In press, *Journal of Abnormal Psychology* [Special section on cognitive bias modification]

Seeing the bigger picture: Training in perspective broadening reduces self-reported affect and psychophysiological response to distressing films and autobiographical memories

Patricia E.S. Schartau

Tim Dalgleish

Barnaby D. Dunn

Medical Research Council Cognition and Brain Sciences Unit

RUNNING HEAD: Training cognitive appraisal

Abstract

Appraising negative experiences in ways that reduce associated distress is a key component of successful emotion regulation. In 4 studies we examined the effects of systematically practicing appraisal skills using a computer-mediated cognitive bias modification (CBM) methodology. In Studies 1 to 3, healthy participants practiced applying appraisal themes linked to 'seeing the bigger picture' to a series of distressing training films, either during each film (Study 1), or immediately after each film (Studies 2 and 3). Control participants watched the same films with no appraisal instructions. Participants who practiced appraisal, compared to controls, exhibited reduced levels of self-reported negative emotional (Studies 1 to 3) and electrodermal (Study 1) responses to a final test film which all participants were instructed to appraise. In Study 4, a comparable effect of appraisal practice was found using distressing autobiographical memories, for participants with higher levels of negative affect. Appraisal practice also led to reduced intrusion and avoidance of the target memories in the week post-study, compared to pre-study levels, and relative to the no-practice controls. The findings are discussed in terms of the broader literature CBM.

Key words: emotion regulation, training, appraisal, cognitive bias modification, CBT

Introduction

Cognitive theories of emotional disorders contend that systematic biases in the way emotional information is processed play a key role in the onset and maintenance of mental ill health (Williams, Watts, MacLeod, & Mathews, 1997). Such theories have received impressive support from two parallel streams of research. In the laboratory, the existence of the proposed biases has been demonstrated across a broad range of clinical conditions and cognitive domains (Williams et al., 1997). In the research clinic, cognitive therapies that seek to reverse cognitive biases deemed to be at the heart of a particular disorder have been developed and their efficacy demonstrated (Beck, 2005).

Historically, laboratory research into cognitive aspects of emotional disorders has focused on demonstrating the existence of cognitive biases and on painting a detailed picture of their boundary conditions and psychological correlates. It has generally then been the job of clinical researchers to translate this basic science into novel therapeutic techniques and interventions. Recently, however, there has been a shift in this historical dialectic with the development of laboratory-based techniques to manipulate the cognitive biases thought to underlie emotional disorders using systematic training methodologies. This program of research into 'cognitive bias modification' (CBM; Koster, MacLeod, & Fox, in press) consequently promises to have a direct clinical application, as opposed to simply providing a springboard for the development of more traditional therapeutic treatments.

CBM methodologies seek to expose participants over a series of trials to a particular way of processing the emotional information presented to them. The rationale is that systematic exposure of this kind will promote a shift in way the participant will process similar information in the future. In CBM participants need not be aware of the particular contingencies to which they are exposed. The rationale is that systematic and repeated practice in the context of these contingencies will bring about bias modification. To illustrate,

Mathews and Mackintosh (2000), in a series of studies, constrained their participants to interpret, repeatedly and systematically, a sequence of ambiguous emotional scenarios in a predominantly negative fashion. Following such training, participants were presented with novel ambiguous scenarios to interpret as they liked. The results revealed that the trained participants had acquired the intended negative interpretation bias, relative to non-trained controls. Furthermore, the acquired bias appears to last for at least several days (Yiend, Mackintosh, & Mathews, 2005). Importantly, acquisition of such biases appears to result in increased vulnerability to develop an anxious mood in response to a subsequent stressor (Wilson, MacLeod, Rutherford, & Mathews, 2006), indicating a causal role for such biases in the etiology of psychological distress.

Similar CBM methods can of course be used to induce positive biases (e.g., Holmes, Mathews, Mackintosh & Dalgleish, 2006), hence their clinical promise. The potential clinical advantages of CBM are twofold. First, CBM techniques can bring about change in cognitive biases that more traditional CBT interventions might usually only hope to influence indirectly; e.g. attentional bias for threat (MacLeod, Rutherford, Campbell, Ebsworthy & Holker, 2002). Secondly, CBM techniques can be self-administered by patients as a form of low-intensity 'therapy' as part of a stepped-care approach (Haaga, 2000), or as a clinical adjunct while on a waiting list for more traditional psychological therapy, at a relatively small cost to health service providers.

To date, CBM studies have imposed very close constraints on the ways in which participants process the information on each training trial. So, for example, in CBM for interpretation (CBM-I) participants are presented with training trials that describe ambiguous specific events whose ambiguity is only resolved as either positive or negative by the final word in the event description (e.g. Mathews & Mackintosh, 2000). The rationale is that trained participants will acquire a generalized tendency to seek interpretations of a particular

valence across a broad range of situations, including those with little overlap in content with the training scenarios.

An alternative is to take the reverse approach by ascertaining what the broader 'rules' underlying any generalized interpretation tendency might look like and then having participants practice using those rules as templates to generate specific interpretations tailored to each of a range of different situations. The prediction would be that practicing such generative skills would facilitate biased processing of novel situations and stimuli. To date, CBM has not utilized this alternative approach and that was consequently the aim of the studies reported here.

Our first challenge was to identify the broad interpretation rules (henceforth 'appraisal themes') that we wanted participants to use as templates to generate specific interpretations (henceforth 'local appraisals'). To do this we used the extensive cognitive behavior therapy (CBT) literature as a starting point. CBT, across a range of clinical problems, lays out *dysfunctional* appraisal themes, or 'cognitive distortions', that are hypothesized to act as templates for the generation of local appraisals to a discrete negative experience and that serve to extrapolate that experience across broader life domains (Beck, 1976; 2005). For instance, in depression a single experience of rejection might be overgeneralized to an appraisal that 'nobody likes me'. Similarly, following trauma, the experience of early symptoms such as intrusive images or nightmares might be interpreted as 'I have lost my mind and will never be okay again' (Dalgleish, 2004; Ehlers & Clark, 2000). There is a wealth of evidence that negative local appraisals such as these are a source of marked distress across different clinical conditions (Power & Dalgleish, 2008) and it is central to the CBT model that such cognitions represent a key maintenance process for emotional disorders (Beck, 1976; 2005). The emotive power of negative appraisals has also been demonstrated across many studies in the mainstream cognition-emotion literature. This is exemplified by

the recent theoretical and empirical work of James Gross and colleagues (see Gross, 2002) but dates back to prototypical empirical studies in the 1960s (e.g. Lazarus & Alfert, 1964).

Many of the proposed broader dysfunctional appraisal themes in the CBT literature reflect some difficulty on the part of those with clinical problems to 'see the bigger picture', by failing, for example, to recognize that situations are not black or white, or all-or-nothing, or that scenarios change over time, and so on. In the present studies we wanted participants to systematically practice appraising events so as to *reduce* the negative affect they experienced. We therefore selected appraisal themes that focused being better able to 'see the bigger picture' regarding a discrete event by adopting a broader perspective towards it that serves to integrate positive or adaptive information.

To this end, three of the four themes that we selected were: 'every cloud has a silver lining', to encapsulate the notion that even if an experience seems negative this does not mean that everything about it is negative; 'broader perspective', to encapsulate the notion that even if a given experience is negative, there are other more positive and mitigating aspects of life; and 'time heals' to encapsulate the notion that things will seem better once the acute negative experience has passed. The final theme, 'bad things happen', was selected to encapsulate the notion that sometimes it is simply necessary to accept that negative events occur.

This selection of themes was then informally validated by asking a group of healthy volunteers ($N = 30$) to identify, in an open-ended way (i.e. they were not provided with any themes), how they tried to think about difficult situations to make them more acceptable and less distressing. All four themes were independently generated by at least some of the participants. Having selected appraisal themes, our study design was to examine whether systematic practice in using these themes as templates to generate local appraisals conferred

benefits in terms of reduced emotional reactivity, in the first instance to distressing film clips (cf., Gross, 2002)

Study One

In Study 1 participants therefore either practiced constructing local functional appraisals from the aforementioned themes to a series of emotive film clips (the 'training' films) containing real life footage (the Appraise group) or were asked to simply watch these training films in a no-practice control condition (the Watch group). Before (pre-training) and after (post-training) viewing the training films, *both* groups were asked to functionally appraise 'test' emotive film clips. The hypotheses were that the Appraise group would show a significant reduction in self-reported emotional and psychophysiological responses (GSR and HR) between the pre-training and post-training test films, and that this reduction would be significantly greater than any changes in the Watch controls, even though for each test clip *all* participants were using appraisal.

Method

Participants

For all of the studies, participants aged between 18 and 65 were recruited from the Cognition and Brain Sciences Unit panel of community volunteers or via advertisement in University of Cambridge colleges. Exclusion criteria were a previous or current self-reported mental health problem, having suffered a motor vehicle accident (MVA) trauma, or scoring above cut-offs on standard measurements of anxiety and depression (except in the case of Study 4). Depression levels were assessed with the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), using a greater than 15 cut-off (*moderately to severely depressed*; Shaw, Vallis, & McCabe, 1985). Anxiety was assessed using the

Spielberger State Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983), with above 40 on either state or trait anxiety set as the cut-off.

Forty-three healthy volunteers were recruited for Study 1. Two participants' data were set aside as they were unable to apply an appraisal theme to one of the distressing test films, leaving 21 participants randomly assigned to the Appraise group (Mean age = 27.71 years, *SD* = 9.81; 13 females) and 20 participants to the Watch group (Mean age = 28.35 years, *SD* = 9.50; 13 females). Groups were comparable for age, gender, state anxiety, trait anxiety, depression and estimated verbal intelligence (measured using the National Adult Reading Test [NART]; Nelson, 1982), *P*s > .25.

Film Task

The two distressing test films (approximately 2.5 minutes long) employed pre- and post-training depicted real life footage of humans and animals experiencing marked distress. They were selected after piloting a set of 22 distressing films. Pilot participants ($N = 15$) rated the films in terms of clarity, complexity, and uncertainty of outcome, as well as for intensity of the evoked emotions of distress, horror, happiness, fear, disgust, surprise, anger, guilt, shame, helplessness, contempt, sympathy, anxiety and arousal, all on 9-point Likert Scales. The two test films were selected as the most comparable, in terms of clarity, complexity, and uncertainty of outcome, $ts \leq 0.17$, $Ps \geq 0.86$, and because they both evoked marked levels of distress and horror (> 6.0 on the 9-point scales), of comparable intensity, $ts \leq 0.40$, $Ps \geq 0.69$. Test film presentation order was counterbalanced.

For the first test film *all* participants were instructed to employ one or more of the four appraisal themes outlined in the Introduction in order to reduce their emotional reaction. Functional appraisal was described as "changing your feelings about the film by changing the way you think about it", and the four themes were:

1. *Bad things happen* - bad things happen in the world and I need to put them behind me and move on.
2. *Silver lining* - there are usually some good aspects to every situation and it is important to focus on these.
3. *Broader perspective* - bad events are rare overall and lots of good things are happening all of the time.
4. *Time heals* - in the (near) future, this will not seem anywhere near as bad as it does now.

Participants were provided with examples for each theme and allowed to familiarize themselves with using the themes as templates to generate specific appraisals on a distressing practice film, prior to assessment with the first test film.

Following the first test film participants viewed a further six distressing 'training' films (each separated by a one-minute 'thought break' to allow participants to relax and clear their mind)(see Figure 1 for the task timeline). These six 'training' films were drawn from the remaining 20 distressing films piloted in advance. They depicted real-life footage of a range of situations (e.g. war, shooting, accidents) and the piloting revealed that they all evoked significant levels of distress.

Participants in the Appraise group practiced applying the appraisal themes to the 6 training films. In contrast, participants in the Watch group were asked to: "attend to the [training] film fully and do not let your mind drift onto other topics. The film clip may elicit negative emotions such as fear or anger in you, or just general distress. Your task is it to just watch the film and to allow yourself to feel the emotions that the film evokes. We would like you to avoid trying to regulate your emotions in any way."

To return affective responses to baseline, all participants viewed a neutral film after the training films. *All* participants were then instructed to use the four appraisal themes for

the second test film before viewing a happy film at the end of the experiment to normalize affect.

To index participants' affective reactions to the test films, subjective ratings were taken after each film for distress, horror and the other emotions used in the pilot study, as experienced during viewing, rated on 9-point Likert scales ranging from 1 "not at all" to 9 "very much so". Psychophysiological responses (heart rate [HR] and galvanic skin response [GSR]) were also recorded while viewing each film (see below). Affect ratings and psychophysiological responses were compared to a baseline (four minute) rest period prior to each test film. Change scores were computed, subtracting rest ratings and responses from film ratings and responses (Stern, Quigley & Ray, 2001).

Analyses of self-reported emotion focused on distress and horror for a number of reasons. First, we wanted to examine the same emotions across all 4 studies and these two emotions were those most intensively evoked during piloting of the different test stimuli used across the studies. Previous research has shown (e.g. Gross, 1998) that analyses of emotions not intensely and reliably evoked by the emotive stimuli are not sensitive to detecting effects of appraisal. Secondly, previous studies have shown that a more generalized emotion term such as 'distress' or 'negative emotions' is a reliable index of shifts in emotion experience in response to experimental manipulations (Richards & Gross, 2000) and, indeed, the notion of subjective units of distress (SUDs) is one commonly used in clinical research (e.g. Dalgleish & Yiend, 2006). Thirdly, some of the films used in the present series of studies have previously been used as analogue traumatic stress inductions (e.g. Holmes, Brewin & Hennessy, 2004) and distress and horror are the two emotion categories historically associated with such trauma experiences (American Psychiatric Association, 1994). Despite an analytic focus on distress and horror, as already noted we elicited ratings in all of the studies for the range of emotions that we used in piloting in line with Gross (1998). The

reasoning behind this was to "make the key comparison less obvious to participants" (Gross, 1998, p. 229) by using the other emotion terms as distractors.

In order to index participants' compliance with the task instructions for the test films, they were asked to rate how much of the time they had attended to each film or distracted themselves, and how much of the time they had appraised or suppressed their emotions (all rated on 100-point visual analogue scales measured as percentage of time). Additionally, participants were asked two factual questions about the content of each film to ensure that they had attended to the content. To verify that participants had successfully followed the appraisal instructions, they were asked to write down the specific appraisal(s) they had used.

To permit evaluation of whether participants complied with instructions, participants in the Appraise condition were asked to judge after each training film whether they had successfully used at least one of the appraisal themes (rating "yes" or "no" and if yes, giving details of which theme was applied). In addition, participants in the Watch group rated the extent to which they had employed appraisal or any other regulation strategy during the six films (rating "yes" or "no" and if yes, giving details of which strategy(ies) were applied).

Psychophysiology

HR and GSR were measured using a BIOPAC™ MP100 unit running Acqknowledge 8.0 software (BIOPAC, 1997) with one GSR 100B amplifier and one electrocardiogram (ECG) 100B amplifier attached. The GSR amplifier was set to direct current, had a sensitivity of 5 μohm/V, with a 10 Hz low-pass filter and a 0.05 Hz high-pass filter. The ECG amplifier gain was set at 5000, the R-wave detector was switched on, and the filter was switched off. Data were acquired at 200 samples per second, providing a temporal resolution of 5ms.
In order to monitor HR, two disposable Ag-AgCl ECG electrodes with clip-on shielded leads were attached to the top sides of each wrist and a further electrode was attached to the right shin of each participant. To measure GSR, two Ag-AgCl GSR electrodes were attached to the

volar surfaces of the medial phalanges of the second and third finger of the non-dominant hand. BIOPAC GSR paste (with the recommended 0.05M NaCl saturation; Grey & Smith, 1984) was used as the electrolyte. Choice of electrode attachment and sampling was based on published research guidelines for GSR (Dawson, Schell, & Filion, 2000), and ECG (Brownley, Hurwitz, & Schneidermann, 2000). With regard to the psychophysiology data, for any participant whose mean score deviated from the group mean by more than three standard deviations, the original Acqknowledge data file was scanned for movement or other artefacts. If a clear artefact was found, these data were set aside from analyses.

Procedure

Testing for all four studies took place in a quiet, softly lit room, with participants seated in a comfortable chair facing the computer monitor. In this study, participants were asked to attend one assessment session of 1.5 hours. They were screened for exclusion criteria using protocol questions, the BDI and the STAI. After completing other questionnaires not reported here, demographic information was taken, the NART was administered, and psychophysiological electrodes were attached. Participants then completed the appraisal training task, programmed in E-prime software (Schneider, Eschman, & Zuccolotto, 2002), and presented on a personal computer with a 40 cm screen that participants viewed from 60 cm away. Participants in all studies were then debriefed and paid a $10/hour honorarium.

Results and discussion

For all studies reported here analyses addressing a priori hypotheses are directional, with alpha set at 0.05, and other analyses are non-directional.

Instruction compliance

The two groups (Appraise, Watch) differed significantly in terms of the number of appraisals they reported using during the training films (as expected), t (20) = 30.47, $P <$

.001, but not in the extent to which they appraised during the test films (since both groups were instructed to employ appraisal, *t*s < 1. Participants generally reported paying attention (> 93% of the time), appraising much of the time (> 72%), not distracting themselves (< 12% of the time), and not suppressing (< 29%) during the test films. There were no differences between groups on these dimensions, *P*s > .22.

Film task

For self-reported emotions (distress, horror) and psychophysiology, there were no significant differences between groups for either the pre-training baseline rest period, the post-training baseline rest period, the change in baseline ratings from pre- to post-training or the pre-training test film, *Fs* ≤ 1.67, *Ps* ≥ 0.10 (see Table 1).

Analyses involving order of test stimuli (in this case the 2 test films) revealed no significant main effects or interactions including this factor in any of the 4 studies, *Fs* ≤ 2.41, *Ps* ≥ 0.10. Consequently, order is not considered further in this article and analyses without order as a factor are presented.

Change scores from baseline to film for the pre- and post-training test films are shown in Figure 2. A MANOVA examining the effects of appraisal practice with baseline to film change scores for horror and distress as the dependent variables revealed a significant multivariate effect of Time, F (2, 38) = 5.34, $p < 0.01$, $\eta_p^2 = 0.22$, with the emotion change scores being higher pre-training, but not of Group, F (2, 38) = 1.17, $p = 0.32$, $\eta_p^2 = 0.06$. The multivariate Time by Group interaction was also significant, F (2, 38) = 5.40, $p < 0.01$, $\eta_p^2 = 0.22$, consistent with a decrease in negative emotional reactivity post-training in the Appraise, as compared to the Watch group. Univariate output showed that this interaction was significant separately for horror, F (1, 39) = 11.05, $p < 0.01$, $\eta_p^2 = 0.22$, and distress, F (1, 39) = 3.99, $p < 0.05$, $\eta_p^2 = 0.09$, in line with our hypotheses.

Paired samples t-tests were run on each group separately to deconstruct these univariate interactions for distress and horror. For the Watch group, as expected, there were no significant differences pre- to post-training in reported changes in horror or distress, *ts* < 1. For the Appraise group, however, there was the hypothesized significant decrease in pre- to post training change scores for both horror, t (1, 20) = 4.54, $p < 0.001$ (one-tailed), $d = 0.95$, and distress, t (1, 20) = 4.25, $p < 0.001$ (one-tailed), $d = 0.82$.

A similar MANOVA with the change scores for HR and GSR as the dependent variables showed a significant multivariate effect for Time, F (2, 34) = 12.04, $p < 0.001$, $\eta_p^2 = 0.42$, with physiological reactivity decreasing from pre-to post-training, but not for Group, $F < 1$. The multivariate Time by Group interaction effect was again significant, F (2, 34) = 3.99, $p < 0.05$, $\eta_p^2 = 0.19$. However, a significant univariate Time by Group interaction effect was only found for GSR, F (1, 35) = 7.00, $p < 0.01$, $\eta_p^2 = 0.17$, in line with our hypotheses, and not for HR, F (1, 35) = 1.66, $p = 0.21$, $\eta_p^2 = 0.05$, although it was in the anticipated direction.

To deconstruct this univariate interaction term for GSR, paired samples t-tests were run for each group separately. These revealed a significant decrease in GSR change scores over time in both the Watch, t (17) = 2.19, $p < 0.05$, $d = 0.39$, and Appraise, t (18) = 4.58, $p < 0.001$ (one-tailed), $d = 1.32$, groups. Thus, both groups' mean GSR change to the test films decreased from pre- to post-training, but this decrease was more marked in the Appraise group.

The results of Study 1 support our hypotheses that participants who practiced perspective broadening by constructing local appraisals from more general adaptive themes, mostly relating to 'seeing the bigger picture', would show significant decreases from pre-training to post-training in both self-reported emotional reactivity (horror, distress) and psychophysiological response (GSR) to novel emotive test films (with large effect sizes, *ds* >

0.8), that were also significantly greater than any changes over time in control participants who received no practice. These data suggest that systematic practice in constructing functional local appraisals from more general appraisal confers significant benefits, even following a single, relatively brief practice session.

A number of aspects of the Study 1 results merit some discussion. The first is that the findings are unlikely to be a function of simple demand effects influencing emotion self-report, as we were also able to show significantly greater reductions in GSR response in the Appraise group, relative to controls, with a non-significant interaction in the expected direction in HR response. This is particularly important as there is scant evidence in the literature for psychophysiological changes accompanying the use of functional appraisal in individuals in the absence of appraisal practice (e.g. Gross, 1998).

Secondly, it is important to note that the group differences on the test films are not simply a result of the Watch control participants becoming more distressed or sensitized over the course of watching the 'training' films, as there was no significant increase for these participants in terms of their emotional responsiveness from the pre- to the post-training test films (indeed the distress ratings and GSR response of the Watch group decreased; see Figure 2). Furthermore, there was no difference between the Watch group and the Appraise group in their baseline levels of emotion prior to the second test film, indicating that the Watch controls were not in a more emotionally aroused state at that juncture.

Study Two

The main aim of Study 2 was to examine the replicability of the findings from Study 1 of a beneficial effect of practicing perspective broadening. We also wanted to examine whether asking participants to appraise following the emotional films, as opposed to while they were playing, would be effective. Our rationale was that such retrospective appraisal should significantly reduce the cognitive load on participants who, in Study 1, had to watch

and comprehend the film while at the same time trying to construct local appraisals. Our concern was that for some individuals appraising contemporaneously with the emotive event may be sufficiently difficult that they would struggle to apply such a skill reliably across a range of distressing situations. We were therefore anxious to examine the effects of practicing a skill that is potentially more tractable.

Finally, for Study 2 we increased the amount of practice that participants in the Appraise group received from six films to thirteen films and split this across two separate sessions on consecutive days. Our aim was to examine whether this increase in the amount of practice led to stronger emotion-reducing effects on the subsequent test film.

Participants

Thirty-five healthy participants were randomly assigned to either the Appraise group or the Watch group. Exclusion criteria were the same as in Study 1. Three participants' data were set aside as one of them was unable to apply an appraisal theme to one of the distressing test films and two participants asked to take a thirty minute break during the study, which would have confounded the results. The final groups (*n*s = 16) were comparable for age, gender, state anxiety, trait anxiety, and NART estimated verbal intelligence, *P* s > .22.

Film Task

The film task was broadly similar to that used in Study 1 (see Figure 1), with the exception that it took place on two consecutive days since the number of training films was increased from six to thirteen in order to try to augment the effect size. Furthermore, the different sets of instructions were not employed *while* viewing the distressing training films but in a 'thought break' following each film, during which participants in the Appraise group appraised the film and participants in the Watch group simply thought about the film without seeking to regulate their emotions in any way. Furthermore, for the test films pre- and post-training, where all participants used appraisal, this also occurred in post-film thought breaks.

As a result of this methodological change we dispensed with psychophysiology analyses for the test films as both groups were now simply watching these. There was thus no strong hypothesis that the Appraise group should show reduced psychophysiological responsiveness *during* the test films, post-training, compared to control participants, as no appraisal was required at this time.

In the absence of psychophysiological data we opted to assess putative demand effects by asking participants to estimate at study onset how successful they thought either appraisal practice (for the Appraise group) or repeated viewing (for the Watch group) would be in reducing negative emotions elicited during the film clips using a 0 (not at all) to 20 (very much so) scale, following Holmes, Mathews, Mackintosh and Dalgleish (2006). This allowed us to examine whether there was a significant association between estimated and actual change, as one might expect if demand effects were in operation.

As in Study 1, the training trials were separated by one-minute breaks to clear one's mind. The additional seven training films employed in Study 2 were drawn from the remaining distressing films piloted for Study 1, and all elicited significant levels of distress.

Procedure

Participants were asked to attend two assessment sessions on consecutive days (at the same time of day), each lasting one hour. The first session ended after the sixth training clip. The procedure was otherwise identical to Study 1.

Results and discussion

Instruction compliance

The two groups differed significantly in terms of the number of appraisals they reported using during the training films (as expected), t (30) = 10.34, $P < .001$, but not in the extent to which they appraised during the test films (> 73% of the time), ts < 1. Participants

generally reported not distracting themselves (< 11% of the time) or suppressing (< 23%) for the test films and there were no differences between groups on these dimensions, *P*s > .40.

Film task

For self-reported distress, and horror, there were no significant differences between groups for either the pre-training baseline period, the post-training baseline period, the change across baselines, or the pre-training post-film thought break, *P*s ≥ 0.19 (see Table 2).

The MANOVA examining the self-report emotion (horror, distress) change data (see Figure 3a) pre- and post-training showed a significant multivariate effect of Time, F (2, 29) = 10.11, $p < 0.001$, $\eta^2 = 0.41$, with emotion change scores being lower post-training, but not of Group, $F < 1$, and a significant Time by Group interaction, F (2, 29) = 11.69, $p < 0.001$, $\eta_p^2 = 0.45$, consistent with a greater reduction in emotional reactivity over time in the Appraise group, relative to the Watch group, as anticipated. Univariate Time by Group analyses showed significant effects for both horror, F (1, 30) = 11.04, $p < 0.01$, $\eta_p^2 = 0.27$, and distress, F (1, 30) = 19.25, $p < 0.001$, $\eta_p^2 = 0.39$.

Paired samples t-tests revealed no significant change in the Watch group from pre-training to post-training for the change scores for horror, t (1, 15) = 1.83, $p = 0.09$, $d = 0.41$, or distress, t (1, 15) = 0.61, $p = 0.55$, $d = 0.13$. However, there was a significant decrease in the mean emotion change scores in the Appraise group from pre-training to post-training for horror, t (1, 15) = 5.12, $p < 0.001$ (one-tailed), $d = 1.18$, and distress, t (1, 15) = 5.58, $p < 0.001$ (one-tailed), $d = 1.00$.

To formally compare the magnitude of these significant training effects in Study 2 with those in Study 1, we conducted a Group by Time by Study MANOVA on the self-reported emotion change scores. The critical three way interaction was not significant, $F < 1$, indicating that the additional training films had not significantly augmented the appraisal practice effect, even though the controlled effect sizes (comparing the Appraise and Control

groups) had increased notably across the two studies: Study 1 – distress, d = .64, horror, d = 1.06; Study 2 – distress, d = 1.60; horror, d = 1.21.

To assess putative demand effects we examined the associations between participants' estimations of emotion change from pre- to post-training with the actual changes in self-reported emotions, for the two groups. These associations were not significant for either horror or distress (Ps > .21).

Study 2 provides a conceptual replication of Study 1 by showing that systematic practice in perspective broadening leads to significant absolute and relative decreases in reported emotional reactivity (horror, distress). This suggests that the effects of appraisal practice are reliable.

The data also extend the findings of Study 1 by indicating that beneficial effects of appraisal practice can accrue when functional appraisals are applied in a one- minute thought break following exposure to the film, rather than during the film itself. Such immediate retrospective appraisal arguably represents a more tractable skill than appraising at the same time as the emotive stimulus.

We did not collect psychophysiological data in Study 2 due to this shift to post-film thought breaks. We therefore assessed demand effects by comparing participants predictions about emotional change with the actual reported change (Holmes et al., 2006). There was no support for an association between expected and actual change, suggesting that expectancy-led demand effects do not play a key role in the pattern of findings.

There was no statistically significant benefit from increasing the number of training films in Study 2 compared to Study 1, although effect sizes were augmented. However, it is important to note that any increase in the appraisal practice effect as a function of a greater number of training films may have been offset by splitting the films across two sessions and/or by moving the appraisal to a post-film thought break.

Study Three

One potential issue emerging from the first 2 studies is that the control participants have simply been instructed to watch the 'training' films and not to perform any explicit cognitive operations on them. A concern therefore is that any effects of appraisal practice in the Appraise participants may simply reflect practice in cognitive manipulation of the film material, rather than in appraisal per se. To examine this, in Study 3 we included a second control group (the Detachment group) who were asked to think about the emotive training films in terms of how technically challenging they would have been to shoot and, in particular, how as the makers of the films, they would have sought to emphasize the films' emotional impact.

Our inspiration for this particular set of detachment instructions was the work of Gross and colleagues (e.g. Gross, 1998; see Introduction). In a number of these studies, participants are instructed to adopt some form of professional or non-immersed stance to the material (what Gross and colleagues call "reappraisal"). For example, Gross (1998) asked participants to "adopt a detached and unemotional attitude" as they watched film and "to try and think about what you are seeing objectively, in terms of the technical aspects of the events you observe" (p. 227). In each case, this form of detachment led to significant reductions in self-reported emotional distress compared to a simple watch condition and also to an expression suppression condition, in which participants had to try not to show any evidence of emotion on their face.

The focus of the present study therefore was to see if practicing functional appraisal provided significantly greater benefits compared to praticing cognitive detachment - itself an effective emotion regulation procedure (e.g. Gross, 1998). It is important to note that the Detachment group were still asked to *appraise* the pre-and post-training test films in the same way as the Appraise (and Watch) groups. We were therefore examining whether practicing

any form of plausible and emotion-reducing cognitive manipulation would lead to benefits in the ability to apply functional appraisals during a later test film, as a preliminary assessment of whether the systematic practice needs to be appraisal-specific.

Our hypotheses were therefore that the Appraise group would show significant reductions in self-reported emotional reactivity from pre- to post-training, and that these reductions would be greater than for both Watch control group and Detachment control groups.

Method

Participants

Forty-eight healthy participants were randomly assigned to Appraise, Watch, or Detachment conditions (*N*s = 16). The groups were comparable for age, gender, state anxiety, trait anxiety, and NART estimated verbal intelligence, *P*s > .10. No participant's data were set aside following application of the exclusion criteria, which were the same as for Studies 1 and 2.

Film Task and procedure

The film task was identical to that used in Study 2, with the addition of a Detachment group who were instructed to imagine during the thought breaks that they had to shoot the scenes depicted in the training films and identify in each training film (and note down) the scene that would have been the most difficult one to shoot from: a) a technical point of view; and b) from an emotional point of view - in other words, the scene for which it would have been most challenging to illustrate the emotional content for the viewer. The other change was that the number of training films was reduced, relative to Study 2, back to the six films used in Study 1, as increasing the number of films had conferred no significant advantage.

Participants attended a 1 hour 40 minute assessment session. The procedure was otherwise identical to Study 1.

Results and discussion

Instruction compliance

The Appraise group, as expected, reported appraising more during the training films than either the Detachment or Watch groups, where nobody reported any appraisal, F (1, 45) = 894.68, $P < .001$. There were no differences in the extent to which participants reported appraising during the thought breaks after the test films (where all three groups were instructed to appraise), Ps > .05. There was a significant group difference for the time reportedly spent suppressing in the pre-training post-film thought break, F (1, 45) = 4.19, $P < .03$, with post hoc analyses revealing that the Watch group reported suppressing significantly less of the time than the Detachment group, $P < .05$.

For the two test films participants reported paying attention (> 96% of the time), not distracting themselves (< 18 % of the time), appraising (>77 % of the time), and modest use of suppression (< 38%). There were no differences between conditions on these dimensions, Ps > .10.

Film task

The Detachment group reported significantly higher levels of distress than the Appraise group at the first baseline rest (See Table 3), $t = 2.27$, $p < 0.05$, with a trend for a similar difference at the second baseline rest, $t = 1.76$, $p = 0.09$. There were no significant differences between groups in change from first to second baseline, Ps > .27.

These differences in the baseline scores suggested that, despite the random allocation to groups, the Detachment group was more emotionally aroused at study entry and this tended to carry over through the study. To guard against regression to the mean effects, we therefore decided that the most conservative approach was to covary out the average baseline scores for horror and distress (means of baselines one and two) for the key analyses.

Therefore, in order to examine the effects of appraisal practice (see Figure 3b), separate ANCOVAs for changes in response to the test films pre- and post-training for horror and distress, with the mean baseline scores as covariates, were performed. In line with our hypotheses, the results of the two ANCOVAs showed a significant Time by Group interaction for horror, F (2, 44) = 4.88, $p < 0.05$, $\eta_p^2 = 0.18$, and a very strong trend for distress, F (2, 44) = 3.11, $p = 0.05$, $\eta_p^2 = 0.12$. No main effects of Time or Group were found, $Fs \leq 1.85$, $Ps \geq 0.18$.

Planned contrasts within these ANCOVAs of differences in the magnitude of the emotion-change scores across time were computed for the different group pairings. As hypothesized, there were significant differences across time between the Appraise and Watch groups (as in Studies 1 and 2) for both horror, mean difference = 3.37, $p < 0.01$ (one-tailed), and distress, mean difference = 2.38, $p < 0.03$ (one-tailed), with the Appraise group showing a greater reduction in reported emotional reactivity from pre- to post-training, relative to the Watch group. There were also the hypothesized significant differences between the Appraise and Detachment groups for horror, mean difference = 2.25, $p < 0.03$ (one-tailed), and distress, mean difference = 2.83, $p < 0.05$ (one-tailed), again with the Appraise group showing a greater reduction in reported emotional reactivity relative to the Detachment group. There were no significant differences between the Watch and Detachment groups, mean differences < 1.50, $Ps > 0.30$.[1]

Finally, we examined changes in self-reported emotional reactivity across time for each group separately (as in Studies 1 and 2). There were no significant differences for either horror or distress in the Watch and Detachment groups, ts ≤ 1.25, Ps ≥ 0.23. For the Appraise group, as hypothesized, change scores for both distress, t (15) = 2.49, $P < 0.05$ (one-tailed), and horror, t (15) = 2.48, $P < 0.05$ (one-tailed), decreased significantly from pre- to post-training.

As in Study 2, our examination of putative demand effects revealed no significant associations between estimated and actual change in emotions from pre- to post-training, *P*s > .09.

The results from Study 3 once again replicated the findings from Studies 1 and 2 of a significant effect of appraisal practice on self-reported emotional reactivity from pre- to post-training, which was also significantly greater than the non-significant changes in a Watch control group. Furthermore, the present data also revealed a significant advantage of appraisal practice over practicing another form of cognitive evaluation of the training films - cognitive detachment – which itself has been shown to be an effective form of emotion regulation (e.g. Gross, 1998). These data suggest that the beneficial effects of appraisal practice are not simply a result of practicing any cognitive manipulation of the stimulus with plausible emotion-reducing effects.

As in Study 2, there was no support from our expected versus actual change analyses to suggest that expectancy-led demand effects accounted for the patterns of data.

Study Four

In the final study we shifted the focus to training with autobiographical memories. Establishing that appraisal practice is effective for events relevant to the individual's life and emotional well-being is an essential step if such an approach is going to be viable in a clinical context.

We also examined effects of appraisal practice over a longer time period, as Studies 1 through 3 had focused on effects immediately post-training. Furthermore, we wanted to examine whether any effects of practice transferred to day-to-day experience. To these ends we compared daily levels of intrusions (e.g. "I thought about it when I did not mean to") and avoidance (e.g. "I tried not to think about it"), of the autobiographical memories used in the experimental session for the seven days preceding the study and for the same period post-

study, using a standardized measure - the Impact of Event Scale (Horowitz, Wilner & Alvarez,1979).

Our first hypothesis was that appraisal practice would bring about reductions in reported emotional reactivity to selected autobiographical memories in the same way as it had for films in Studies 1 through 3, and that these effects would be greater than for participants receiving no practice. Our second hypothesis was that the appraisal practice group would experience significantly reduced intrusion and avoidance of the target memories on the IES in the seven day follow-up period compared to their pre-study scores, and that this reduction would be greater than any changes in the no-practice controls.

Method

Participants

Thirty-three volunteers were recruited. Exclusion criteria were broadly as for Studies 1 to 3, this time with respect to processing memories rather than films, with two exceptions: First, participants with higher anxiety and depression scores were now included. Scores of up to 23 on the BDI-I (*moderately depressed*; Shaw et al., 1985) and scores of up to 60 on the STAI scales were permitted. Second, participants who did not generate sufficiently distressing memories (rating one or more of the test memories as less than 30 on a 100 point visual analogue SUD scale, ranging from 0 “not at all distressing” to 100 “extremely distressing”) were excluded. This cut off was applied because piloting revealed that allowing scores of thirty or less introduced floor effects into the data. Three participants had to be excluded for this reason, leaving final sample sizes of 15 participants randomly assigned to each group. The groups were comparable for age, gender, state anxiety, trait anxiety, depression and estimated verbal intelligence, ts < 1, Ps > .50.

Memory Task and procedure

The Memory Task was similar in conception to the Film Task (see Figure 1), with the exception that the distressing films were replaced by autobiographical memory scenarios. The procedure was otherwise the same as for Study 1. Memory scenarios were generated a week before the experimental session when participants were asked to produce five 100-word narratives describing distressing memories from their personal past. Out of the five memories, two were used as the pre- and post-training test memories, one as a practice memory and two as training memories. Each memory was rated by the participant in terms of: associated distress, both when the setting event occurred, and currently, on 100-point visual analogue SUD scales, ranging form 0 "not at all distressing" to 100 "extremely distressing" (Dalgleish & Yiend, 2006); how frequently it is currently the object of appraisal (where 0 = "not appraised at all", 1 = "appraised once per year", 2 = "once per month", 3 = "once per week", 4 = "several times per week" and 5 = "every day"); and how much current traumatic stress it generates (over the previous week) according to the Impact of Event Scale (IES; Horowitz et al., 1979). The IES is a self-report measure designed to assess current symptoms of intrusion (e.g. "I thought about it when I did not mean to") and avoidance (e.g. I tried not to talk about it") for any specific life event. Respondents are asked to rate how true certain statements are on a scale of 0 = "not at all", 1 = "a little bit", 2 = "moderately", 3 = "quite a bit" and 4 = "extremely" with regard to a specific life event, taking the past seven days into account. The IES has good reliability and validity (Horowitz et al., 1979). The two most closely matched memories (with regard to their IES scores and current SUD ratings) were selected by the experimenter as the two test memories for each participant. To examine longer-term effects of appraisal practice, participants also completed the IES for the two test memories at one week post-study. The rationale for looking at both test memories was based on the assumption that any skills acquired through appraisal practice would continue to

operate over the seven days post-study and consequently one might expect to see benefits in the processing of not only the post-training test memory, but also of the pre-training memory.

In the 'training phase', participants read the first training memory five times, with each reading followed by a one-minute thought break. Participants in the Appraise group were instructed to apply the four appraisal themes (as used in Studies 1 to 3) once each in the first four thought breaks. During the fifth thought break, participants were asked to reapply the theme that had worked best for them. This process was repeated for the second training memory. Participants in the Control group were asked to simply think about their respective memories during the thought breaks without regulating their emotions in any way. To return affective responses to baseline, after the training period all participants read a neutral script (a description of how to bake a cake) and at the end of the experiment a happy script.

Results and discussion

Instruction compliance

As intended, the two groups differed significantly in terms of the number of appraisals they reported using during the training phase, $t\ (14) = 25.56$, $P < .001$, but did not differ significantly in the extent to which they appraised during the post-memory thought breaks for the two test memories (where both groups were instructed to employ appraisal), $ts \leq 1.05$, Ps $> .3$. For the test memories, participants generally reported distracting themselves little (<24% of the time), suppressing only modestly (<31% of the time), and appraising (>75% of the time). There were no differences between groups on these dimensions, Ps $> .1$.

Memory Task

Regarding the pre-study data for the two test memories there were no significant differences across groups for either memory, Ps $> .1$, and both memories were reported as a source of marked current distress (SUD scores past and current > 60; IES scores > 17). For both groups, the second test memory was rated as less distressing across the different

measures than the first. However, critically this was comparable across the two groups, *ts* ≤ 0.91, *Ps* ≥ 0.37.

As in previous studies, change scores indexed participants' subjective emotional reaction (distress, horror), this time focusing on the differences between the baselines and post-memory thought breaks (see Table 4 and Figure 4). There were no significant differences between groups for either the pre-training baseline rest period, the post-training baseline rest period, the change across baselines, or the pre-training post-memory thought break, *ts* ≤ 2.00, *Ps* ≥ 0.06.

To examine the effects of appraisal practice on self-reported emotional reactivity, a mixed model MANOVA was performed. The multivariate output revealed no main effect of Time, $F < 1$, or Group, $F\ (2, 27) = 1.81$, $p = 0.18$, $\eta_p^2 = 0.12$, but a strong trend for a Time by Group interaction, $F\ (2, 27) = 2.95$, $p = 0.07$, $\eta_p^2 = 0.18$. Univariate output showed the hypothesized significant interaction effects for distress, $F\ (1, 28) = 5.64$, $p < 0.05$, $\eta_p^2 = 0.17$, and horror, $F\ (1, 28) = 2.99$, $p < 0.05$, $\eta_p^2 = 0.10$, with the Appraise group showing greater reductions in emotional reactivity relative to the Control group from pre- to post-training.

Paired samples t-tests revealed no significant decrease in the mean emotion change scores for distress, $t\ (1, 14) = 1.43$, $p = 0.18$, $d = 0.61$, or horror, $t\ (1, 14) = 0.83$, $p = 0.42$, $d = 0.28$, in the Control group from pre-training to post-training. However, in line with our hypotheses, there was a significant decrease over time in changes in self-reported distress, $t\ (1, 14) = 2.36$, $p < 0.05$ (one-tailed), $d = 0.67$, and horror, $t\ (1, 14) = 1.78$, $p < 0.05$ (one-tailed), $d = 0.56$, in the Appraise group.

As in Studies 2 and 3, our examination of putative demand effects revealed no significant associations between estimated and actual change in emotions from pre- to post-training for either group, *P*s > .19. Furthermore, within the Appraise group the correlations between actual and estimated change were non-significant, *r*s < |.36|, *P*s > .18, with the

correlation involving distress actually going in the direction opposite to that which would be expected if demand effects were in operation.

IES pre- and post-study

Three control participants were excluded from this analysis since they did not complete and return the IES 7-days post-study.

A mixed-model MANOVA with time (pre- versus post-study) as the within-subjects factor, group as the between-subjects factor and the Total IES scores for each of the two test memories as dependent variables (see Figure 4b and Table 5) revealed a significant multivariate effect of Time, F (2, 24) = 4.16, $p < 0.05$, $\eta_p^2 = 0.26$, with IES scores reducing post-study, and a Time by Group interaction, F (2, 24) = 3.91, $p < 0.05$, $\eta_p^2 = 0.25$, consistent with a greater reduction in IES scores for the Appraise group relative to the Controls. Univariate output revealed a Time by Group interaction for the pre-training memory, F (1, 25) = 6.74, $p < 0.05$, $\eta_p^2 = 0.21$, and a trend for the post-training memory, F (1, 25) = 3.37, $p = 0.08$, $\eta_p^2 = 0.12$.

Breaking down these interaction effects, paired-samples t-tests revealed no significant change in IES scores across time for the Control group for either memory, $ts < 1$. However, as hypothesized, IES scores significantly decreased for both memories for the Appraise group: Pre-training memory, t (1, 14) = 4.13, $p < 0.01$ (one-tailed), $d = 0.75$; post-training memory, t (1, 14) = 2.61, $p < 0.01$ (one-tailed), $d = 0.46$.

The results of Study 4 critically showed that the effects of practicing perspective broadening that were observed in Studies 1 through 3 for film material were also present when using autobiographical memories, with significant reductions in self-reported emotional reactivity (horror, distress) in the Appraise group from pre- to post-training, that were greater than the non-significant changes in the Control participants. Furthermore, appraisal practice led to a reduction in self-reported intrusion and avoidance of the target personal memories in

the week following the study, with significant decreases in IES scores across this period (relative to the week pre-study) that were significantly greater than the non-significant effects in the Controls. Interestingly, these effects of appraisal practice on IES scores were evident not only for the post-training test memory but also for the pre-training test memory (and in fact were numerically stronger for the latter), suggesting that any skills accrued as a result of appraisal practice acted on this latter memory occurred in the seven days post-study.

As in Studies 2 and 3, our expectancy versus actual change analyses provided no support for expectancy-led demand effects for the patterns of pre- to post-training data within the experimental session.

General Discussion

The present series of 4 studies examined the viability of using CBM methods to examine the effects of systematic practice in the use of functional cognitive appraisals related to 'seeing the bigger picture' on negative emotional reactions to both autobiographical and non-autobiographical distressing stimuli.

Study 1 showed beneficial effects of appraisal practice with distressing film material in terms of reduced self-reported emotional reactivity and galvanic skin response (GSR). Studies 2 and 3 replicated the findings of Study 1, this time with participants appraising immediately after the film rather than while it was playing. In addition, evidence was presented in Study 3 indicating that individuals practiced in generating appraisals subsequently benefited from a greater reduction in emotional reactivity to a distressing film relative to individuals trained in another form of emotion-regulation (detachment) indicating that the effects are somewhat specific to appraisal practice.

Study 4 showed that practice in appraising autobiographical memories replicated the pattern of findings for film material in Studies 1 through 3. Furthermore, maladaptive processing of the memories according to the Impact of Event Scale reduced significantly

following appraisal practice from the pre-study week to the post-study week, with no such reduction in no-practice control participants.

Across all studies appraisal practice effects were significant both within the appraisal practice group (i.e. significant decreases in emotion, psychophysiology, and IES scores from pre- to post-training) and when comparing practice and control groups. The former effects are particularly important as such systematic practice would have limited clinical utility if it did not produce strong and reliable within-subjects effects. Furthermore, they indicate that the significant between-group effects are not simply a function of any change in emotional reactivity from pre-to-post training in the controls (which in any case was non-significant in all studies).

As discussed in the Introduction, a key way in which the present studies differ from previous CBM investigations was the focus on asking participants to practice generating local appraisals, tailored to the specific emotional stimulus under evaluation, as a function of more general appraisal templates that were chosen so as to broadly reflect the ability to see the bigger picture. The fact that such robust and reliable training effects of practice on emotional reactivity were obtained using this approach is extremely encouraging as it suggests that individuals can benefit from practicing skills with a potentially broad range of application, as opposed to simply practicing with large sets of very specific, local appraisals, each with limited applicability.

The broad appraisal themes selected for the purposes of the present research were deliberately very general. It may well be the case that for particular groups of individuals trying to deal with relatively circumscribed emotive material, and/or solely for autobiographical material (without the need to accommodate film material), other themes may be more appropriate and this is a matter for future investigations. However, a more pressing issue is to examine the ability to translate the present methodology to a clinical

sample. The first 3 studies reported here excluded individuals with even moderate levels of anxiety and depression. Study 4 was more inclusive but still used a sub-clinical sample. If the present approach is to have clinical viability, its reliability in a clinical group needs to be established.

The present studies are not without their limitations. First, across all studies there was a confound in the instructions given to the appraisal practice participants who were asked both to use appraisal and to alter their emotions to the training clips/memories, compared with the control groups who were requested to do neither. It therefore remains possible that it is the request to alter emotions during the training phase, rather than the request to appraise, that is driving the current findings. That said, we are reasonably confident that any such an effect is not operating via simple demand, as: 1) we found a greater reduction in GSR reactivity to films after training, compared with before training, for the training group in Study 1; 2) there was no significant correlation between expectancies about change and actual change in Studies 2-4; and, 3) for the pre- and post-training test stimuli that provide the basis for our key findings, all participants were given *exactly the same instructions* regarding appraisal. Nevertheless, future research should unconfound the instructions for the training material either by referencing emotion change in any control condition or by not referencing it in the appraisal practice conditions.

A second limitation is that we did not assess the propensity to use the recently practiced appraisal themes in the participants who practiced (versus control participants) by giving them a free hand in how to process the post-training test material and then assessing any use of appraisal. Instead, we explicitly instructed *all* participants to apply the appraisal themes that were the focus of the study. The rationale for this was to provide a more conservative test of the effects of practice on emotional reactivity and related indices by actually giving the controls the same tools as the participants who practiced. This approach

therefore differs from other CBM methodologies (e.g. Mathews & Mackintosh, 2000), where the degree to which a pattern of bias has been acquired is explicitly assessed. It is important that future research on perspective broadening should address this shortfall by assessing both whether a different or better way of appraising has been successfully acquired and also what the effects (e.g. on emotional reactivity) of such acquisition might be.

In summary, the present studies show that practice in tailoring general appraisal themes reflecting 'seeing the bigger picture' to fit specific emotion-inducing autobiographical and non-autobiographical events leads to benefits in terms of self-reported emotions, psychophysiology and intrusion and avoidance of novel material over the following week. Appraisal practice therefore has considerable promise in terms of integrating CBM Methods into the clinic, and future research needs to target the viability of these techniques in clinical groups.

References

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. NY: Meridian.

Beck, A. T. (2005). The current state of cognitive therapy - A 40-year retrospective. *Archives of General Psychiatry, 62*, 953-959.

BIOPAC, (1997). BIOPAC MP100. Psychophysiological Data Acquisition System: Biopac Systems, Inc.

Brownley, K.A., Hurwitz, B.E., & Schneidemann, N. (2000). Cardiovascular Psychophysiology. In J.T. Caccioppo, G. Tassinari, & G.G. Bernston (Eds.), *Handbook of Psychophysiology*, (2nd ed., pp. 224- 265). Cambridge, UK: Cambridge University Press.

Dalgleish, T. (2004). Cognitive approaches to Posttraumatic Stress Disorder (PTSD): The evolution of multi-representational theorizing. *Psychological Bulletin, 130*, 228-260.

Dalgleish, T., & Yiend, J. (2006). The effects of suppressing a negative autobiographical memory on concurrent intrusions and subsequent autobiographical recall in dysphoria. *Journal of Abnormal Psychology, 115*, 467 – 473.

Dawson, M.E., Schell, A.M., & Filion, D.L. (2000). The electrodermal system. In J.T. Caccioppo, L.G. Tassinary, & G.G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed., pp. 200 - 223*)*. Cambridge, UK: Cambridge University Press.

Ehlers, A., & Clark, D.M. (2000). A cognitive model of post traumatic stress disorder. *Behavior Research and Therapy, 38,* 319 – 345.

Grey, S.J., & Smith, B.L. (1984). A comparison between commercially available electrode gels and purpose-made gel, in the measurement of electrodermal activity. *Psychophysiology, 21*, 551-557.

Gross, J.J. (1998). Antecedent- and response- focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology, 74*, 224-237.

Gross, J.J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology, 39,* 281-291.

Haaga, D.A.F. (2000). Introduction to the special section on stepped care models in psychotherapy : Stepped care models in psychotherapy. *Journal of Consulting and Clinical Psychology,* 68, 547-548.

Holmes, E. A., Brewin, C. R., & Hennessy, R. G. (2004). Trauma Films, Information Processing, and Intrusive Memory Development. *Journal of Experimental Psychology General*, *133*, 3-22.

Holmes, E., Mathews, A., Mackintosh, B. & Dalgleish, T. (2006). Positive imagery during interpretation training improves mood and positive interpretative bias. *Behavior Therapy, 37*, 237-247.

Horowitz, M. J., Wilner, N., & Alvarez, W. (1979). Impact of Event scale: A measure of subjective stress. *Psychosomatic Medicine, 41*, 209_218.

Jamison, K. R., Gerner, R. H., Hammen, C., & Padesky, C. (1980). Clouds and Silver Linings - Positive Experiences Associated with Primary Affective-Disorders. *American Journal of Psychiatry, 137*, 198-202.

Koster, E., MacLeod, C. & Fox, E. (in press). Editorial to Special Section on cognitive bias modification (CBM). *Journal of Abnormal Psychology.*

Lazarus, R.S., & Alfert, E. (1964). Short-circuiting of threat by experimentally altering cognitive appraisal. *Journal of Abnormal and Social Psychology, 69,* 195-205.

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association

through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology, 111*(1), 107-123.

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

Nelson, E. (1982). *National Adult Reading Test (NART):* The Manual. Windsor, UK: NFER-Nelson.

Power, M.J. & Dalgleish, T. (2008). *Cognition and emotion: From order to disorder* (2nd Ed.). Hove, U.K.: Psychology Press.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime 1.1 users reference guide.* Pittsburgh: Psychology Software Tools Inc.

Shaw, B.F., Vallis, T.M., & McCabe, S.B. (1985). The assessment of the severity and symptom patterns in depression. In E.E. Beckham & W.R. Leber (Eds.), *Handbook of depression: Treatment, assessment and research.* Homewood, IL: Dorsey Press.

Spielberger, C.D., Gorsuch, R.L., Lushene, R., Vagg, P.R., & Jacobs, G.A. (1983a). *Manual for the State-Trait Anxiety Inventory.* Palo Alto, CA: Consulting Psychologists Press.

Stern, R.M., Ray, W.J., & Quigley, K.S. (2001). *Psychophysiological recording* (2nd ed.), Oxford, UK: Oxford University Press.

Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1997). *Cognitive psychology and emotional disorders* (2nd ed.). Chichester: Wiley.

Wilson, E.J., MacLeod, C., Mathews, A., & Rutherford, E.M. (2006). The causal role of interpretive bias in anxiety reactivity. *Journal of Abnormal Psychology, 115,* 103 – 111.

Yiend, J., Mackintosh, B., & Mathews, A. (2005). Enduring consequences of experimentally induced biases in interpretation. *Behaviour Research and Therapy, 43*, 779-797.

Author note

Patricia Schartau, Tim Dalgleish and Barney Dunn, Emotion Research Group, Medical Research Council Cognition and Brain Sciences Unit, Cambridge, England. We would like to thank Andrew Mathews for comments and help with this work. This research was funded by the U.K. Medical Research Council. Correspondence concerning this article should be addressed to Tim Dalgleish, MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, U.K., or by electronic mail to: tim.dalgleish@mrc-cbu.cam.ac.uk

Footnote

1- Because the Detachment group reportedly spent significantly more time suppressing during the pre-training test film than the Watch group, we repeated the analyses with time spent suppressing during this test film as an additional covariate. The results remained the same.

Table 1

Mean (*SD*) self-reported emotion scores ($N = 41$), galvanic skin response (GSR; in μsiemens) and heart rate (HR; in beats per minute) ($n = 37$) data by group for the pre-and post-training baselines and distressing test films, for Study 1.

	Watch group ($n = 20$)				Appraise group ($n = 21$)			
	Base 1	Pre	Base 2	Post	Base 1	Pre	Base 2	Post
Horror	1.50 (0.76)	5.65 (2.35)	2.05 (1.43)	6.25 (2.20)	2.05 (1.80)	6.67 (2.06)	2.38 (1.94)	4.71 (2.61)
Distress	1.65 (0.93)	5.40 (2.39)	2.25 (1.59)	5.80 (2.53)	2.14 (1.53)	6.38 (1.72)	2.52 (1.89)	5.05 (2.01)
	($n = 18$)				($n = 19$)			
GSR	4.96 (2.69)	6.16 (3.32)	6.72 (3.64)	7.64 (4.11)	6.63 (3.40)	8.10 (3.70)	8.57 (3.47)	9.23 (3.59)
HR	86.11 (16.43)	83.25 (14.00)	81.54 (12.44)	79.97 (10.25)	79.90 (8.88)	79.30 (7.17)	84.87 (18.22)	82.38 (17.70)

Note

Base 1/2 = first/second baseline rest

Pre/Post = pre/post-training test film

Table 2

Mean (*SD*) scores for each target emotion by group for pre-and post-training baselines and post-film thought breaks for Study 2 (*N* = 32).

	Watch group (*n* = 16)				Appraise group (*n* = 16)			
	B1	Pre	B2	Post	B1	Pre	B2	Post
Horror	1.50	3.69	1.81	3.69	1.25	4.88	2.31	3.13
	(1.03)	(2.18)	(1.38)	(2.36)	(0.77)	(2.80)	(1.54)	(2.31)
Distress	1.75	3.31	2.00	3.94	1.75	4.75	2.31	3.38
	(1.29)	(1.81)	(1.46)	(2.81)	(0.93)	(2.08)	(1.20)	(2.22)

Note

B1/B2 = first and second baseline rest scores

Pre/Post = pre-/post-training post-film thought break scores

Table 3

Mean (*SD*) scores for each target emotion by group for pre-and post-training baselines and post-film thought breaks for Study 3 (*N* = 48).

	Watch group (*n* = 16)				Appraise group (*n* = 16)				Detachment group (*n* = 16)			
	B1	Pre	B2	Post	B1	Pre	B2	Post	B1	Pre	B2	Post
Horror	2.25	4.50	2.50	5.13	1.69	4.81	2.31	3.00	2.25	4.69	3.63	4.87
	(1.65)	(2.16)	(1.93)	(2.36)	(1.08)	(2.43)	(1.85)	(1.83)	(1.29)	(2.15)	(2.47)	(1.75)
Distress	2.13	4.06	2.56	4.63	1.50	4.31	2.31	3.44	2.69	4.69	3.38	4.94
	(1.59)	(2.26)	(2.06)	(2.47)	(1.03)	(2.36)	(1.54)	(1.59)	(1.81)	(2.07)	(1.86)	(2.02)

Note

B1/B2 = first and second baseline rest scores

Pre/Post = pre-/post-training post-film thought break scores

Table 4

Mean (*SD*) scores for each target emotion by group for pre-and post-training baselines and post-memory thought breaks for Study 4 ($N = 30$).

	Control group ($n = 15$)				Appraise group ($n = 15$)			
	B1	Pre	B2	Post	B1	Pre	B2	Post
Horror	1.87	1.80	1.20	1.67	2.07	3.33	1.60	2.00
	(1.36)	(1.37)	(0.41)	(1.23)	(1.39)	(2.74)	(1.12)	(1.60)
Distress	3.13	3.47	2.20	3.73	3.00	5.00	2.47	3.40
	(1.92)	(1.20)	(1.32)	(2.25)	(2.04)	(2.20)	(1.68)	(1.45)

Note

B1/B2 = first and second baseline rest scores

Pre/Post = pre-/post-training post-film thought break scores

Table 5

Mean (*SD*) IES Total scores by group pre-and post-study for Study 4 for the two test memories (n = 27).

	Control group (n = 12)		Appraise group (n = 15)	
	Pre-experiment	Post-experiment	Pre-experiment	Post-experiment
IES Total score for pre-study memory	23.75 (19.54)	23.08 (17.74)	26.20 (14.89)	15.07 (14.60)
IES Total score for post-study memory	21.08 (19.66)	22.58 (19.70)	19.27 (12.88)	13.27 (13.18)

Figure 1

Timeline of the study procedure for Studies 1 to 4.

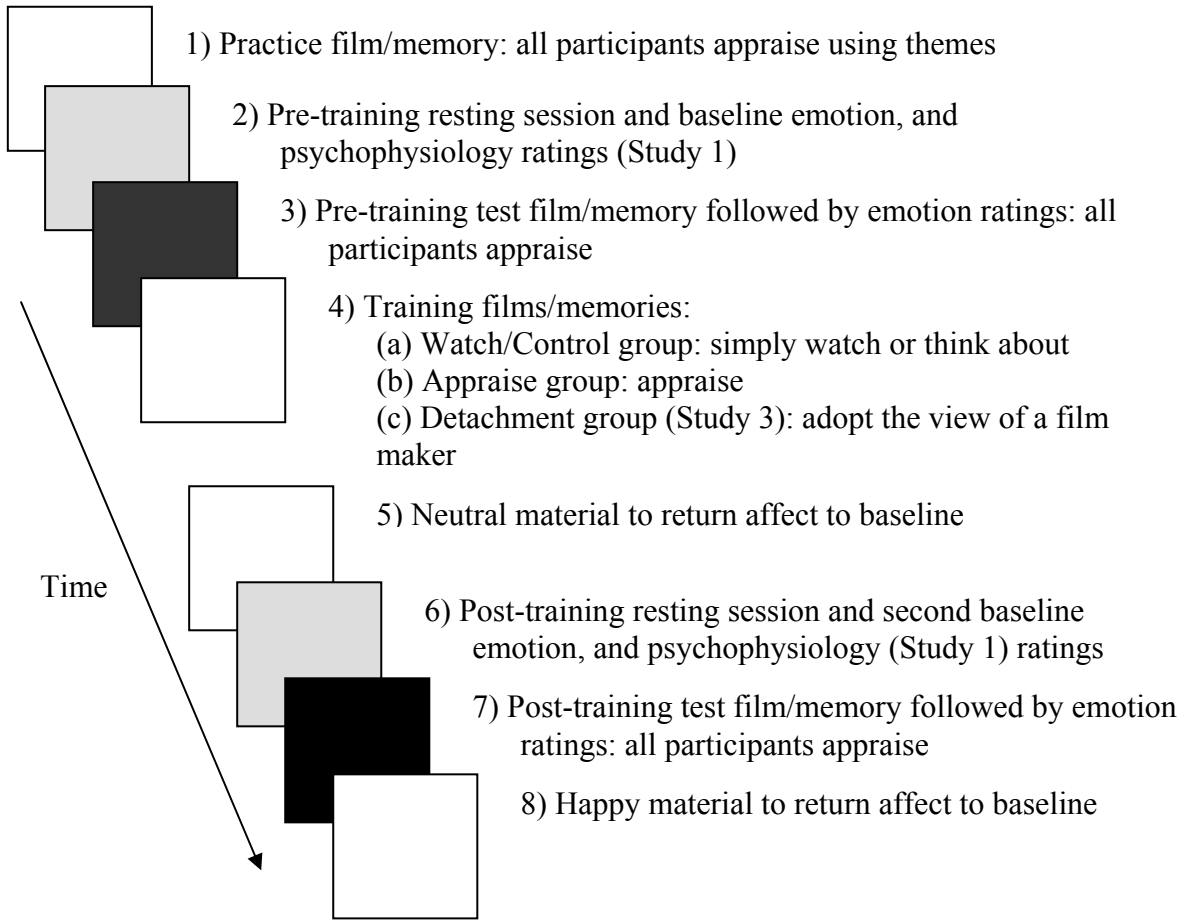
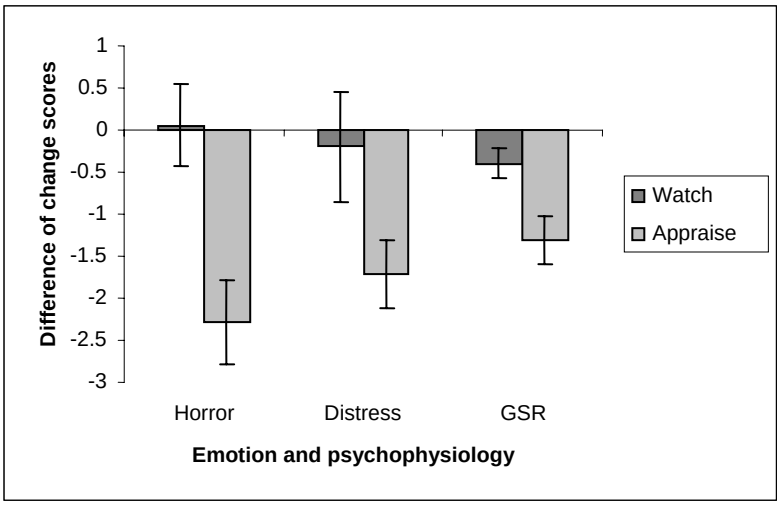


Figure 2

Mean (S.E.M.) difference of change scores in self-reported ratings of horror and distress (*N* = 41), and GSR (*n*= 37), post-training minus pre-training, for Study 1.



Note

GSR changes indexed in μsiemens.

Figure 3a

Mean (S.E.M.) difference of change scores in self-reported ratings of horror and distress (post-training minus pre-training) for Study 2 (N = 32).

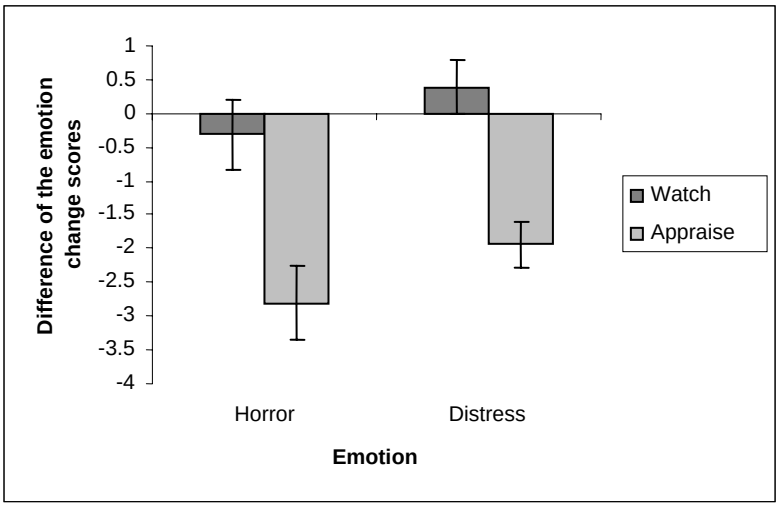


Figure 3b

Mean (S.E.M.) difference of change scores in self-reported ratings of horror and distress (post-training minus pre-training) for Study 3 (N = 48).

Figure 4a

Mean (S.E.M.) difference of change scores in self-reported ratings of horror and distress (post-training minus pre-training) for Study 4 (N = 30).

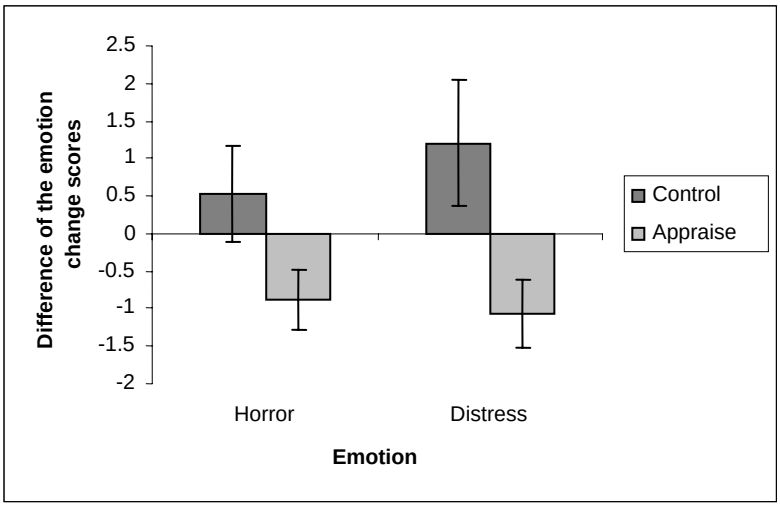


Figure 4b

Mean (S.E.M.) change in IES scores by group (post-study minus pre-study) for each of the two test memories for Study 4 (N = 27).