2007). Beginning change in these processes prior to structured therapy may therefore be beneficial. Further, the low intensity format of training suggests that it may be easy to administer to those on waiting lists, or in circumstances where access to clinical psychology is limited, or for those who are difficult to engage in traditional treatment options. Structured assessment of cost effectiveness and the combination of training with currently endorsed treatment options will demonstrate in which context AET will be most beneficial.

The most effective format for the delivery of training also needs to be determined. Most evidence for training efficacy has been obtained using group-format protocols (e.g., Ekkers et al., 2011; Korrelboom et al., 2012; Neshat-Doost et al., 2013; Moradi et al., 2014). However, given the relative simplicity of both MEST and COMET, there is great potential for fully automated delivery. Moving these interventions online is associated with several advantages, including protecting the fidelity with which they are delivered, but also offering a low-cost and potentially scalable option. One COMET programme has already been trialled in a fully online format (although no significant gains were observed; Schneider et al., 2015), and imagery CBM has been consistently delivered fully online (although again, evidence for treatment efficacy has been variable; Williams, Blackwell, Mackenzie, Holmes and Andrews, 2013, 2015). Further exploration of whether group-based programmes may also be efficacious in a self-guided format may help to improve the accessibility of training, but will also determine whether group-based effects are an active component of AET.

The body of evidence for AET in anxiety and stress- related disorders is in need of further development. This review has demonstrated that MEST offers a promising treatment option for PTSD, and an exploratory RCT comparing MEST to an active treatment condition is now needed. Of key importance will be the exploration of the utility of AET for other stress-related disorders (e.g., complicated grief), and whether training types that have received support in other disorders (e.g., COMET) may prove effective in treating stress-related disorders. Current small scale trials indicate that the AET protocols so far studied (COMET) are not efficacious in treating anxiety disorders, although further studies are needed before firm conclusions may be made regarding treatment efficacy. Guidance on the development of novel interventions (MRC, 2000, 2008) suggests that further consideration of the underlying mechanisms through which training may impact anxiety symptoms is needed.

Current research with anxiety-disordered samples has focussed on COMET, but used different training protocols for each disorder type (i.e., protocol targeting representations regarding OCD or panic symptoms). Use of a training protocol which targets maladaptive representations that are evident across anxiety disorders may prove effective. The worry and rumination COMET protocol has been shown to be efficacious in treating depression symptoms, and may therefore prove beneficial to anxiety disorders that are also characterised by these cognitive patterns (e.g., generalised anxiety disorder, panic disorder). Further, the reviewed research utilized obsessive-compulsive and panic disordered samples, and has not yet explored treatment efficacy in generalised anxiety, or social anxiety. This may prove a fruitful area for future research. In fact, some CBM protocols targeting interpretive bias which involve the generation of autobiographical episodes have shown promising results on symptom outcomes in subclinical samples (e.g. Bowler et al., 2012) or pathological cognitive processes in clinical groups (e.g. Generalized Anxiety Disorder, Hayes, Hirsch, Krebs, & Mathews, 2010), and investigation of their efficacy in reducing symptoms in specific clinical samples is warranted. However, at present, it appears that current AET techniques may need to be further refined for use in the treatment of anxiety disorders.

Although we have demonstrated that further investigation of individual training paradigms is warranted, another possibility is the combination of currently promising techniques into a unitary programme with a view to enhancing treatment effects. For example, simultaneously targeting overgeneralization and reduced positivity may produce greater change in symptomology, and training maladaptive memory representations that may be evident across disorders may also enhance transdiagnostic effects. One programme –memory flexibility training (MemFlex) – that is currently under evaluation (Hitchcock et al., 2015) seeks to directly target both memory biases, and the results of this trial will indicate the promise of simultaneous training. Similarly, our review indicates that some AET protocols have been applied across different clinical disorders (e.g., MEST in depression and PTSD, COMET in depression and anxiety disorders). The biases targeted by AET are transdiagnostic indicators of outcome (e.g., overgeneralization predicts prognosis for both PTSD and depression; Bryant et al., 2007, Sumner et al., 2010), thus the potential effect of AET programmes on multiple comorbidities also warrants further investigation, and may offer implications for the treatment of dual diagnoses.

9.1. Limitations

This review may have been affected by some limitations. First, we only examined the impact of training on self-report symptoms, and did not consider effects on diagnostic status, or clinician-rated instruments. Outcomes may differ when symptoms are rated by those other than the individual experiencing the disorder, thus future reviews may wish to consider the impact of training on both clinician-rated and self-report measures. In addition, future reviews may wish to consider the size of the effect when training is compared to active versus non-active control conditions. The current studies have primarily consisted of smaller, exploratory trials, in line with recommendations for the development of complex interventions (MRC, 2000, 2008), and have not always included active control conditions as these comparisons are made later in the phase-based development of new interventions. As the size of a between-group effect will almost invariably be a function of the control condition used, care must be taken in comparing effect size estimates, particularly in relation to the wider field of psychological interventions in which waitlist controls have tended to dominate (e.g. 55 out of 94 studies in one recent meta-analysis of CBT for depression; Cuijpers et al., 2013). Once the research in this area is more developed, future meta-analyses should seek to determine the effect of AET relative to different forms of comparison groups. Continued research in this field, and particularly the completion of fully-powered definitive randomised-controlled trials, will allow future reviews to further investigate the efficacy of AET programmes.

10. Conclusions

Autobiographical episodic training represents a prototypical example of the development of a science-informed and evaluated intervention. The current literature provides a substantial evidence base for further evaluation of training as a treatment option for depression, although further research is needed to determine whether such programmes may also be effective in anxiety and stress-related disorders. At this point in time, autobiographical episodic training is an exciting avenue for future research into accessible, cost-effective, and low-intensity treatment options for those experiencing emotional disturbance.

Funding

This work was funded by the UK Medical Research Council.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.cpr.2016.12.003.

References[3]

Addis, D. R., Hach, S., & Tippett, L. J. (2016). Do strategic processes contribute to the specificity of future simulation in depression? *British Journal of Clinical Psychology*. http://dx.doi.org/10.1111/bjc.12103.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Arean, P. A., Perri, M. G., Nezu, A. M., Schein, R. L., Christopher, F., & Joseph, T. X. (1993). Comparative effectiveness of social problem-solving therapy and reminiscence therapy as treatments for depression in older adults. *Journal of Consulting and Clinical Psychology*, *61*, 1003–1010.

Atkins, D., Best, D., Briss, P. A., Eccles, M., Falck-Ytter, Y., Flottorp, S., ... Zaza, S. (2004). Grading quality of evidence and strength of recommendations. *British Medical Journal (Clinical Research Ed.)*, *328*, 1490–1498. http://dx.doi.org/10.1136/bmj.328.7454.1490.

Balshem, H., Helfand, M., Schünemann, H. J., Oxman, A. D., Kunz, R., Brozek, J., ... Guyatt, G. H. (2011). GRADE guidelines: 3. Rating the quality of evidence. *Journal of Clinical Epidemiology*, *64*, 401–406. http://dx.doi.org/10.1016/j.jclinepi.2010.07.015.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571.

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects: Hoeber medical division.* Harper & Row.

Beck, A. T., Steer, R., & Brown, G. (1996). *Manual for the Beck depression inventory-II.* San Antonio, TX: Psychological Corporation.

Becker, E. S., Vanderhasselt, M. -A., & Vrijsen, J. N. (2015). Memory training in depression. *Current Opinion in Psychology*, *4*, 48–52. http://dx.doi.org/10.1016/j.copsyc.2015.02.004.

Beike, D. R., Brandon, N. R., & Cole, H. E. (2016). Is sharing specific autobiographical memories a distinct form of self-disclosure? *Journal of Experimental Psychology: General*. http://dx.doi.org/10.1037/xge0000143.

Berna, C., Lang, T. J., Goodwin, G. M., & Holmes, E. A. (2011). Developing a measure of interpretation bias for depressed mood: An ambiguous scenarios test. *Personality and Individual Differences*, *51*, 349–354. http://dx.doi.org/10.1016/j.paid.2011.04.005.

Blackwell, S. E., Browning, M., Mathews, A., Pictet, A., Welch, J., Davies, J., ... Holmes, E. A. (2015). Positive imagery-based cognitive bias modification as a web-based treatment tool for depressed adults: A randomized controlled trial. *Clinical Psychological Science*, *3*, 91–111. http://dx.doi.org/10.1177/2167702614560746.

Blackwell, S. E., & Holmes, E. A. (2010). Modifying interpretation and imagination in clinical depression: A single case series using cognitive bias modification. *Applied Cognitive Psychology*, *24*, 338–350. http://dx.doi.org/10.1002/acp.1680.

Bowler, J. O., Mackintosh, B., Dunn, B. D., Mathews, A., Dalgleish, T., & Hoppitt, L. (2012). A comparison of cognitive bias modification for interpretation and computerized cognitive behavior therapy: Effects on anxiety, depression, attentional control, and interpretive bias. *Journal of Consulting and Clinical Psychology*, *80*, 1021–1033. http://dx.doi.org/10.1037/a0029932.

Brand, A. N., Jolles, J., & Gispen-de Wied, C. (1992). Recall and recognition memory deficits in depression. *Journal of Affective Disorders*, *25*, 77–86. http://dx.doi.org/10.1016/0165-0327(92)90095-N.

Brewin, C. R. (2006). Understanding cognitive behaviour therapy: A retrieval competition account. *Behaviour Research and Therapy*, *44*, 765–784. http://dx.doi.org/10.1016/j.brat.2006.02.005.

Bryant, R. A., Sutherland, K., & Guthrie, R. M. (2007). Impaired specific autobiographical memory as a risk factor for posttraumatic stress after trauma. *Journal of Abnormal Psychology*, *116*, 837–841. http://dx.doi.org/10.1037/0021-843X.116.4.837.

[3] *=included in the systematic review.

Christopher, G., & MacDonald, J. (2005). The impact of clinical depression on working memory. *Cognitive Neuropsychiatry*, *10*, 379–399. http://dx.doi.org/10.1080/13546800444000128.

Clarke, P. J., Notebaert, L., & MacLeod, C. (2014). Absence of evidence or evidence of absence: Reflecting on therapeutic implementations of attentional bias modification. *BMC Psychiatry*, *14*, 1–6. http://dx.doi.org/10.1186/1471-244x-14-8.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155. http://dx.doi.org/10.1037/0033-2909.112.1.155.

Coles, M. E., & Heimberg, R. G. (2002). Memory biases in the anxiety disorders: Current status. *Clinical Psychology Review*, *22*, 587–627. http://dx.doi.org/10.1016/S0272-7358(01)00113-1.

Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*, 261–288. http://dx.doi.org/10.1037/0033-295X.107.2.261.

Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry*, *58*, 376–385.

Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis.* New York: Routledge.

Dalgleish, T., Bevan, A., McKinnon, A., Breakwell, L., Mueller, V., Chadwick, I., ... Werner-Seidler, A. (2014). A comparison of Memory Specificity Training (MEST) to education and support (ES) in the treatment of recurrent depression: Study protocol for a cluster randomised controlled trial. *Trials*, *15*, 293–302. http://dx.doi.org/10.1186/1745-6215-15-293.

Dalgleish, T., Navrady, L., Bird, E., Hill, E., Dunn, B. D., & Golden, A. -M. (2013). Method-of-loci as a mnemonic device to facilitate access to self-affirming personal memories for individuals with depression. *Clinical Psychological Science*, *1*, 156–162. http://dx.doi.org/10.1177/2167702612468111.

Dalgleish, T., Williams, J. M. G., Golden, A. M. J., Perkins, N., Barrett, L. F., Barnard, P. J., ... Watkins, E. (2007). Reduced specificity of autobiographical memory and depression: The role of executive control. *Journal of Experimental Psychology: General*, *136*, 23–42. http://dx.doi.org/10.1037/0096-3445.136.1.23.

Dalgleish, T., & Werner-Seidler, A. (2014). Disruptions in autobiographical memory processing in depression and the emergence of memory therapeutics. *Trends in Cognitive Science*, *18*, 596–604. http://dx.doi.org/10.1016/j.tics.2014.06.010.

Debeer, E., Raes, F., Williams, J. M. G., & Hermans, D. (2011). Context-dependent activation of reduced autobiographical memory specificity as an avoidant coping style. *Emotion*, *11*, 1500–1506. http://dx.doi.org/10.1037/a0024535.

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*, 455–463. http://dx.doi.org/10.1111/j.0006-341X.2000.00455.x.

Ekkers, W., Korrelboom, K., Huijbrechts, I., Smits, N., Cuijpers, P., & van der Gaag, M. (2011). Competitive memory training for treating depression and rumination in depressed older adults: A randomized controlled trial. *Behaviour Research and Therapy*, *49*, 588–596. http://dx.doi.org/10.1016/j.brat.2011.05.010.

Foa, E. B., Keane, T. M., Friedman, M. J., & Cohen, J. A. (2008). *Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies.* USA: Guilford Press.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., ... Charney, D. S. (1989). The Yale-Brown obsessive compulsive scale: Development, use, and reliability. *Archives of General Psychiatry*, *46*, 1006–1011. http://dx.doi.org/10.1001/archpsyc.1989.01810110048007.

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: Current status and future directions. *Annual Review of Clinical Psychology*, *6*, 285–312. http://dx.doi.org/10.1146/annurev.clinpsy.121208.131305.

Guyatt, G. H., Oxman, A. D., Kunz, R., Woodcock, J., Brozek, J., Helfand, M., ... Akl, E. A. (2011). GRADE guidelines: 7. Rating the quality of evidence—Inconsistency. *Journal of Clinical Epidemiology*, *64*, 1294–1302. http://dx.doi.org/10.1016/j.jclinepi.2011.03.017.

Hayes, S., Hirsch, C. R., Krebs, G., & Mathews, A. (2010). The effects of modifying interpretation bias on worry in generalized anxiety disorder. *Behaviour Research and Therapy*, *48*, 171–178. http://dx.doi.org/10.1016/j.brat.2009.10.006.

Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*, 1539–1558. http://dx.doi.org/10.1002/sim.1186.

Higgins, J.P.T., Deeks, J.J., & Altman, D.G. (editors). Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions. Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Higgins, J. P. T., & Green, S. (2011). Cochrane handbook for systematic reviews of interventions. Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Hirsch, C. R., Meeten, F., Krahé, C., & Reeder, C. (2016). Resolving ambiguity in emotional disorders: The nature and role of interpretation biases. *Annual Review of Clinical Psychology*, *12*, 281–305. http://dx.doi.org/10.1146/annurev-clinpsy-021815-093436.

Hitchcock, C., Hammond, E., Rees, C., Panesar, I., Watson, P., Werner-Seidler, A., & Dalgleish, T. (2015). Memory flexibility training (MemFlex) to reduce depressive symptomatology in individuals with major depressive disorder: Study protocol for a randomised controlled trial. *Trials*, *16*, 494. http://dx.doi.org/10.1186/s13063-015-1029-y.

Holmes, E. A., Blackwell, S. E., Burnett Heyes, S., Renner, F., & Raes, F. (2016). Mental imagery in depression: Phenomenology, potential mechanisms, and treatment implications. *Annual Review of Clinical Psychology*, *12*, 249–280. http://dx.doi.org/10.1146/annurev-clinpsy-021815-092925.

Holmes, E. A., James, E. L., Kilford, E. J., & Deeprose, C. (2010). Key steps in developing a cognitive vaccine against traumatic flashbacks: Visuospatial tetris versus verbal pub quiz. *PloS One*, *5*, e13706. http://dx.doi.org/10.1371/journal.pone.0013706.

Holmes, E. A., Lang, T. J., & Shah, D. M. (2009). Developing interpretation bias modification as a 'cognitive vaccine' for depressed mood — Imagining positive events makes you feel better than thinking about them verbally. *Journal of Abnormal Psychology*, *118*, 76–88. http://dx.doi.org/10.1037/a0012590.

Holmes, E. A., Mathews, A., Mackintosh, B., & Dalgleish, T. (2008). The causal effect of mental imagery on emotion assessed using picture-word cues. *Emotion*, *8*, 395–409. http://dx.doi.org/10.1037/1528-3542.8.3.395.

Jing, H. G., Madore, K. P., & Schacter, D. L. (2016). Worrying about the future: An episodic specificity induction impacts problem solving, reappraisal, and well-being. *Journal of Experimental Psychology: General*. http://dx.doi.org/10.1037/xge0000142.

Joormann, J., Siemer, M., & Gotlib, I. H. (2007). Mood regulation in depression: Differential effects of distraction and recall of happy memories on sad mood. *Journal of Abnormal Psychology*, *116*, 484–490. http://dx.doi.org/10.1037/0021-843X.116.3.484.

Kleim, B., & Ehlers, A. (2008). Reduced autobiographical memory specificity predicts depression and posttraumatic stress disorder after recent trauma. *Journal of Consulting and Clinical Psychology*, *76*, 231–242. http://dx.doi.org/10.1037/0022-006x.76.2.231.

Kleim, B., Graham, B., Fihosy, S., Stott, R., & Ehlers, A. (2014). Reduced specificity in episodic future thinking in posttraumatic stress disorder. *Clinical Psychological Science*, *2*, 165–173. http://dx.doi.org/10.1177/2167702613495199.

Korrelboom, K., Maarsingh, M., & Huijbrechts, I. (2012). Competitive memory training (COMET) for treating low self-esteem in patients with depressive disorders: A randomised clinical trial. *Depression and Anxiety*, *29*, 102–110. http://dx.doi.org/10.1002/da.20921.

Korrelboom, K., Peeters, S., Blom, S., & Huijbrechts, I. (2014). Competitive memory training (COMET) for panic and applied relaxation (AR) are equally effective in the treatment of panic in panic-disordered patients. *Journal of Contemporary Psychotherapy*, *44*, 183–190. http://dx.doi.org/10.1007/s10879-013-9259-3.

Korrelboom, K., van der Gaag, M., Hendriks, V. M., Huijbrechts, I., & Berretty, E. W. (2008). Treating obsessions with competitive memory training: A pilot study. *The Behavior Therapist*, *31*, 29–36.

Korrelboom, K., van der Weele, K., Gjaltema, M., & Hoogstraten, C. (2009). Competitive memory training for treating low self-esteem: A pilot study in a routine clinical setting. *The Behavior Therapist*, *32*, 3–8.

Lang, T. J., Blackwell, S. E., Harmer, C. J., Davison, P., & Holmes, E. A. (2012). Cognitive bias modification using mental imagery for depression: Developing a novel computerized intervention to change negative thinking styles. *European Journal of Personality*, *26*, 145–157. http://dx.doi.org/10.1002/per.855.

Lester, K. J., Mathews, A., Davison, P. S., Burgess, J. L., & Yiend, J. (2011). Modifying cognitive errors promotes cognitive well-being: A new approach to bias modification. *Journal of Behavior Therapy and Experimental Psychiatry*, *42*, 298–308. http://dx.doi.org/10.1016/j.jbtep.2011.01.001.

Luborsky, L., Diguer, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., ... Schweizer, E. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice*, *6*, 95–106. http://dx.doi.org/10.1093/clipsy.6.1.95.

MacLeod, C., & Grafton, B. (2016). Anxiety-linked attentional bias and its modification: Illustrating the importance of distinguishing processes and procedures in experimental psychopathology research. *Behaviour Research and Therapy*. http://dx.doi.org/10.1016/j.brat.2016.07.005.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*, 17–24.

Matt, G. E., Vázquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review*, *12*, 227–255. http://dx.doi.org/10.1016/0272-7358(92)90116-P.

Mathews, A., & Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of abnormal psychology*, *109*, 602–615. http://dx.doi.org/10.1037/0021-843X.109.4.602.

Medical Research Council (2000). *A framework for development and evaluation of RCTs for complex interventions to improve health.* London: Medical Research Council.

Medical Research Council (2008). *Developing and evaluating complex evaluations: New guidance.* London: Medical Research Council.

Messer, S. C., Angold, A., Costello, E. J., Loeber, R., Van Kammen, W., & Stouthamer-Loeber, M. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development. *International Journal of Methods in Psychiatric Research*, *5*, 251–262.

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*, e1000097. http://dx.doi.org/10.1371/journal.pmed.1000097.

Moradi, A. R., Moshirpanahi, S., Parhon, H., Mirzaei, J., Dalgleish, T., & Jobson, L. (2014). A pilot randomized controlled trial investigating the efficacy of memory specificity training in improving symptoms of posttraumatic stress disorder. *Behaviour Research and Therapy*, *56*, 68–74. http://dx.doi.org/10.1016/j.brat.2014.03.002.

Morgan, J. (2010). Autobiographical memory biases in social anxiety. *Clinical Psychology Review*, *30*, 288–297. http://dx.doi.org/10.1016/j.cpr.2009.12.003.

Morina, N., Deeprose, C., Pusowski, C., Schmid, M., & Holmes, E. A. (2011). Prospective mental imagery in patients with major depressive disorder or anxiety disorders. *Journal of Anxiety Disorders*, *25*, 1032–1037. http://dx.doi.org/10.1016/j.janxdis.2011.06.012.

Moscovitch, D. A., Gavric, D. L., Merrifield, C., Bielak, T., & Moscovitch, M. (2011). Retrieval properties of negative vs. positive mental images and autobiographical memories in social anxiety: Outcomes with a new measure. *Behaviour Research and Therapy*, *49*, 505–517. http://dx.doi.org/10.1016/j.brat.2011.05.009.

Neshat-Doost, H. T., Dalgleish, T., Yule, W., Kalantari, M., Ahmadi, S. J., Dyregrov, A., & Jobson, L. (2013). Enhancing autobiographical memory specificity through cognitive training: An intervention for depression translated from basic science. *Clinical Psychological Science*, *1*, 84–92. http://dx.doi.org/10.1177/2167702612454613.

National Institute for Health and Clinical Excellence (November 2012). The guidelines manual. London: National Institute for Health and Clinical Excellence. Available from: www.nice.org.uk

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology*, *109*, 504–511. http://dx.doi.org/10.1037/0021-843X.109.3.504.

Peterson, C., Semmel, A., von Baeyer, C., Abramson, L. Y., Metalsky, G. I., & Seligman, M. E. P. (1982). The attributional style questionnaire. *Cognitive Therapy and Research*, *6*, 287–300. http://dx.doi.org/10.1007/BF01173577.

Raes, F., Williams, J. M. G., & Hermans, D. (2009). Reducing cognitive vulnerability to depression: A preliminary investigation of MEmory Specificity Training (MEST) in inpatients with depressive symptomatology. *Journal of Behavior Therapy and Experimental Psychiatry*, *40*, 24–38. http://dx.doi.org/10.1016/j.jbtep.2008.03.001.

Robinaugh, D. J., & McNally, R. J. (2013). Remembering the past and envisioning the future in bereaved adults with and without complicated grief. *Clinical Psychological Science*, *1*, 290–300. http://dx.doi.org/10.1177/2167702613476027.

Rude, S. S., Wenzlaff, R. M., Gibbs, B., Vane, J., & Whitney, T. (2002). Negative processing biases predict subsequent depressive symptoms. *Cognition and Emotion*, *16*, 423–440. http://dx.doi.org/10.1080/02699930143000554.

Schacter, D. L., & Addis, D. R. (2007). The cognitive neuroscience of constructive memory: Remembering the past and imagining the future. *Philosophical Transactions of the Royal Society: Biological Sciences*, *362*, 773–786. http://dx.doi.org/10.1098/rstb.2007.2087.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience*, *8*, 657–661. http://dx.doi.org/10.1038/nrn2213.

Schneider, B. C., Wittekind, C. E., Talhof, A., Korrelboom, K., & Moritz, S. (2015). Competitive Memory Training (COMET) for OCD: A self-treatment approach to obsessions. *Cognitive Behaviour Therapy*, *44*, 142–152. http://dx.doi.org/10.1080/16506073.2014.981758.

Serrano, J. P., Latorre, J. M., Segura, L. R., Bravo, B. N., Aguilar Corcoles, M. J., Ricarte, J. J., & Gatz, M. (2012). Life review therapy using autobiographical retrieval practice for older adults with clinical depression. *Psicothema*, *24*, 224–229.

Starr, L. R., & Davila, J. (2012). Responding to anxiety with rumination and hopelessness: Mechanism of anxiety-depression symptom co-occurrence? *Cognitive Therapy and Research*, *36*, 321–337. http://dx.doi.org/10.1007/s10608-011-9363-1.

Sumner, J. A., Griffith, J. W., & Mineka, S. (2010). Overgeneral autobiographical memory as a predictor of the course of depression: A meta-analysis. *Behaviour Research and Therapy*, *48*, 614–625. http://dx.doi.org/10.1016/j.brat.2010.03.013.

Torkan, H., Blackwell, S., Holmes, E., Kalantari, M., Neshat-Doost, H., Maroufi, M., & Talebi, H. (2014). Positive imagery cognitive bias modification in treatment-seeking patients with major depression in Iran: A pilot study. *Cognitive Therapy and Research*, *38*, 132–145. http://dx.doi.org/10.1007/s10608-014-9598-8.

Watkins, E. R., Baeyens, C. B., & Read, R. (2009). Concreteness training reduces dysphoria: Proof-of-principle for repeated cognitive bias modification in depression. *Journal of Abnormal Psychology*, *118*, 55–64. http://dx.doi.org/10.1037/a0013642.

Watkins, E. R., Taylor, R., Byng, R., Baeyens, C., Read, R., Pearson, K., ... Shadish, W. (2012). Guided self-help concreteness training as an intervention for major depression in primary care: A phase II randomized controlled trial. *Psychological Medicine*, *42*, 1359–1371. http://dx.doi.org/10.1017/S0033291711002480.

Werner-Seidler, A., & Moulds, M. L. (2011). Autobiographical memory characteristics in depression vulnerability: Formerly depressed individuals recall less vivid positive memories. *Cognition & Emotion*, *25*, 1087–1103. http://dx.doi.org/10.1080/02699931.2010.531007.

Werner-Seidler, A., & Moulds, M. L. (2012). Mood repair and processing mode in depression. *Emotion*, *12*, 470–478. http://dx.doi.org/10.1037/a0025984.

Williams, A. D., Blackwell, S. E., Mackenzie, A., Holmes, E. A., & Andrews, G. (2013). Combining imagination and reason in the treatment of depression: A randomized controlled trial of internet-based cognitive bias modification and internet-CBT for depression. *Journal of Consulting and Clinical Psychology*, *81*, 793–799. http://dx.doi.org/10.1037/a0033247.

Williams, A. D., O'Moore, K., Blackwell, S. E., Smith, J., Holmes, E. A., & Andrews, G. (2015). Positive imagery cognitive bias modification (CBM) and internet-based cognitive behavioural therapy (iCBT): A randomized controlled trial. *Journal of Affective Disorders*, *178*, 131–141. http://dx.doi.org/10.1016/j.jad.2015.02.026.

Williams, J. M. G., Barnhofer, T., Crane, C., Herman, D., Raes, F., Watkins, E. R., & Dalgleish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin*, *133*, 122–148. http://dx.doi.org/10.1037/0033-2909.133.1.122.

Yiend, J., Lee, J. -S., Tekes, S., Atkins, L., Mathews, A., Vrinten, M., ... Shergill, S. (2014). Modifying interpretation in a clinically depressed sample using 'cognitive bias modification-errors': A double blind randomised controlled trial. *Cognitive Therapy and Research*, *38*, 146–159. http://dx.doi.org/10.1007/s10608-013-9571-y.