Dear Author,

Here are the proofs of your article.

- You can submit your corrections online, via e-mail or by fax.
- For online submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and email the annotated PDF.
- For fax submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the journal title, article number, and your name when sending your response via e-mail or fax.
- Check the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- Check the questions that may have arisen during copy editing and insert your answers/corrections.
- Check that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the Edited manuscript.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please do not make changes that involve only matters of style. We have generally introduced forms that follow the journal’s style. Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections within 48 hours, we will send you a reminder.
- Your article will be published Online First approximately one week after receipt of your corrected proofs. This is the official first publication citable with the DOI. Further changes are, therefore, not possible.
- The printed version will follow in a forthcoming issue.

Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL: http://dx.doi.org/[DOI]. If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information go to: http://www.springerlink.com.

Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us if you would like to have these documents returned.
<table>
<thead>
<tr>
<th><strong>Metadata of the article that will be visualized in OnlineFirst</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Please note:</strong> Images will appear in color online but will be printed in black and white.</td>
</tr>
<tr>
<td><strong>Article Title</strong></td>
</tr>
<tr>
<td><strong>Article Sub-Title</strong></td>
</tr>
<tr>
<td><strong>Article Copyright</strong></td>
</tr>
<tr>
<td><strong>Journal Name</strong></td>
</tr>
<tr>
<td><strong>Corresponding Author</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Email</strong></td>
</tr>
<tr>
<td><strong>Author</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Email</strong></td>
</tr>
<tr>
<td><strong>Author</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Email</strong></td>
</tr>
<tr>
<td><strong>Author</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Email</strong></td>
</tr>
<tr>
<td><strong>Author</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Email</strong></td>
</tr>
</tbody>
</table>
Abstract
Depression has been widely associated with a cognitive deficit leading to the negative interpretation of ambiguous information. Recently, cognitive bias modification (CBM) procedures have shown that such negative biases are causally related to emotional vulnerability. However, research using CBM has been notably lacking in depression. This is the first double blind randomised controlled study investigating the effect of cognitive bias modification-errors (CBM-errors), on depression and its influence on mood and resilience to stress. CBM-errors is a new form of cognitive bias modification for interpretation, which targets the full range of cognitive errors as well as interpretation biases. Forty clinically depressed participants were randomly allocated to a positive training or neutral text reading control group. Participants trained to make positive interpretations subsequently interpreted novel ambiguous information in a positive manner compared to controls. The results suggest that a positive cognitive bias can be induced in clinically depressed individuals using a simple computerised intervention. There was little evidence of corresponding benefits in terms of mood or response to stress, suggesting that multiple sessions are likely to be needed to confer symptom related change. A systematic investigation of the optimum number and timing of multiple sessions is now called for.

Keywords (separated by '-')
Cognitive bias modification - Depression - Randomised controlled trial
Dear Author

During the process of typesetting your article, the following queries have arisen. Please check your typeset proof carefully against the queries listed below and mark the necessary changes either directly on the proof/online grid or in the ‘Author’s response’ area provided below.

<table>
<thead>
<tr>
<th>Query</th>
<th>Details required</th>
<th>Author’s response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Please confirm the section headings are correctly identified.</td>
<td>📩</td>
</tr>
<tr>
<td>2.</td>
<td>Lang et al. (2011) has been changed to Lang et al. (2012) so that this citation matches the list.</td>
<td>📩</td>
</tr>
<tr>
<td>3.</td>
<td>References Blackwell and Holmes (2011), Lang et al. (2009), Salemink et al. (2009), Watkins and Brown (2002), Wells and Beevers (2010), Jacobson and Truax (1991) are cited in text but not provided in the reference list. Please provide references in the list or delete these citations.</td>
<td>📩</td>
</tr>
<tr>
<td>4.</td>
<td>McManus et al. (2012) has been changed to van McManus et al. (2011) so that this citation matches the list.</td>
<td>📩</td>
</tr>
<tr>
<td>5.</td>
<td>References Beard and Amir (2008), Hayes et al. (2010), Hirsch et al. (2009), Holmes et al. (2009a, b), Kaltenthaler et al. (2004), Koster et al. (2009), MacLeod et al. (1996), Richards (2004), Schartau et al. (2009), Schmidt et al. (2009), Wenzel and Rubin (2004), Yiend et al. (2011) are given in list but not cited in text. Please cite in text or delete from list.</td>
<td>📩</td>
</tr>
<tr>
<td>6.</td>
<td>Please provide year for the reference Yiend.</td>
<td>📩</td>
</tr>
</tbody>
</table>
Original Article

Modifying Interpretation in a Clinically Depressed Sample Using ‘Cognitive Bias Modification-Errors’: A Double Blind Randomised Controlled Trial

Jenny Yiend · Jong-Sun Lee · Sinem Tekes · Louise Atkins · Andrew Mathews · Manouk Vrinten · Christian Ferragamo · Sukhwinder Shergill

Abstract  Depression has been widely associated with a cognitive deficit leading to the negative interpretation of ambiguous information. Recently, cognitive bias modification (CBM) procedures have shown that such negative biases are causally related to emotional vulnerability. However, research using CBM has been notably lacking in depression. This is the first double blind randomised controlled study investigating the effect of cognitive bias modification-errors (CBM-errors), on depression and its influence on mood and resilience to stress. CBM-errors is a new form of cognitive bias modification for interpretation, which targets the full range of cognitive errors as well as interpretation biases. Forty clinically depressed participants were randomly allocated to a positive training or neutral text reading control group. Participants trained to make positive interpretations subsequently interpreted novel ambiguous information in a positive manner compared to controls. The results suggest that a positive cognitive bias can be induced in clinically depressed individuals using a simple computerised intervention. There was little evidence of corresponding benefits in terms of mood or response to stress, suggesting that multiple sessions are likely to be needed to confer symptom related change. A systematic investigation of the optimum number and timing of multiple sessions is now called for.

Keywords  Cognitive bias modification · Depression · Randomised controlled trial

Introduction  Cognitive theories of depression postulate that depressed individuals have a tendency to interpret ambiguous information in a negative manner (e.g. Beck 1967). Such biases pervade cognitive processing and prevent a realistic appraisal of everyday events, leading to the maintenance of psychopathology (e.g. Yiend and Mackintosh 2004). Despite decades of research into the causes and optimal treatments for clinical depression a chronic shortage of resources precludes readily available treatment. There is an overwhelming “urgency of addressing depression as a public health priority” (Moussavi et al. 2007) and the need for accessible, cost-effective treatments is paramount (Layard 2005). Further compounding the situation, recent research suggests that the prognosis of depression as seen in primary care settings is worse than previously thought (Yiend et al. 2009).

An experimental technique, cognitive bias modification for Interpretation (CBM-I) offers a potentially cost effective and widely accessible solution. Experimental manipulations of cognitive biases have confirmed their causal role in sustaining clinical disorders (Mathews and MacLeod 2005), and interest now focuses on the adaptation of these techniques for potential treatment. CBM-I studies have shown that processing biases can be induced through repeated processing of emotionally ambiguous information that encourages either negative or benign interpretations. This subsequently leads to congruent biases in the interpretation of new material (Mathews and Mackintosh 2000). Potential interventions based on this technique use only non negative or positive inductions to normalise pre-existing negative biases in patients or those with vulnerability to psychological disorders.

Cognitive bias modification methods offer many potential advantages over existing therapist-delivered psychological...
interventions. They are a more convenient, flexible mode of treatment, not requiring meetings with a therapist. They offer the potential for delivery using modern technologies (e.g. internet or mobile phone) and require minimal supervision. They could therefore become highly cost effective and widely accessible. CBM methods are also less demanding and more acceptable to patients than traditional therapies. This is because personal thoughts and beliefs are not directly interrogated and there is no need for social interaction or stigmatising visits to outpatient clinics. Similarly patient insight is not required because CBM seeks to target the underlying maintaining cognitive bias directly and therefore patient engagement is likely to be easier. In sum, CBM methods offer a high gain, low cost treatment option because they can circumvent many of the practical and psychological requirements that disadvantage competing psychological interventions.

The success of CBM methods targeting attentional bias (CBM-A) in anxiety disorders is well established (Hakamata et al. 2010; Beard et al. 2012). For instance, a meta analysis (Beard et al. 2012) reported large effect sizes for attentional bias modification in samples with high levels of social anxiety, GAD, phobias and worry. However other research suggests that CBM-I has significantly larger effect sizes than CBM-A (Hallion and Ruscio 2011). There is little research specifically investigating the use of CBM-I techniques in depression (MacLeod 2012). In their recent meta analysis of CBM studies Hallion and Ruscio (2011) found only three studies on clinical depression and all used attentional bias modification (CBM-A) rather than CBM-I (Baert et al. 2010; Koster et al. 2010; Wells and Beever 2010). This is slightly surprising given that the evidence base for naturally occurring biases in depression is considerably stronger for biased interpretative processes (e.g. Cowden Hindash and Amir 2012; Wisco et al. 2010) than it is for biased attentional effects (see recent meta-analysis of Peckham et al. 2010). Nevertheless, beneficial emotional effects of CBM-A were demonstrated in a sample of mild to moderately depressed students (Wells and Beever 2010) and shown to reduce recurrence risk in patients with remitted depression (Brown et al. 2012). However findings have been somewhat inconsistent when using CBM-A with moderate to severe depression. Some studies have observed an increase in depressive symptoms and others have failed to effectively change, or sustain change in, attentional bias (Baert et al. 2010; Haeffel et al. 2012).

Only two other studies to date, not included in the above meta analysis, have investigated CBM-I in a clinically depressed sample. Blackwell and Holmes (2010) conducted a single case series investigating the impact of 1 week of daily sessions of CBM-I. Four out of seven people demonstrated improvements in cognitive bias or mood immediately after the sessions and this was largely maintained at 2-week follow up. Although promising, single case series are not designed to allow wider generalisation of results. The second study compared the impact of 1 week of daily sessions of positive CBM in thirteen depressed participants, to a matched depressed control group who were exposed to positive and negative information equally (Lang et al. 2012). Significant improvements on a cognitive bias measure, the Scrambled Sentences task (Wenzlaff 1988, 1993), and depressive and intrusive symptoms were reported immediately after positive CBM, with improvements dropping to trend level by 2-week follow-up.

Although promising, the findings of Lang et al. (2012) leave several questions unanswered. Most importantly no measure of biased cognition was given at baseline. Researchers in the field now recognise the importance of assessing the mechanism one is seeking to change, before attempting to change it. There is considerable variance in the level of naturally occurring biased cognition in any sample and without baseline assessment it is impossible to rule out group differences apparently relating to training, which are in fact a result of failed randomisation. This is especially true with small sample sizes where randomisation frequently fails unless some method of minimisation, balancing prognostic factors, is used. It is clear then that there remains a gap in the literature with regard to CBM-I for clinical depression. The current study sought to contribute to filling this gap by conducting a single session double blind randomised controlled trial using a clinically depressed sample and a new version of CBM-I, called CBM-errors (Lester et al. 2011). CBM-errors is designed to target the entire range of cognitive errors first described by Beck et al. (1979).

Cognitive bias modification-errors offers several advantages over alternative CBM-I methods. First, the content was generated from a dedicated exercise to accumulate real world clinical exemplars from the content of therapy sessions (Lester et al. 2011). It therefore has stronger face validity and end user relevance than researcher generated content. Second it is likely that other CBM versions have omitted important types of cognitive error categories, which are nevertheless ubiquitous in clinical settings. CBM-errors targets the full range of inferential biases identified by Beck et al. (1979) and therefore includes categories of inference not previously incorporated within CBM-I content (most notably personalization, see Table 1). Finally, many versions of CBM-I are designed for a specific disorder such as social anxiety and are therefore not ideally positioned for use in depression. By developing a version of CBM which targets cognitive errors rather than a specific disorder, one can bypass traditional diagnostic labels and instead work.
towards a method which targets underlying functional
maintaining mechanisms of potential transdiagnostic appli-
cability (Harvey et al. 2004).
In CBM-errors, training scenarios are used to provide
practice in the benign resolution of thinking errors believed
by many clinicians (e.g. Beck et al. 1979) to contribute to
maintaining emotional disorders. The randomised controlled
trial (RCT) is widely accepted as the gold standard research
method for establishing efficacy of putative new interven-
tions, giving utmost priority to minimising any source of
potential bias. In this sense it is a ‘proof of principle’, rather
than an attempt to establish ‘real world’ usefulness (effec-
tiveness). CBM-errors has previously been shown to pro-
mote positive inferences, reduce vulnerability to stress and
improve self perception of performance in a sample of
healthy volunteers showing elevated levels of negative affect
and cognition (Lester et al. 2011). This intervention aims to
reduce the cognitive errors (Beck et al. 1979) most com-
monly targeted during therapy (Ilardi and Craighead 1999)
as underlying mechanisms that sustain a range of disorders
(Johnson et al. 1992), including depression.
To the standard CBM-errors reported by Lester et al.
(2011) we added a component designed to prompt positive
future-directed cognition. It has been suggested that
depression is characterised by a deficit in processing
positive information rather than an excess of negative
cognitions, especially when projecting into the future
(Stöber 2000). The reduced ability to think positively about
the future arises from a difficulty in accessing mental
representations of those experiences (MacLeod and

<table>
<thead>
<tr>
<th>Possible cognitive error</th>
<th>Definition</th>
<th>Example modification item (clinician generated content, adapted into CBM format designed to counteract the example error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selective abstraction</td>
<td>Focusing on a detail taken out of context, while ignoring other more salient features of the situation and conceptualizing the whole experience on the basis of this fragment</td>
<td>You are a well praised and respected employee. Your boss has trusted you with a large and important project. You are slightly behind schedule and wonder if you will still be … p-a-sed (praised). Do you think your boss will fire you? (no)</td>
</tr>
<tr>
<td>Minimization</td>
<td>Errors in evaluating the significance or magnitude of an event, transforming neutral or positive experiences into negative ones and rejecting positive experiences as not good enough</td>
<td>You have to take an exam before a new company will employ you. You have revised hard but aren’t sure how well you will do. When the results are given you are told you’ve passed and you think this was r-a-suring (reassuring). Was passing the exam down to luck? (no)</td>
</tr>
<tr>
<td>Personalization</td>
<td>Tendency to place all experiences in one of two opposite categories, e.g. flawless or defective rather than viewing them as existing on a continuum. In describing oneself, the extreme negative categorization is selected</td>
<td>You are performing the lead role in a local play, which you have been practising for many months. On the opening night you remember almost all of your lines but stumble over one. You think the audience will think your performance is … m-r-ellous (marvellous). Are you unhappy with your performance? (no)</td>
</tr>
<tr>
<td>Personalization</td>
<td>Tendency to relate external events to oneself when there is no basis for making such a connection</td>
<td>You decide to ring your family to share some good news. When you speak to them you are excited and keen to talk, but they ask if you could phone back later. You think that your family must be … in-eres-ed (interested). Are your family too busy to talk right now? (no)</td>
</tr>
<tr>
<td>Overgeneralization</td>
<td>Errors in evaluating the significance or magnitude of an event, anticipation of extreme adverse outcomes and considering the most unfavorable of all possible outcomes of a situation</td>
<td>Your best friend invites you to lunch. Over coffee you start to have a deep and meaningful conversation about your lives, achievements and goals. You think that everyone else’s life sounds wonderful and that yours is completely … fa-u-lo-s (fabulous). Are you satisfied with how your life has been? (yes)</td>
</tr>
<tr>
<td>Personalization</td>
<td>Drawing a general rule or conclusion on the basis of one or more isolated incidents and applying the concept across the board to related and unrelated situations</td>
<td>You have spent ages plucking up the courage to ask the person you fancy out on a date. When you finally ask, they politely turn you down because they already have a prior engagement. You think you will spend the rest of your life being … b-l-d (bold). Are you pleased you were brave enough to ask the person out on a date? (yes)</td>
</tr>
<tr>
<td>Articulation</td>
<td>Drawing a specific conclusion in the absence of evidence to support the conclusion or when the evidence is contrary to the conclusion</td>
<td>Your partner is very caring and supportive towards you. Last Saturday evening you bickered and argued over something silly and your partner seemed a little irritated. You think that your partner finds you … i-o-ab-l (lovable). Does your partner still care about you? (yes)</td>
</tr>
</tbody>
</table>
Salaminou 2001), which may result in the belief that such events are less likely to occur. Therefore, we reasoned that encouraging positive prospective cognition as part of the CBM-errors intervention could be especially beneficial for clinical depression.

**Methods**

**Participants**

Participants were recruited through a combination of advertisement in King’s College University campuses, online depression support and self-help websites, and GP surgeries within local boroughs. Advertisements asked for ‘people who suffer from depression to take part in a study examining whether changing the way we think about the future has an impact on our thinking style and mood’. Telephone screening was conducted to establish eligibility. Inclusion criteria required participants to meet screening criteria for Major Depressive Disorder (current episode) on the Mini-International Neuropsychiatric Interview (MINI; Sheehan et al. 1998; administered by phone) or score above 15 on the BDI-II (administered by email). Participants were also required to be fluent English-speaking and between the ages of 18 and 70. Exclusion criteria were: receipt of current psychotherapy, changes to medication within the last 2 months, significant Axis I or Axis II comorbidity, previous head injury involving loss of consciousness for more than 3 min, current major physical illness (e.g. heart disease, stroke). Of 182 potential participants responding to advertisement, 111 were not eligible. Of these, 49 were ineligible due to significant comorbidities (37 bipolar, 7 OCD and 5 drug/alcohol abuse); 11 due to medication changes; 22 due to receiving current psychotherapy; 21 due to not meeting screening criteria for MDD and 8 due to major physical illness and/or previous head injury. Of the remaining 71 eligible, a further 31 subsequently declined to participate.

The 40 eligible, consenting participants were aged between 18 and 70, mean = 43.12, SD = 12.2. A CONSORT diagram is shown in Fig. 1.

**Baseline Measures**

**Credibility and Expectancy Questionnaire (CEQ)**

One question (“how successful do you think this session will be in reducing your symptoms?”) was extracted from the CEQ (Devilly and Borkovec 2000) with a scale ranging from 1 (“not at all useful”) to 9 (“very useful”). The question was asked as part of the sociodemographic questions to assess participants’ expectations of the session.

**MINI Clinical Interview (M.I.N.I; Sheehan et al. 1998)**

The M.I.N.I is a short semi-structured diagnostic clinical interview, designed to establish both DSM-IV and ICD-10 diagnoses and to be valid for use by non-clinicians after a brief period of training. The interviewers were given a 1 day training with a psychiatrist, which included a seminar, role-playing and practicing administration of the interview with actors as well as patients attending psychiatric assessment. The M.I.N.I has high internal reliability and test–retest reliability (Sheehan et al. 1998). The Major Depressive Disorder module was used to ascertain the proportion of the sample meeting criteria for current Major Depressive Episode at the time of testing.

**Mood Measures**

The BDI-II (Beck et al. 1996) and the Major Depression Inventory (MDI; Bech et al. 2001) were used to assess the severity of depressive symptoms. Positive and negative affect was measured using the trait version of the Positive and Negative Affect Schedule (PANAS-T; Watson et al. 1988). Visual analogue scales (VAS; Aitken 1969) measured state “sad” or “anxious” mood on two separate 10-cm scales. Mood is measured as a proportion between 0, “not at all” to 1, “extremely”.

**Scrambled Sentences Task (SST; Wenzlaff 1988, 1993)**

The baseline measure of biased interpretation was the SST with cognitive load. It consisted of 20 scrambled sentences of six words each which, when reordered, permitted either a positive or negative sentence formation. Although the SST may not be a ‘process pure’ measure of interpretation, it was chosen as a reliable task, related to interpretation, which has an established sensitivity to individual differences in depression. For example, the negativity bias on the SST is a known predictor of depressive symptoms (Rude et al. 2003). Following a practice, participants were instructed to use five out of the six words in each list to create a grammatically correct sentence (e.g. from ‘winner am born I loser a’, ‘I am a born winner’ or ‘I am a born loser’). Participants were required to remember a six-digit number while performing the task.

**Stressor Task**

A video task was used to assess baseline resilience to stress. Two mildly stressful video clips of life-threatening accidents (approximately 1 min each) were taken from a real life documentary. Participants saw clip 1 at baseline and clip 2 at test, with order of presentation counterbalanced across participants, within each group. Clips were
taken from those used in previous studies (Hoppitt et al. 2010a, b). Two 10 cm visual analogue scales (Aitken 1969) were used before and after the video, with adjectives “sad” or “anxious”, as described above. An index of mood change was calculated by subtracting ‘after’ from ‘before’ values.

Intervention

Active Condition

CBM-error training used 72 items from Lester et al. (2011) divided into four blocks of 18 items each, with an optional rest between blocks and included approximately equal numbers of each type of error. Blocks 1 and 2 were non-negative, whereas 3 and 4 were overtly positive (see Mathews et al. 2007; Lester et al. 2011). Table 1 gives the range of errors targeted and provides examples of both source and modification materials.

Item content covered seven categories: academic, family, mood and health, relationships, social activities, hobbies and work. In order to maximise the personal relevance of the intervention all participants were asked to indicate at screening which one of these categories was least relevant to them. For those subsequently assigned to the active condition, randomisation included automatic allocation to a personalised version of the intervention programme, that omitted all items in a participant’s least relevant category. In addition, a picture related to the topic of each passage was added to every trial, after debriefing comments from pilot work suggested that this would assist participants to imagine themselves in the relevant situation.

Participants were presented with 72, three-line scenarios appearing one sentence at a time that was followed by a positive word fragment, which resolved the ambiguity of the descriptions (for full description of training format see Yiend et al. 2005). A question forcing positive response (requiring yes/no response) appeared, which was reinforced by providing feedback to question (correct/incorrect). See Fig. 2 for an intervention trial example.

Frame 3 of Fig. 2 highlights how we attempted to control participants’ engagement in training items by providing a fixed 5 s imagery period on each trial, together with a future directed sentence stem (whose content was uniquely related to each trial). This was designed to constrain and direct participants’ engagement with processing the positive meaning of ambiguous training passages. This new addition to our previous CBM-errors intervention was also designed to prompt positive future-directed cognition and comprised the instruction ‘Now imagine the situation that you have just read and what happened next…’ followed by a suggested positive continuation sentence and a fixed 5 s pause to allow processing of the positive continuation. The total duration of training was approximately 45 min.

---

1 Arbitrary inference, selective abstraction, minimization, dichotomous thinking, overgeneralisation, personalization, catastrophising/magnification.

---

**Fig. 1** CONSORT diagram illustrating the flow of participants through each stage of the trial.
Control Condition

The control condition followed an identical trial design and procedure to the intervention, including all imagery components, pictures, and imagination stem and future thinking, as shown in Fig. 2. The only difference was the content of the reading items which were unambiguous and emotionally neutral (e.g. ‘You turn the kettle on and wait for the water to boil. You get a teabag out of the tin, which you put into a mug, and pour the boiling water onto the teabag. Next, you add the m_ _k (milk). Have you made a cup of tea?’).

Primary Outcome

Similarity Ratings Test (SRT; Mathews and Mackintosh 2000)

Our chosen primary outcome measure was the SRT due to its widespread previous use in interpretation bias modification studies. It is usually given once only due to the surprise element inherent in the task. The present version consisted of 10 novel ambiguous test items taken from Lester et al. (2011) involving potential cognitive errors. The task followed the same format as that described in detail previously (Mathews and Mackintosh 2000; Yiend et al. 2005), but in brief this comprised two parts: the presentation of novel ambiguous items and the recognition test. Ambiguous items were similar in form to the training items, but novel in content and each had a designated ‘title’. Below the title, three sentences described an emotionally ambiguous social situation, ending in a word fragment completion followed by a question both of which—unlike the training items—maintained the ambiguity of the preceding text. Participants were thus able to apply their own spontaneous interpretation to the meaning of these test passages. For example:

The wedding reception

Your friend asks you to give a speech at her wedding reception. You prepare some remarks and when the time comes, get to your feet. As you speak, you notice some people in the audience start to…l_ _gh (laugh). ‘Did you stand up to speak?’ (factual question, correct answer ‘Yes’).

Ten such items were shown consecutively. The second part of the SRT task, the recognition test, followed immediately and involved the presentation of corresponding title-sentence pairs. Participants rated each sentence on a Likert scale to indicate its similarity to the previous passage (1 = ‘very different’, 4 = ‘very similar’). Each title appeared four times, alongside four different types of sentence: two target sentences matched the positive and negative meanings of the previous passage respectively (these probed participants’ spontaneous interpretation of the ambiguous passage) and two foil sentences did not match the passage directly, but were positively and negatively valenced (these measured response bias and assessed any valence priming effects of training). In the example given above, the following sentences were rated:

The wedding reception

As you speak, people in the audience find your efforts laughable. (negative target)

As you speak, people in the audience laugh appreciatively. (positive target)

As you speak, some people in the audience start to yawn. (negative foil)
As you speak, people in the audience applaud your comments. (positive foil)

Secondary Outcomes

Two additional, parallel versions of the Scrambled Sentences Task and video stressor task, were used as secondary outcomes. Versions were counterbalanced across participants within each Group.

Follow Up Measures

Automatic Thoughts Questionnaire-Revised (ATQ-R; Hollon and Kendall 1980)

The ATQ-R was chosen as measure of biased cognition suitable for use at follow up. It is a self report instrument which assesses the frequency of negative cognitions related to the self. It contains 40 items, 10 positive (e.g. “I am proud of myself”) and 30 negative (e.g. “I’m worthless”) rated on a 5-point (1 = not at all, 5 = all the time) indicating frequency of occurrence over the past week.

Other follow up measures assessed mood and symptoms using the BDI-II, PANAS and MDI measures which are described above.

Procedure

Participants were randomly assigned to either the experimental (N = 19) or control arm (N = 21) of the study using a procedure of randomisation with minimisation on sociodemographic factors and screening instrument (MDI, BDI) scores (Pocock and Simon 1975). A blinded researcher, outside the research team, used a randomisation computer programme to automatically allocate participants to pre-specified conditions, which included task counterbalancing. Participants were blind to their condition of assignment. They were told that: We are testing two procedures to see whether either is effective in reducing depressed mood. You will be assigned to one of them. Similar procedures have sometimes been helpful for non-depressed volunteers, but neither has yet been shown effective in clinical trials. This rationale was used by the team and in all written documentation associated with the study, following ethical approval. All programmes and materials were arbitrarily coded, with relevant randomisation codes being communicated to the team prior to each session. The intervention/control procedure was administered by one researcher in the team, while two independent researchers shared equally the administration of pre and post measures. Using these procedures all members of the research team, as well as participants themselves, were blinded to the assigned condition, and remained so until after the final follow up questionnaire had been returned.

Participants attended a single session conducted between February and April 2012 at the Institute of Psychiatry, London. After consent procedures, sociodemographic questions included level of education, imagination type; ‘is your imagination primarily auditory/visual (i.e. image-based) or verbal?’ as measured by the CEQ and diagnostic interview (including age of onset). Thereafter participants completed baseline measures: BDI-II, MDI, PANAS-T in a fixed random order, followed by the SST and stressor task given in counterbalanced order across participants. Participants then received their allocated condition. Both were presented on a 15-in. laptop screen using E prime version 2.0 and lasted approximately 45 min. Thereafter participants completed a 5 min emotionally neutral and unrelated filler task, involving reading neutral text and answering questions. This was administered to allow dissipation of any short-term mood effects induced by the intervention or control. Finally outcome measures were completed: the SRT followed by VAS, SST and video stress task in counterbalanced order. Participants were informed that debriefing and payment would be given after the 4-week follow-up. The duration of the entire session was approximately 2 h.

The mean duration from testing to follow up was 28.97 days (SD = 8.07, range = 20–65). Thirty four of 40 (85 %) participants returned the following paper based measures by post: BDI-II, MDI, PANAS-T, ATQ-R. Participants also completed a feedback form including questions regarding the aims of the study and how much they liked/thought the session helped them, rated on a scale ranging from 1 to 7 (1 = not at all/really disliked, 7 = extremely/really liked). The full debrief and unblinding procedure took place once data analyses were completed.

All procedures and measures were approved by the King’s College London Psychiatry, Nursing and Midwifery Research Ethics committee and participants were given £20 for their participation.

Results

Analytical Approach

Data was analysed using the IBM Statistical Package of Social Sciences (SPSS) version 20.0. Prior to analysis, all data were screened for missing values and outliers. Variables were checked for the assumptions of parametric testing prior to applying statistical tests. Normality was assessed using Kolmogorov–Smirnov tests, visual inspection of histograms and boxplots. Skewness and kurtosis...
values were scrutinised for each dependent variable. Levene’s Test was used to assess and, if necessary, adjust for homogeneity of variance in pairwise comparisons.

Participants

The M.I.N.I (Sheehan et al. 1998) diagnostic interview revealed that across the whole sample 38 out of 40 (95 %) met full diagnostic criteria for current episode MDD at the time of testing. The two not meeting full diagnostic criteria were excluded from all analyses. A further two participants were observed by researchers during the session to be greatly distracted and not engaging correctly with tasks to the extent that the validity of their data was compromised. An a priori decision was therefore made to exclude them from all analyses. 2 The final sample therefore comprised 36 participants, 17 and 19 in CBM-errors and control groups respectively. As shown in Table 2, these groups were comparable at baseline on all measures. Importantly, groups were matched at baseline on interpretation bias as measured by the Scrambled Sentences Task (see Table 2).

Moreover, scores indicated the presence of a clear negative bias (more sentences resolved in the negative direction than positive in both groups), as would be expected in a clinically depressed sample. Overall mean BDI-II score was 30.06 (SD = 7.44, range = 17–45.50), which indicated a severe level of depressive symptoms (Beck et al. 1996). Screening data showed that one participant had multiple sclerosis and hypoglycaemia.

State Mood

State mood was measured prior to any of the outcome measures to ensure group differences could not be considered a secondary consequence of any immediate mood induced effects of the active intervention. Groups were comparable in state sadness, t(33) = − .77, p = .448 (Experimental M = .32, SD = .22, Control M = .37, SD = .20) and anxiety t(33) = − .90, p = .373 (Experimental M = .25, SD = .17, Control M = .32, SD = .22).

Primary Outcome: Similarity Rating Test

Mean recognition ratings from the Similarity Rating Test were calculated for each participant across the four different conditions: non-error target, error target, positive foil and negative foil. These means were entered into a mixed-

model ANOVA with Group (Experimental, Control) as a between-participants factor, and Sentence Type (Target, Foil) and Sentence Valence (error/positive, non-error/ positive) as a within-participants factor.

The analysis revealed large main effects of Sentence Type F(1,31) = 47.97, p = .0005, ηp² = .61, (target M = 2.3, SE = .07, foil M = 1.7, SE = .07) and Sentence Valence F(1,31) = 60.70, p = .0005, ηp² = .66 (error/positive M = 1.7, SE = .05, non-error/positive M = 2.3, SE = .08). There was a significant interaction between Group, Sentence Type and Sentence Valence F(1,31) = 6.54, p = .016, ηp² = .17 (Intervention group: non error target M = 2.74, SE = .14, error target M = 1.90, SE = .11, non error foil M = 1.96, SE = .15, error foil M = 1.54, SE = .09. Control: non error target M = 2.48, SE = .12, error target M = 2.06, SE = .10, non error foil M = 1.93, SE = .13, error foil M = 1.35, SE = .07).

A mixed model ANOVA (Group × Sentence Valence) was conducted separately for each level of Sentence Type, targets (which reflect interpretation bias) and foils (which reflect response bias). The analysis revealed a significant Group × Sentence Valence interaction for targets, F(1,31) = 4.21, p = .049, ηp² = .12, but not foils, F(1,31) = 0.87, p = .358 (means given above).

Figure 3 shows that the interaction on target items reflected significantly higher ratings for non error over error interpretations in the intervention group compared to controls. Further follow up t-tests revealed a non significant trend for the intervention groups to endorse non error interpretations more strongly than did controls, t(31) = 1.36, p = .092, intervention group: M = 2.74, SD = .70 versus control group: M = 2.48, SD = .39.

As shown in Fig. 3(a), the presence of the significant cross over interaction (Group × Sentence Valence for targets alone, supported by the higher order Group × Sentence Type × Sentence Valence interaction) indicated a significant group difference on our primary outcome measure. This interaction was clearly carried (irrespective of follow up tests) by the intervention group showing a significantly stronger bias (steeper gradient) than controls in favour of non error compared to error interpretations.

Secondary Outcome Measures

Scrambled Sentences Task. 5

Scoring of this task followed procedures described in previous publications (e.g. Rude et al. 2003) supplemented by personal communication with authors. Total negativity

5 Three participants had missing data on this measure.
and positivity scores were obtained by calculating the proportion of sentences corresponding to each valence and dividing by the total number of completed sentences (maximum possible = 20). Each participant thus obtained a score that was a proportion of 1 (Rude et al. 2003). Only sentences that were exact expected matches of the positive or negative unscrambled sentence were counted. Consequently, and in line with previous studies, a number of ‘errors’ and ‘invalid’ responses were also recorded, for example where sentences were grammatically incorrect, used fewer than 5 words or otherwise did not match the one of the two possible designated sentences. As a result positivity and negativity proportions are not simply direct inverses of each other, but may reflect different directional biases.  

Table 2 Mean participant characteristics or counts by group with standard deviations in parentheses

<table>
<thead>
<tr>
<th>Group</th>
<th>Statistic</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental (n = 17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control (n = 19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sociodemographics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>42 (12.5)</td>
<td>43 (13.0)</td>
</tr>
<tr>
<td>Gender (female)</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Education (university or above)</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>Age of onset</td>
<td>22.7 (13.4)</td>
<td>22.2 (10.6)</td>
</tr>
<tr>
<td>Clinical characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI-II</td>
<td>30.8 (7.8)</td>
<td>29.7 (7.3)</td>
</tr>
<tr>
<td>MDI</td>
<td>32.9 (7.0)</td>
<td>32.3 (5.5)</td>
</tr>
<tr>
<td>PANAS-positive</td>
<td>18.8 (6.4)</td>
<td>20.8 (6.0)</td>
</tr>
<tr>
<td>PANAS-negative</td>
<td>28.8 (7.2)</td>
<td>26.4 (8.2)</td>
</tr>
<tr>
<td>VAS—sad</td>
<td>0.5 (0.3)</td>
<td>0.5 (0.2)</td>
</tr>
<tr>
<td>VAS—anxious</td>
<td>0.4 (0.2)</td>
<td>0.4 (0.2)</td>
</tr>
<tr>
<td>Baseline measures</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SST Positive</td>
<td>.24 (.11)</td>
<td>.37 (.25)</td>
</tr>
<tr>
<td>Negative</td>
<td>.51 (.19)</td>
<td>.42 (.26)</td>
</tr>
<tr>
<td>Vulnerability to stress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAS sad change</td>
<td>0.00 (0.11)</td>
<td>0.00 (0.10)</td>
</tr>
<tr>
<td>VAS anxious change</td>
<td>0.04 (0.13)</td>
<td>0.02 (0.13)</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auditory/visual imaginationa</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Expectation of change (CEQ)</td>
<td>4.9 (2.0)</td>
<td>3.9 (1.5)</td>
</tr>
<tr>
<td>Do you think the session helped you?b</td>
<td>2.9 (1.5)</td>
<td>2.8 (1.9)</td>
</tr>
<tr>
<td>How much did you like session?b</td>
<td>3.7 (1.8)</td>
<td>3.1 (1.2)</td>
</tr>
<tr>
<td>Likeability of session</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 3 Mean similarity ratings for different directions of a targets and b foils by group (Experimental, Control), as measured at test (error bars ± 1 standard error)

593 and positivity scores were obtained by calculating the proportion of sentences corresponding to each valence and dividing by the total number of completed sentences (maximum possible = 20). Each participant thus obtained a score that was a proportion of 1 (Rude et al. 2003). Only sentences that were exact expected matches of the positive or negative unscrambled sentence were counted. Consequently, and in line with previous studies, a number of ‘errors’ and ‘invalid’ responses were also recorded, for example where sentences were grammatically incorrect, used fewer than 5 words or otherwise did not match the one of the two possible designated sentences. As a result positivity and negativity proportions are not simply direct inverses of each other, but may reflect different directional biases.  

6 Alternative, less stringent scoring procedures, as reported in some papers, did not result in significant changes to the pattern of results.  

6FL01
A mixed model ANOVA was conducted on these SST proportion scores, for the negativity and positivity measures separately, with one within-participants factor, Time (pre-, post-intervention) and one between-participants factor, Group (experimental, control). For positivity scores the Group \times Time interaction revealed a non significant trend, \( F(1,31) = 2.82, p = 0.103, \eta^2_p = .08 \), as did negativity scores, \( F(1,31) = 3.17, p = 0.085, \eta^2_p = .09 \). Inspection of the condition means for both types of score (see Table 3), suggested that in both cases the direction of trends was in line with hypotheses, showing an increase in the proportion of positive sentences, and a reduction in negative, from baseline to test in the intervention group only. Hypothesis-driven follow-up t-tests confirmed this, showing no significant changes from baseline to test in controls (all \( ts < 1, ps > 0.3 \)), but a significant increase in positive, \( t(14) = 2.06, p = 0.029 \), and near significant decrease in negative sentences, \( t(14) = 1.68, p = 0.057 \), across time in the intervention group.

We suspected the lack of outright significance in these analyses was due to the compromised power resulting from the loss of data on this measure in our already reduced final sample (3 additional participants did not complete this measure as required at the two time points). We therefore conducted the above analyses on the full sample (\( n = 36 \)) after imputing means for the relevant missing data. This indeed showed clearly significant effects on negativity bias scores, including a significant Group \times Time interaction, \( F(1,38) = 4.92, p = 0.033 \), supporting a significant reduction in negative bias score in the intervention group, \( t(18) = 2.33, p = 0.016 \), but not the control group, \( t(20) = -0.84, p = 0.206 \). However, as imputation of means is a controversial practice, we present these findings purely as context for the preceding analyses.

**Stress Vulnerability**

State mood change in response to the stressful video clips was calculated by subtracting the post viewing score from the pre viewing score separately for VAS sad and VAS anxious scales. Higher scores thus reflected an increase in negative mood. This provided two change index scores for each adjective, one at baseline and the other at test. Group (experimental, control) \times Time (baseline, test) ANOVAs were conducted for each adjective separately. For VAS-sad there was a trend main effect of time, \( F(1,32) = 3.16, p = 0.085, \eta^2_p = .09 \), suggesting that the video elicited less sadness at test than at baseline. The interaction between 

<table>
<thead>
<tr>
<th>Group</th>
<th>Positive</th>
<th>Negative</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>.24 (.11)</td>
<td>.51 (.19)</td>
<td>.34 (.20)</td>
<td>.44 (.23)</td>
</tr>
<tr>
<td>Test</td>
<td>.37 (.24)</td>
<td>.44 (.25)</td>
<td>.37 (.20)</td>
<td>.49 (.20)</td>
</tr>
</tbody>
</table>

Table 3 Mean proportion of positive and negative sentences on the scrambled sentences task at baseline and test (standard deviations in parentheses)

Time and condition was non significant, \( F(1,32) = .69, p = .411 \). For VAS-anxious there was no significant main effect, nor Group \times Time interaction.

**Follow Up**

### Loss to Follow Up

One additional participant was lost to follow up from each group (see Fig. 2). There were no significant differences in average duration of questionnaire return across the groups, \( t(30) = -1.23, p = .228 \) (median = 28, range = 20–41 days).

Analyses of BDI-II, MDI and PANAS scores were conducted using 2 \times 2 Group (experimental, control) \times Time (baseline, follow up) ANOVAs. There were no significant Group \times Time interactions, all \( Fs < 2.5, ps > 0.16 \). Analysis of ATQ-R scores were conducted for positive and negative subscales separately, comparing experimental and control groups using independent t tests. There were no significant group differences, \( ts < 0.7, ps > 0.5 \).

In post hoc analyses we classified participants receiving CBM-errors into responders or non responders according to whether they showed clinically significant reduction in BDI-II scores from baseline to follow-up. Clinically significant reduction was defined as requiring both a shift to a lower depression severity category (i.e. clinically significant change) and a reduction greater than 7.16 on the BDI-II (i.e. reliable change). We then compared responders and non responders across baseline characteristics. Responders showed a near significant trend towards lower levels of baseline depression than non-responders (MDI: 28.67 vs 35.22), \( t(13) = 2.08, p = .057, d = 1.08 \).

8 The reliable change index (RCI) was calculated following the method of Jacobson and Truax (1991). RCI is obtained by multiplying standard error of difference between two sets of test scores (here 3.65, taken from normative date on the BDI-II) by 1.96 (z-value for significance level), which gave a value of 7.154. Thus any change in BDI-II score of 7.16 or above would be unlikely to occur in absence of actual change.

Combining both types of score to give a single bias score for analysis resulted in a Group \times Time, \( F(1,31) = 3.06, p = .09 \).
Discussion

The present study investigated a single session of computerised CBM-errors in a clinically depressed sample using a rigorous double blind randomised controlled methodology. Participants who were assigned to CBM-errors, perceived novel ambiguous situations (designed to invite cognitive errors) in a more positive, error-free manner, than did matched controls who received an unambiguous, neutral reading programme. This confirms the hypothesised potential for CBM-errors to reduce cognitive errors in individuals currently suffering from depression. Non significant trends suggested potential benefits on a related measure of biased cognition, the scrambled sentences task. There were no significant group differences on resilience to stress or on follow up measures of mood, symptoms and self reported cognition.

There is a growing literature evidencing cognitive bias modification (CBM) techniques as potential low cost, theoretically driven interventions for a range of clinical disorders, including anxiety and depression. Of the two main versions of CBM, that which targets attentional biases (CBM-A) and that which targets interpretation biases (CBM-I), the former has arguably generated the widest application. However, in clinical depression CBM-A has not, to date, proven promising (Hallion and Ruscio 2011).

Two studies using CBM-I for clinical depression have yielded better results but have either used single case methods (Blackwell and Holmes 2011) or have not examined whether comparison groups were matched for pre-existing interpretation biases at baseline (Lang et al. 2012). Therefore the main aim of the present study was to show that CBM-I can successfully reduce negative interpretation bias in clinical depression when groups were carefully matched on a range of measures (including pre-existing interpretation bias) and a strict double blind randomised controlled methodology is used. To this extent the aims of the present work were fulfilled. Effect sizes on the primary outcome for cognitive change (the similarity rating task) were large (non-error targets $\bar{r}^2 = .62$, means 2.74 vs. 2.48) and comparable to those found using this technique in a subclinical sample (Lester et al. 2011: 2.92 vs. 2.38), confirming the reliability and translational potential of CBM-errors.

On our secondary measure of cognitive change, the scrambled sentences task, a near significant trend provided tentative evidence that participants receiving the intervention showed an increase in the proportion of positive, and decrease in the proportion of negative, sentences generated, at test compared to baseline, relative to controls who did not show this pattern of change over time. It is likely that our study was underpowered to see clear results on this measure, due to missing data. Indeed a sample size calculation for the negativity bias measure, based on variance estimates from the present data (where observed effect size $f = 0.31$) showed that 80% power to detect significance (assuming alpha set at 0.05) would be achieved by a sample size of 16 per group. This measure has been shown to be a significant predictor of depressive symptoms (Rude et al. 2003) and thus the ability of any intervention to induce change on this measure has encouraging implications for the likely consequences for depressive symptomatology.

Notably, by 4 weeks, there was no longer any evidence of enduring cognitive change from this single session intervention, according to self report responses on the Automatic Thoughts Questionnaire. This is perhaps not surprising, and could be attributed to a number of different factors, but the time course of decay of cognitive effects will continue to be an essential parameter to map out in future work, as will the potential benefits and optimum timing of ‘top-up’ sessions. It is important that early studies, such as this one, include follow up assessments in order to estimate feasibility parameters including response rates, selection of most appropriate outcomes and testing of data collection methods (Lancaster et al. 2004). It is also unclear the extent to which the positive findings here.

In other respects the data presented here leave many unanswered questions. Perhaps the most surprising result was the lack of evidence for group differences in stress resilience, given the precedent of previous single session CBM-I studies that have demonstrated this (e.g. Mackintosh et al. 2006; Wilson et al. 2006). There are several possible explanations. One is that observing effects on emotional reactivity could depend on the comparison conditions used. Studies using similar stress tasks report that participants trained to make threatening interpretations display significant elevations of state anxiety following a stressor, whereas those in benign conditions do not show pronounced changes (Wilson et al. 2006; Lester et al. 2011). Therefore, as noted by Mackintosh et al. (2006), it is possible that positive training compared to neutral control in the current study was insufficiently powerful to elicit any differential stress effects.

Another possible explanation is that the anxiety related content of the stressor task was insufficiently matched to the predominantly dysphoric concerns of our sample. However, mitigating against this explanation, the same stressor task proved sensitive to differences among the mildly depressed sample of the second study reported by Lester et al. (2011). Conversely, a study measuring CBM effects which did use videos content matched to the dysphoric concerns of the sample (depicting scenes of bereavement and bullying) found no significant differences between positively and negatively trained groups (Lang...
An equally likely explanation is that our stress task was simply insufficiently intense to elicit meaningful mood change. Practical constraints dictated that only one very short (1 min) video clip could be used, resulting in smaller than expected mood change scores during the course of viewing, which is likely to have compromised sensitivity. Thus one clear recommendation for future research is to optimise stressor tasks during piloting to avoid ceiling or floor effects, both of which appear to have compromised stress resilience findings in the literature to date.

Notwithstanding the above limitations, this study is distinctive as the first randomised controlled trial of the version of CBM-I, known as CBM-errors. CBM-errors was specifically designed to target and reduce the interpretations reflecting the full range of cognitive errors originally identified by Beck et al. (1979; arbitrary inference, selective abstraction, minimization, dichotomous thinking, overgeneralisation, personalization, catastrophising/magnification) and typically seen and worked within routine clinical practice. Situations inviting possible cognitive errors were generated in earlier work using exemplars taken directly from clinicians and patients (Lester et al. 2011). Thus the content of the material used in CBM-errors has inherent face validity. The present study should be considered a starting point for the further development of CBM-errors as a putative intervention for clinical depression. In this context, the present study is ‘proof of principle’, establishing that this particular technique can effectively reduce cognitive errors in clinically depressed patients. Any ‘real’ benefits to patients in terms of mood, resilience or longer term cognition have yet to be fully tested. We did not anticipate significant change in trait mood or symptoms in a clinical sample after only one short intervention session, which was not presented to patients as a treatment per se. Nevertheless we took the opportunity to measure trait mood and symptoms at test and 4 week follow up, as a useful referent point for future work and because this is recommended in pilot and feasibility studies such as this (Lancaster et al. 2004) It is likely that experimentally induced improvement in maladaptive cognitive mechanisms would require multiple sessions of implementation to accrue lasting experiential benefits observable to depressed patients, as might be expected with any other psychologically driven therapy.

While most CBM experts would agree with the need for multiple sessions to effect symptom change, as yet there has been no attempt to systematically investigate how many and of what duration would be optimum. To date studies have somewhat arbitrarily chosen the number, timing and length of sessions, without any clear evidence base. These have typically ranged from a session a day over eight consecutive days (Salemink et al. 2009) to four sessions over a 2 week period (Mathews et al. 2007). Related to this is the question of the likely decay of induced cognitive effects, its timecourse and it the optimum approach to ‘top up,’ should that be needed. In addition, optimising and improving the strength of transfer effects of CBM techniques will become an increasingly important issue, as translational work continues. Other evidence suggests that behavioural interventions can be more effective at instigating belief change than purely cognitive ones (van McManus et al. 2011) which raises another possible direction for future CBM research. As with face to face cognitive therapies, one way of improving efficacy and transfer may be to add supplementary behavioural reinforcement. We would advocate, as one of the next priorities for the CBM field, a systematic approach to investigating these questions (number timing and length of sessions; need for ‘top up’; improving effect sizes and transfer) and establishing an appropriate evidence base for the further implementation of translational studies.

The literature also shows that there are often important individual differences in treatment response (e.g. van Doorn et al. 2012). This is likely to be the case with CBM techniques and future work needs to pay closer attention to this. Indications from our post hoc analyses were in line with previous research suggesting CBM interventions yield most benefit for those with milder depression levels (Baert et al. 2010). In our sample a near significant trend suggested the same. The cognitive deficits associated with depression are well known (e.g. Watkins and Brown 2002) and although one attraction of CBM is its relatively low cognitive demands compared to traditional psychological therapies, it may nevertheless be proportionately more difficult to engage those with more severe depression. Thus one might usefully focus on simplifying and shortening the cognitive demands of text based CBM while seeking to maintain the key mechanism of bias manipulation. One way to do this would be reducing item length to one sentence; indeed pilot work in our laboratory suggests this is feasible. Another way might be the use of auditory presentation as some are already exploring (Holmes and Mathews 2005, 2010). Our participants, severely depressed but not receiving treatment, represent a wider population of individuals for whom a more accessible, cost-effective, intervention such as CBM could be enormously beneficial.

In summary, this study reports the first double blind randomised controlled trial investigating the effect of CBM-errors in a clinical sample. Our data suggest that a positive cognitive bias can be induced in clinically depressed individuals using a simple computerised intervention. That there was little evidence of corresponding benefits in terms of mood or response to stress suggests that multiple sessions are likely to be needed to confer
symptom related change. This in turn raises the important question of how many sessions, over what timescale and for what duration will be optimal? To date, the field has approached this aspect, of translating a basic experimental manipulation into a clinical intervention, in a rather ad-hoc fashion. A systematic investigation of these questions is now called for.

References


