HOW ARE COGNITIVE THERAPY SKILLS USED (OR NOT USED) YEARS AFTER TREATMENT?
SUBJECTIVE AND PHYSIOLOGICAL MEASURES

by

Benjamin Sajowitz Paul

B.A., Psychology, Philosophy, University of Southern California, 2008

Submitted to the Graduate Faculty of

the School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

SCHOOL OF ARTS AND SCIENCES

This thesis was presented

by

Benjamin Sajowitz Paul

It was defended on

July 7, 2011

and approved by

Thomas W. Kamarck, Ph.D., Professor, Department of Psychology

Andrew Koffmann, Ph.D., Clinical Associate Professor, Department of Psychology

Kathryn A. Roecklein, Ph.D., Assistant Professor, Department of Psychology

Thesis Director: Greg J. Siegle, Ph.D., Associate Professor, Department of Psychiatry

HOW ARE COGNITIVE THERAPY SKILLS USED (OR NOT USED) YEARS AFTER TREATMENT?

SUBJECTIVE AND PHYSIOLOGICAL MEASURES

Benjamin Sajowitz Paul, M.S.

University of Pittsburgh, 2011

Theorists have proposed that core Cognitive Therapy skills, such as reappraising negative thoughts, are useful not only for recovering from depression but also for staying well years after recovery. However, surprisingly little empirical work has tested (1) whether CT skills are used years after therapy, (2) whether continued CT skill use is associated with mood benefits, and (3) whether continued CT skill use involves significant effort and emotional engagement. To address these issues, we measured frequency of CT skill use in former CT patients, and we tested the association between CT skill use and adaptive mood outcomes as well as the association between CT skill use and engagement with dysfunctional thoughts, as indexed via pupil dilation.

Twenty-two formerly depressed individuals aged 24-70 who responded to research-protocol CT 2-6 years ago completed the Skills of Cognitive Therapy – Patient (SoCT-P) self-report measure and underwent two experimental blocks in which a sad mood induction was provided and then pupil dilation was measured during self-paced presentation of self-relevant dysfunctional statements. During the first block, participants were asked to cope however they normally would, while in the second block, participants were asked to reappraise each statement. We examined (1) mean SoCT-P, (2) correlations between SoCT-P and adaptive mood outcomes

(lack of cognitive reactivity and transience of induced sadness), and (3) the correlation between SoCT-P and pupil dilation in response to dysfunctional statements.

Results indicated that, on average, participants reported using CT skills between “half” and “most” of the time. SoCT-P was non-significantly associated with adaptive mood outcomes, with moderate effect sizes. In the first block, SoCT-P was negatively correlated with pupil dilation during the period immediately before and after presentation of each statement.

This study provides the first evidence that CT skills are still used years after therapy. More power is needed to test whether continued CT skill use is associated with positive mood outcomes, although results were in the predicted direction. Finally, as CT skill use was negatively correlated with an index of cognitive and emotional load, results may argue against the notion that continued CT skill use involves deployment of effortful compensatory strategies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

It is often suggested that the skills taught in Cognitive Therapy for depression (CT) are important long after patients have recovered from depression. In particular, some have proposed that patients who respond to CT will continue to use and benefit from the skills they learned even after completing therapy (A. T. Beck, 1979; R. J. DeRubeis, Siegle, & Hollon, 2008; Hollon, 2003; Hollon, Stewart, & Strunk, 2006; R. B. Jarrett & Kraft, 1997). Under this perspective, core CT skills, such as reappraising negative thoughts, are useful not only for recovering from mental illness but also for staying well after recovery. Yet surprisingly little empirical work has tested this perspective. A better understanding of post-acute use of CT skills is important for (1) determining whether the same skills that are useful in recovering from depression with CT are also important for maintaining euthymic mood after CT, and (2) our knowledge of how former CT patients use their skills, i.e., whether such individuals automatically process negative stimuli in an adaptive manner or whether significant effort and emotional engagement is involved to counter stable negative information processing tendencies. The current study sought to fill these gaps by investigating the use, usefulness, and effort involved in deploying a skill thought to be particularly important in CT (A. T. Beck, 1979), cognitive reappraisal of negative thoughts.

Cognitive reappraisal, i.e., changing the emotional impact of a stimulus by changing one's mental interpretation of that stimulus (Gross & John, 2003), is generally an effective strategy to down-regulate negative affect. This has been evidenced by a growing body of research indicating that instructing participants to change their interpretation of negative stimuli results in less negative affect compared to control conditions (Gross, 1998; Hofmann, Heering, Sawyer, & Asnaani, 2009; Oliveira, 2008; Teasdale & Fennell, 1982; L. E. Williams, Bargh, Nocera, & Gray, 2009). Further evidence comes from findings that individuals who report frequent cognitive reappraisal also report better mental health and respond more adaptively to stressors (Gross & John, 2003; Mauss, Cook, & Gross, 2007; Mauss, Evers, Wilhelm, & Gross, 2006; L. E. Williams et al., 2009).

Yet depressed individuals are deficient in their use of this beneficial strategy. Depressed individuals report using this skill to a lesser extent than healthy controls (Gotlib & Joormann, 2010; S. A. Moore, Zoellner, & Mollenholt, 2008; Siegle, P. M. Moore, & Thase, 2004), and depressive symptomatology is negatively correlated with use of reappraisal (Aldao, Nolen-Hoeksema, & Schweizer, 2010).

Therefore, one would expect that depressed individuals would benefit from learning to reappraise negative thoughts more frequently. Indeed, CT patients who receive more training in cognitive skills such as reappraisal do experience greater symptom improvement (Strunk, Brotman, & R. J. DeRubeis, 2010), suggesting that learning cognitive skills is beneficial for depressed individuals. In line with this interpretation, individuals who report greater CT skill use at therapy mid-point are more likely to achieve therapy response by the end of acute treatment (R. B. Jarrett, Vittengl, L. A. Clark, & Thase, 2011), underscoring the potential role that CT skills may play in recovery from depression. Furthermore, decrease in dysfunctional thinking during CT, which may be related to increasing reappraisal of dysfunctional thoughts, is predictive of subsequent decline in depressive symptoms (R. J. DeRubeis et al., 1990). The evidence suggests that learning to reappraise more frequently during CT is generally beneficial for depressed individuals.

Thus, teaching cognitive reappraisal in CT appears to be associated with increased use of this skill, which is in turn associated with less negative mood. However, surprisingly little research has directly tested whether individuals continue to make use of CT skills and whether continued skill use would be beneficial at maintaining therapy gains.

Initial data concerning frequency of CT skill use following therapy comes from a small qualitative study (Glasman, Finlay, & Brock, 2004). Nine individuals who had responded to community-delivered therapy described as cognitive-behavioral in nature, were administered a semi-structured interview within ten months of therapy completion. Eight of nine individuals (89%) reported that they still reappraise automatic negative thoughts as they learned to do in CT, while the other individual reported that CT techniques were never relevant to him. Thus, initial small-scale data would predict that post-therapy use of CT skills is fairly common.

Although no research has examined CT skill use longer than ten months after therapy completion, data from another form of psychotherapy may suggest that skill use substantially declines years after treatment. For instance, in a form of psychotherapy that encourages practice of silent meditation several times a week, approximately 90% of individuals report frequent practice at the end of treatment, and 60% continue frequent practice six months after treatment completion, but by four years post-treatment only 30% continue to meditate frequently. Even the 30% figure is inflated by frequent use of add-on meditation programs after acute treatment completion (Bondolfi et al., 2010; J. J. Miller, Fletcher, & Kabat-Zinn, 1995; Reibel, Greeson, Brainard, & Rosenzweig, 2001) Based on this evidence, we hypothesized that CT skill use would significantly decline but not cease completely following treatment termination.

However, it is important to investigate not only frequency of post-therapy skill use but also whether post-therapy skill use incurs benefits. Although a growing literature have demonstrated that CT skill use at therapy completion is predictive of lesser depressive symptomatology at follow-ups (Gibbons et al., 2009; Powers, Thompson, & Gallagher-Thompson, 2008; Strunk, R. J. DeRubeis, Chiu, & Alvarez, 2007), it has not been empirically established that CT skill use after therapy remains important. Theoretically, Beck (1979) suggested that after patients complete therapy, they would continue to benefit from applying CT skills, such as reappraisal, to "the inevitable problems [they] will face" (p. 319). He even encouraged recovered patients to induce a negative mood so as to demonstrate that the CT techniques they had learned would effectively dissipate the mood. Jarrett and Kraft (1997) expanded on this notion by developing a system of Continuation CT that promotes continued practice and generalization of acute CT skills. In both cases, the investigators tacitly assume that the skills that are useful for recovering from depression are also useful for staying well following recovery.

In fact, some preliminary support has recently emerged to bolster this assumption. In a survey of ninety older formerly depressed individuals who had responded to CBT eighteen months ago, those who continued to report few depressive symptoms also reported that CT skills were helpful to them (Powers et al., 2008). This suggests that individuals may at least attribute continued benefits to continued use of CT skills, if not actually benefit from them. Based on this evidence, as well as the aforementioned evidence that cognitive reappraisal is generally linked with adaptive mood outcomes (Gross, 1998; Gross & John, 2003; Hofmann et al., 2009; Mauss et al., 2007, 2006; Oliveira, 2008; Teasdale & Fennell, 1982; L. E. Williams et al., 2009), we hypothesized that continued CT skill use would be associated with adaptive mood outcomes years after CT.

Benefits of continued CT skill use have also been discussed in literature concerning "enduring effects" of CT. In interpreting evidence that CT's positive effect on mood may endure even after treatment termination, Hollon (2003) hypothesized that continued skill use may play a causal role in this process. DeRubeis et al. (2008) elaborated a potential neural model of this theory, in which CT may teach patients to voluntarily activate prefrontal cortex to regulate negative emotions, and once patients learn this skill, they can continue to incur continued benefits by implementing it after therapy completion.

Overall, a salutary effect of continued CT skill use has been assumed by some, hypothesized by others, but rarely tested. Yet there have been some theoretical proposals that do not attribute beneficial effects to continued skill use per se. For example, Hollon et al. (2006) raised the possibility that deeper-level changes in core beliefs may be most responsible for enduring effects, rather than explicit skill deployment. In particular, initial CT skill use may lead to deeper-level changes in negative beliefs about the self, and it is the latter change that is proximally responsible for enduring gains (cf. Barber & R. J. DeRubeis, 1989). Indeed, there is some evidence that CT reduces negative elaborations about the self , potentially related to deactivation of core beliefs, rather than increasing explicit deployment of CT skills (Gibbons et al., 2009). This is consistent with evidence that after CT, knowledge about one's own core beliefs is a more powerful predictor of continued mood benefits than is identification and reappraisal of negative thoughts (Strunk et al., 2007). Thus, some accounts would hold that CT skills per se may not provide enduring benefits but that deep-level changes incurred by earlier use of CT skills may be responsible.

Aside from the questions of the use and usefulness of CT skills post-therapy, a final question concerns the effort and emotional engagement involved in deploying such skills. Some have suggested that after continued voluntary reappraisal of negative thoughts, the process becomes automatized such that it no longer requires significant effort (Barber & R. J. DeRubeis, 1989; Hollon et al., 2006; Persons, 1993). Empirical evidence for this claim is scant, aside from a qualitative study that found that some individuals (*n* not given) describe continued use of CT skills as perpetual "hard work" while others report having internalized coping strategies to the point where hard work was no longer necessary (Glasman et al., 2004). Thus, the role of effort in CT skill use is somewhat ambiguous. However, non-psychotherapy studies have more extensively assessed the effort of implementing cognitive reappraisal. Such studies have consistently found that cognitive reappraisal involves significant effort and emotional engagement (Johnstone, van Reekum, Urry, Kalin, & Davidson, 2007; van Reekum et al., 2007; Urry et al., 2006). In line with this literature, we hypothesized that CT skill use would involve significant effort and emotional engagement.

Thus, the current study investigates open questions regarding the use, usefulness, and degree of effort and emotional engagement involved in continued CT skill use. Our hypotheses were: (1) CT skills are still used years after therapy but to a lesser extent than at treatment completion, (2) CT skills are "useful," i.e., CT skill use will be positively associated with adaptive mood outcomes, and (3) CT skill use will involve significant effort and emotional engagement.

Answering these questions may have large-scale implications. If former patients report that they generally do not use CT skills anymore, CT skills might not be relevant to their situation as recovered depressed individuals, and different techniques that work without negative thoughts could be promoted for this time period, such as Mindfulness-Based Cognitive Therapy (Teasdale et al., 2000).

If those who continue to deploy CT skills are better able to deal with negative moods, then it may be useful to devise methods to keep patients motivated to deploy CT skills even when they are no longer feeling depressed. If not, responding to negative thoughts could be conceived of as analogous to an antibiotic that is helpful for acute illness but ineffective following the illness's resolution, and other techniques may be more relevant following CT completion.

Finally, if CT skills are used, and if deploying such skills requires significant effort and emotional engagement, then it may be useful to devise methods to keep patients motivated to deploy arduous CT skills after treatment termination. If not, then therapists may consider extending therapy to the point where new information processing tendencies become automatized.

To help address these issues, we conducted laboratory assessments on individuals who have previously responded to CT for depression. We examined frequency of CT skill use with the Skills of Cognitive Therapy – Patient (SoCT-P; R. B. Jarrett et al., 2011) self-report measure. We tested the level of effort and emotional engagement in CT skill users by presenting individuals with distressing sentences and assessing whether those who use CT skills had greater effort and emotional engagement than those who do not use CT skills. Effort and emotional engagement was indexed via pupil dilation, as the pupil reliably dilates when a task becomes more difficult, greater attention is brought online (Steinhauer & Hakerem, 1992), or more intense emotional stimuli are presented (Steinhauer, Boller, Zubin, & Pearlman, 1983). Previous studies have used pupil dilation reliably to index effort and emotional engagement in a variety of tasks, including reappraisal in healthy and depressed individuals (Johnstone et al., 2007; van Reekum et al., 2007; Urry et al., 2006), emotional processing in depression (Siegle, Granholm, Ingram, & Matt, 2001; Siegle, Steinhauer, Carter, Ramel, & Thase, 2003; Siegle, Steinhauer, Stenger, Konecky, & Carter, 2003; Siegle, Steinhauer, Thase, Stenger, & Carter, 2002), and emotional processing in individuals at-risk for depression (Steidtmann, Ingram, & Siegle, 2010).

To determine whether those who use CT skills are better able to deal with negative moods, we examined the relationship of self-reported skill use with laboratory measures of recovery from a sad mood and with change in dysfunctional attitudes following a mood induction ("cognitive reactivity"). This measure, which taps the ease to which dysfunctional attitudes become activated under stressful circumstances, can also be used as an analog for relapse risk (Z. V. Segal, Gemar, & S. Williams, 1999; Z. V. Segal et al., 2006), allowing us to assess whether CT skill use may be related to lower relapse vulnerability.

As secondary goals, we tested whether CT skill use was associated with decreased viewing time of negative sentences, and whether engagement with negative thoughts is lower in those who have stayed well post-CT relative to those who have relapsed.

2.0 METHODS

2.1 PARTICIPANTS

Former participants of a study of Cognitive Therapy for relapse prevention in depression (MH58356, MH58397; Thase, Jarrett, PIs; described below) and an ongoing study of response to CT for depression (MH074807; Siegle, PI; described below) comprised the subject pool for the present study.

We sent letters and called individuals from these studies who had responded to CT within the last 2-6 years. This time frame was thought to be long enough to examine whether CT teaches enduring skills but short enough that participants could still remember learning those skills.

Following Jarrett and Thase (2010), response was considered 17-item Hamilton Rating Scale for Depression (HRSD) <= 12 and absence of DSM-IV Major Depressive Episode (MDE) as rated by the Structured Clinical Interview for DSM-IV (SCID; First, Spitzer, Gibbon, & J. B. W. Williams, 1997). We excluded individuals who did not respond to CT because they are less likely to have gained CT skills at all (R. B. Jarrett et al., 2011) and therefore it would be difficult to interpret any subsequent reports of CT skill use. We chose to exclude individuals who met SCID (DSM-IV) criteria for current MDE at the time of the testing session because, as previously mentioned, our investigation concerns the use of CT skills during euthymic periods.

With one exception, we also excluded individuals experiencing any Axis I pathology, so that results would not be attributable to mental processes of other conditions. For instance, since intrusive recollections and dissociation occur more frequently in Post-traumatic Stress Disorder than depression (Reynolds & Brewin, 1999), we considered these processes to be qualitatively different from processes involved in depression vulnerability and hence we excluded individuals who met criteria for PTSD. The one exception was that we chose to include an individual with Binge Eating Disorder that did not cause significant functional impairment and did not appear to play a prominent role in this participant's general emotional responding.

Of the 102 individuals contacted, 22 participants (21.6%) were entered into the study (see CONSORT diagram, Figure 1). Thus, participants were 22 formerly depressed adults who had responded to research-protocol CT (described below) approximately 2-6 years ago (Table 1). All participants provided informed consent using forms approved by the University of Pittsburgh Institutional Review Board.

Table 1. Sample characteristics

Characteristic	Complete sample (*n* = 22)
Female, *n* (%)	15 (68%)
Years since Acute CT completion, *M (SD), min-max*	4.50 (1.10), 1.82-5.96
Age, in years, *M (SD), min-max*	48.86 (12.02), 24-70
Marital status, *n* (%)	
Single	12 (55%)
Married or cohabiting	7 (32%)
Separated, divorced, or widowed	3 (14%)
Education, *n* (%)	
Graduated high school or high school equivalent	2 (9%)
Part college	6 (27%)
Graduated 2 year college	3 (14%)
Graduated 4 year college	5 (23%)
Part graduate/professional school	1 (5%)
Completed graduate/professional school	5 (23%)
Ethnicity, *n* (%)	
African-American	4 (18%)
Caucasian	18 (82%)
Current QIDS score (two missing), *M (SD), min-max*	2.65 (2.66), 0-10
Current Axis I Diagnoses, *n* (%)	
Binge Eating Disorder	1 (5%)
None	21 (95%)
Current antidepressant medication, *n* (%)	3 (14%)
Current psychotherapy/counseling, *n* (%)	3 (14%)
Past MDEs (one missing, excluding six "too numerous or indistinct to count"), *M (SD), min-max*	3.4 (1.3), 2-6
Previous Research-protocol Treatment, *n* (%)	
Acute CT	6 (27%)
Acute CT + Cont. CT	6 (27%)
Acute CT + Cont. Fluoxetine or Placebo	9 (41%)
Experienced >= 1 relapse-recurrence following previous treatment	9 (41%)

2.2 PROCEDURE OVERVIEW

First we describe an overview of the experimental procedure, before providing a complete description of study instruments and tasks.

The first four participants completed the entire protocol in one day. As a result of feedback, the protocol was split into two sessions on separate days for all remaining participants (n = 18). In Session 1, participants were interviewed with the SCID to determine eligibility and they were interviewed with the Longitudinal Interval Follow-up Evaluation (LIFE) to assess relapse-recurrences since last research assessment. Participants completed the SoCT-P to describe their current CT skill use as well as a larger battery of self-report assessments. To prepare stimuli for Session 2, participants selected self-relevant dysfunctional statements from a pool of items, generated mood induction scripts, and practiced tasks (all described below).

In Session 2, participants underwent two blocks in which a sad mood induction was provided and then pupil dilation was measured during presentation of self-relevant dysfunctional statements (Figure 2). During the first block ("Naturalistic Block"), participants were asked to imagine the statements as their own thoughts and to mentally deal with them in whatever way they normally would. Participants then completed a computerized CT refresher course in reappraising negative thoughts before undergoing the second block ("Reappraisal Block"). In this block, participants were asked to provide cognitive reappraisals for each thought according to the guidelines of the refresher course. The experiment concluded with a positive mood induction to aid return to a euthymic state.

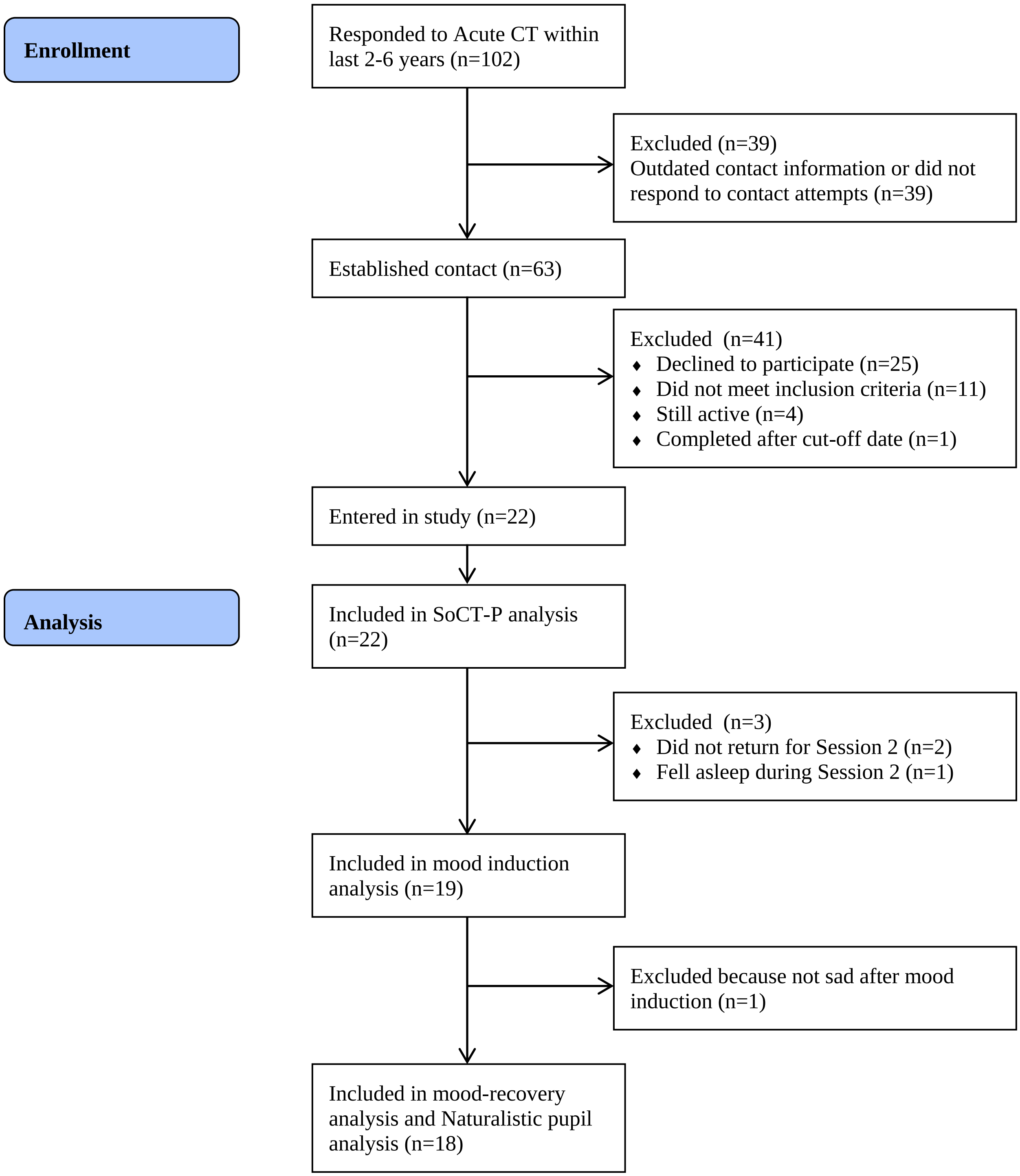


Figure 1. CONSORT chart: participant flow from initial contact to study participation

3.0 DESIGN

3.1 INSTRUMENTS

3.1.1 Instruments for screening

3.1.1.1 Structured Clinical Interview for DSM-IV Disorders (SCID)

The SCID (First et al., 1997) was used to screen out individuals currently meeting criteria for any major Axis I diagnosis, including Major Depressive Disorder. The SCID-I has demonstrated inter-rater reliability ranging from .57 to 1.0, with Major Depressive Disorder demonstrating a reliability of .80 (Zanarini et al., 2000). In terms of 7-10 day test-retest reliability, the SCID-I has demonstrated reliability ranging from .35 to .78, with Major Depressive Disorder demonstrating a reliability of .61. Miller et al. (2001) found that the SCID-I had greater convergence than an unstructured interview with a best estimate diagnosis established by expert diagnosticians using all data available, including longitudinal data. In the current study, all diagnoses were made by the first author after receiving SCID training.

3.1.2 Instruments for CT skill use

3.1.2.1 Skills of Cognitive Therapy – Patient Version (SoCT-P)

The SoCT-P (R. B. Jarrett et al., 2011) is an eight-item self-report measure that assesses understanding and use of principles and skills that are taught in CT, over the last month. Responses range from 1 ("never") to 5 ("always or when needed"). The measure has demonstrated good internal consistency, correlation with observer version, and predictive validity. Two items appear to tap cognitive reappraisal of negative thoughts: "I looked for alternative explanations when I had negative thoughts" and "I weighed the evidence for and against negative thoughts" while other items may tap reappraisal more indirectly, e.g., "I identified automatic negative thoughts and completed thought logs" and "I tested negative automatic thoughts or beliefs by setting up experiments." In the current study, we conducted sensitivity analyses that used only the two clearly reappraisal-relevant items.

3.1.3 Instruments used in stimulus generation

3.1.3.1 Sad mood scripts

To provide stimuli for the two negative mood inductions, participants composed two vignettes describing extremely sad times in their lives involving significant guilt (refer to Instructions to generate memories for sad mood inductions). We asked for memories involving not only sadness but also guilt, so as to increase salience of dysfunctional thoughts that may warrant deployment of CT skills. This allowed us to rule out sad events that might not activate dysfunctional thoughts, such as the death of a pet. Each mood induction used a different sad script, to ensure that the emotional impact of a memory would not be blunted upon repetition.

3.1.3.2 Sad music selections

The combination of reliving a sad autobiographical memory and listening to sad music has been recommended as more effective than just one or the other (Eich, Ng, Macaulay, Percy, & Grebneva, 2007). Sad music and sad memory recall has been found to be effective at inducing transient sad moods (D. M. Clark & Teasdale, 1985; Z. V. Segal et al., 1999, e.g., 2006). To provide stimuli for the two negative mood inductions, two pieces of sad music were idiosyncratically chosen from several possible musical selections (refer to Musical selections for sad mood induction). A widely used negative mood induction is the orchestral piece, "Russia under the Mongolian Yoke," played at half-speed (D. M. Clark & Teasdale, 1985). Consistent with recent recommendations (Kreutz, Ott, Teichmann, Osawa, & Vaitl, 2008) and results from our laboratory's pilot testing, we allowed participants to preview several instrumental pieces of music used in a previous study (Gemar, Kapur, Z. V. Segal, G. M. Brown, & Houle, 1996) and to choose the pieces that would be most effective at inducing a sad mood. Each mood induction used a different sad script, to ensure that the emotional impact of a memory would not be blunted upon repetition.

3.1.3.3 Happy music selections

To provide stimuli for the positive mood induction, one piece of music was idiosyncratically chosen from several possible musical selections (refer to Musical selections for happy mood induction). Consistent with recent recommendations (Kreutz et al., 2008) and results from our pilot testing, we allowed participants to listen to preview several instrumental pieces of music and to choose the piece that would be most effective at inducing a happy mood.

3.1.3.4 Inventories of dysfunctional thoughts

In order to present only self-relevant stimuli, participants idiosyncratically selected the most distressing negative statements from the Automatic Thoughts Questionnaire, Crandell Cognitions Inventory, and the Depression subscale of the Cognitions Checklist. In particular, participants were asked to imagine themselves in a sad mood and to rate how much sadder each thought would make them feel, on a scale from 1 ("no more sad") to 5 ("extremely more sad") (Instructions to select self-relevant dysfunctional thoughts). We asked participants to make ratings while imagining a sad mood because previous studies have demonstrated that recovered depressed individuals generally find few thoughts to be distressing while in their typical mood state (Z. V. Segal & Ingram, 1994). The 30 most personally distressing statements for each participant were randomly split into two non-overlapping subsets of 15 statements. Each experimental block (described below) used one of the two subsets, to ensure that the emotional impact of the statements would not be lessened upon repetition.

Automatic Thoughts Questionnaire (ATQ)

The ATQ (Hollon & Kendall, 1980) is a self-report measure that assesses frequency of 30 negative self-statements. The ATQ has demonstrated high internal reliability, with Cronbach alpha ranging from .89-.97, as well as convergent validity with the Beck Depression Inventory (Harrell & Ryon, 1983; Hollon & Kendall, 1980). Depressed outpatients receive higher scores on this measure than healthy controls and psychiatric controls (Harrell & Ryon, 1983; Hollon, Kendall, & Lumry, 1986), and the measure has also demonstrated sensitivity to treatment effects (Garratt, Ingram, Rand, & Sawalani, 2007; Simons, Garfield, & Murphy, 1984). Sample items are "I hate myself" and ‘‘I’m a failure."

Crandell Cognitions Inventory (CCI)

The CCI (Crandell & Chambless, 1986) is a self-report measure that assesses frequency of 45 negative thoughts associated with depression. The CCI assesses four categories of cognitions, corresponding to Beck's categories of cognitive distortions involving detachment, inferiority, helplessness, and hopelessness. The scale has demonstrated internal consistency with Cronbach alpha of .95 and convergent validity with the Beck Depression Inventory, as well as the ability to discriminate between depressed patients, psychiatric controls, and healthy controls. Sample items are "I'm always letting myself down" and "I mess everything up."

Depression subscale of Cognitions Checklist (CCL-D)

The CCL-D (A. T. Beck, G. Brown, Steer, Eidelson, & Riskind, 1987) is a 14-item scale assessing frequency of depressotypic cognitions. It has demonstrated high internal consistency with a Cronbach alpha of .90-.92 along with good convergent validity with the Hamilton Rating Scale for Depression and the Beck Depression Inventory. Discriminant validity has been demonstrated against the Hamilton Anxiety Rating Scale and the Beck Anxiety Inventory (A. T. Beck et al., 1987; Steer, A. T. Beck, D. A. Clark, & J. S. Beck, 1994). Sample items are "I will never overcome my problems" and "Nothing ever works out for me anymore."

3.1.4 Instruments used in the experimental procedure

3.1.4.1 Sad mood induction

Before presentation of dysfunctional statements, participants underwent a seven-minute negative mood induction in which they were asked to relive a sad autobiographical memory involving significant guilt while listening to idiographically selected sad music (see above). Participants completed the induction alone in a dimly lit room.

3.1.4.2 Sadness Scale of Positive and Negative Affect Schedule – Extended Form (PANAS-SAD)

Before and after presentation of dysfunctional statements, participants rated their current emotional state with the PANAS-SAD (Watson & L. A. Clark, 1999). Participants rated the extent to which they were feeling "sad," "blue," "downhearted," "alone," and "lonely," with responses ranging from 1 ("very slightly or not at all") to 5 ("extremely"). The PANAS sadness scale has demonstrated high internal consistency (Cronbach alpha = 0.87), good convergent validity with the Profile of Mood States ($r = 0.85$), and good construct validity as demonstrated by convergence between self and peer ratings.

3.1.4.3 Task-related Challenging

To verify that participants high in self-reported CT skill use actually challenged negative thoughts during laboratory tasks, participants were asked to complete a novel measure of task-related challenging negative thoughts. After viewing the negative statements, participants rated the frequency with which they reappraised those statements in a manner endorsed by Cognitive Therapy theorists. Six items were presented, four of which were adapted from descriptions found in "Mind Over Mood" (Greenberger & Padesky, 1995), and two of which were items from the SoCT-P as adapted for the current study. Sample items are, "How often did you think the sentences were too extreme to be true?" and "How often did you weigh the evidence for and against the sentences?" Responses ranged from 1 ("not at all") to 5 ("almost always").

3.1.4.4 Training in Challenging Negative Thoughts

To retrain all participants to reappraise negative thoughts, we used the first module of a computerized CT program that has shown efficacy at increasing CT knowledge and reducing depression ("Good Days Ahead"; J. H. Wright et al., 2005). Efficacy has been demonstrated in a randomized controlled trial in which computerized CT ameliorated depression to an equivalent extent as standard CT and superior to a wait-list control. Acquisition of CT knowledge was superior in computerized CT compared to both wait-list control and standard CT (J. H. Wright et al., 2005). In the current study, participants completed the first module ("Basic Principles") in approximately 30-40 minutes. The module uses video vignettes to illustrate negative thoughts, positive thoughts, the relationship of cognitions with emotions and behavior, and the process of "making positive changes to your thinking," i.e., reappraising negative thoughts. Participants were guided through completion of all homework assignments in the module. In these assignments, participants identified self-relevant negative and positive thoughts and then completed simplified thought records in which they provided reappraisals for the negative thoughts. Study personnel provided corrective feedback for any reappraisals that did not acknowledge the original thought and provide a response that was more positive or realistic.

3.1.4.5 Positive mood induction

To aid participants' return to a euthymic state, the experiment concluded with a positive mood induction, using 3-7 minutes of positive, ideographically selected music (see above) paired with recollection of a happy autobiographical memory (refer to Instructions to generate memory for happy mood induction).

3.1.5 Instruments used in secondary analyses

3.1.5.1 The Quick Inventory of Depressive Symptomatology (QIDS)

The QIDS was used to assess whether current results were confounded by current level of residual depressive symptomatology. This 16-item self-report measure has demonstrated high internal consistency with Cronbach alpha of .87 and good convergent validity with interviewer-rated measures of depressive symptomatology (Rush et al., 2006).

3.1.5.2 Longitudinal Interval Follow-up Evaluation (LIFE)

Data from the LIFE (Keller et al., 1987) were used in secondary analyses involving post-acute-CT relapse. The LIFE is a widely used semi-structured interview used to assess the longitudinal course of DSM-IV psychiatric disorders. It has demonstrated inter-rater intraclass correlation of .86-.88 for time to relapse-recurrence, and long-term test-retest reliability of approximately .6 for MDD (Keller et al., 1987; Warshaw, Dyck, Allsworth, Stout, & Keller, 2001). In the current study, all diagnoses were made by the first author after receiving training.

3.2 PREVIOUS PROTOCOLS

Former participants of both CT studies were adult patients with unipolar depression (relapse prevention study: 18-70 years old; study of CT response: 18-55 years old). The relapse prevention study (full methodology described in R. B. Jarrett & Thase, 2010) required one or more prior episodes of depression with complete inter-episode recovery or antecedent dysthymia, while the study of CT response also accepted individuals in their first episode.

In both studies, inclusion criteria were DSM-IV (SCID) diagnosis of MDD and a score 14 or above on the 17-item HRSD at the first two interviews. Additionally, the study of acute CT response required a score of 15 or above on the Beck Depression Inventory II at the initial screening session. In both studies, participants were excluded if they had severe or poorly controlled concurrent medical disorders that may cause depression or if they required medication that could cause depressive symptoms, if they had DSM-IV (SCID) psychotic or organic mental disorder, bipolar I or II disorder, active alcohol or drug dependence, primary obsessive compulsive disorder or primary eating disorder (where primary refers to the diagnosis associated with the most functional impairment), active suicide risk, or had previously failed to respond to a trial of at least 8 weeks of CT conducted by a certified therapist. The relapse prevention study also excluded participants who had previously not responded to 6 or more weeks of 40 mg of fluoxetine.

In both studies participants completed informed consent, underwent a supervised medication withdrawal period if medicated, then had 16-20 sessions of CT over 12-14 weeks, delivered according to the guidelines of Beck (1979). In the relapse prevention study, therapists were required to maintain Cognitive Therapy Scale scores (Young & A. T. Beck, 1980) >= 40. Experienced faculty members led weekly supervision in which groups watched and completed the CTS for randomly selected, videotaped sessions. Therapy supervisors and their teams used the CTS to provide feedback to therapists.

In contrast, in the study of acute CT response, scores >= 40 on the Cognitive Therapy Scale were achieved only on 41 of 56 (73%) of rated videotapes (M = 43.9, SD = 10.0, min = 19, max = 61). Thus, competence in delivering CT was not as well established in the study of acute CT response.

The protocols had different continuation conditions. In the study of acute CT response, participants were not assigned research-protocol treatment following acute CT, although non-responders and partial-responders were referred for non-research follow-up treatment. Relapse-recurrence was assessed at a four-month follow-up assessment using the LIFE. In the relapse prevention study, responders were randomly assigned to continuation treatment based on their risk level. "Low-risk" responders with minimal residual symptomatology were evaluated for depressive relapses with the LIFE for 32 additional months (delivered every four months), while "high-risk" responders were randomized to one of three interventions for 8 months followed by 24 months of LIFE assessment delivered every four months. Interventions were (1) Continuation-CT, a form of CT designed to help euthymic individuals to consolidate and generalize skills learned in CT, (2) Continuation-Placebo, in which an inert drug was delivered in a double-blind fashion, and (3) Continuation-Fluoxetine, in which 10-40 mg of fluoxetine was delivered in a double-blind fashion.

3.3 TASK

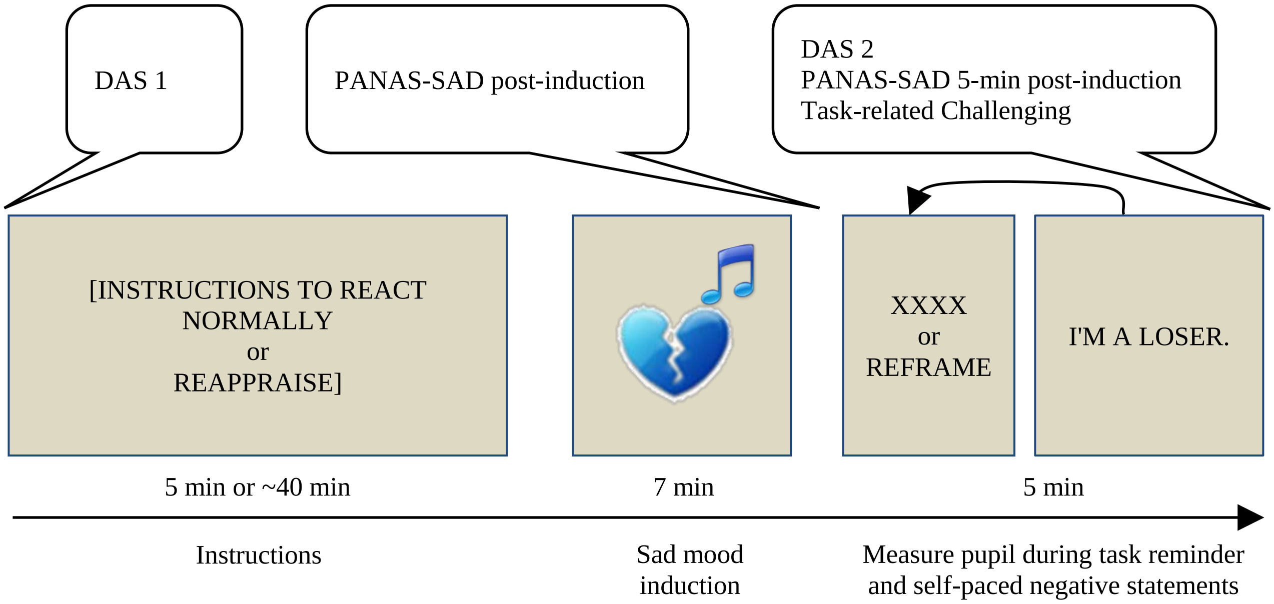
Session 2 consisted of two experimental blocks presented sequentially (Figure 2). For all participants, the "Naturalistic Block," in which participants were asked to process dysfunctional thoughts as they normally would, preceded the "Reappraise Block," in which participants were instructed to deploy CT skills by cognitively challenging each dysfunctional thought, e.g., thinking of ways in which the statement is not true or mentally responding to the thought with a more realistic statement. The block order was fixed so that instructions to reappraise would not bias individuals' naturalistic responses.

Each block began with instructions on how to cognitively respond to the negative statements that would be presented on a computer screen. In the Naturalistic Block, participants were asked to imagine the statements as their own thoughts and to deal with them as they normally would (refer to Instructions for laboratory task, Naturalistic Block). These instructions were designed to provoke the typical manner in which participants cope with negative thoughts in their everyday life, and hence to maximize ecological validity of results. In contrast, during the Reappraise Block, participants were asked explicitly to reappraise each sentence (refer to Instructions for laboratory task, Reappraisal Block) according to the guidelines of the ~40-minute course of computerized CT. We included this block to test whether individual differences demonstrated under naturalistic conditions were upheld in a more tightly controlled condition with explicit instructions to deploy CT skills.

After receiving the instruction, in the Naturalistic Block participants completed the Dysfunctional Attitudes Scale (DAS) to provide a baseline measure of dysfunctional attitudes used to calculate cognitive reactivity. In both blocks, participants were then administered a negative mood induction intended to activate the kind of dysfunctional interpretations of stressful life events that is thought to underlie vulnerability to depression recurrence (Z. V. Segal et al., 1999). Since a stress provocation of some sort is generally necessary to reveal latent dysfunctional processing and relapse-recurrence vulnerability in recovered depressed individuals (Z. V. Segal & Ingram, 1994), we reasoned that CT skills would be critical to assess in the context of a stress provocation.

After the mood induction, mood was assessed after the induction with the Sadness subscale of the Positive and Negative Affect Schedule – Extended Form (PANAS-SAD). Pupil dilation was then measured (as described below) to assess effort and emotional engagement during presentation of the previously selected 15 most self-relevant dysfunctional statements.

To ensure that all pupil measurements were task-related, participants were instructed to advance to the next statement whenever they found themselves getting off-task or distracted (cf. Nolen-Hoeksema & Morrow, 1993). This was meant to decrease the likelihood that participants would engage in significant mind-wandering during pupil assessment. In the Naturalistic Block, each statement was preceded by a 2500-ms fixation mask (row of Xs) to establish a baseline measurement for each trial. In the Reappraise Block, each statement was preceded by 2000-ms reminder text ("REFRAME"), followed by 500-ms of display of a row of "X"s. For those participants who advanced through all 15 items before 5 minutes, statements were repeated. After 5 minutes, participants completed the PANAS-SAD again to assess recovery of sad mood, followed by a battery of self-report measures, including the Challenging Negative Thoughts Scale to assess challenging of task sentences and, in the Naturalistic Block only, the alternative form of the DAS to compute the cognitive reactivity score.



Participants underwent two blocks in which a sad mood induction was provided and then pupil dilation was measured during 5 minutes of self-paced presentation of self-relevant dysfunctional statements. During the first block ("Naturalistic Block"), participants were asked to imagine the statements as their own thoughts and to deal with them in whatever way they normally would. During the second block ("Reappraisal Block"), participants were asked to provide cognitive reappraisals for each thought according to the guidelines of a computerized CT refresher course. DAS 1 = first presentation of Dysfunctional Attitudes Scale, DAS 2 = alternative form of Dysfunctional Attitudes Scale. DAS was measured only during the Naturalistic Block. PANAS-SAD = Sadness scale of the Positive and Negative Affect Schedule.

Figure 2. Task design

A positive mood induction was administered to aid participants' return to a euthymic state.

Participants who were part of the original one-session protocol (*n* = 4) were compensated $20 for their time. Participants who were part of the two-session protocol were compensated $35 for attending both sessions (*n* = 16) or $15 if they did not return for the second session (*n* = 2).

All subjects were paid an additional $10 bonus at the second session if they arrived on-time to both sessions.

3.4 PHYSIOLOGICAL DATA COLLECTION

Stimuli were displayed in black on a light gray computer screen to minimize dilation to changes in illumination associated with stimulus onset and offset. A row of Xs was displayed before stimulus onset to separate trials and to minimize any subsequent light reflect. Participants sat approximately 28.75" from the bottom of the stimulus. Stimuli were sentence-case letters approximately .25 inches high, subtending .009° of visual angle. Data were collected using the ISCAN RK464 pupillometer. The pupillometer consisted of a video camera and infrared light source pointed at a participant's left eye and a device that tracked the location and size of the pupil using these tools. Pupil size was recorded at 60 Hz and passed digitally from the pupillometer to a computer that stored the acquired data along with signals marking the beginning and end of trials. Data collection was managed using ISCAN's included software. Testing occurred with the participant alone in a moderately lit room. We developed and presented the sentences task using E-Prime 2 software (Schneider, Eschman, & Zuccolotto, 2002). Because the mood induction task had already been developed for the original version of E-Prime, we used the older software version to present mood induction tasks. Differences in software version are not thought to affect study results.

3.5 PREPROCESSING

Data were cleaned using our lab's standard methodology (Siegle, Ichikawa, & Steinhauer, 2008), including blink and artifact rejection. Eye blinks were identified and replaced by linear interpolations from 4 samples before to 4 samples after the identified blink. Trials were removed from consideration if they consisted of over 50% blinks or had pupil dilation values greater than 4 standard deviations from the within-subject mean. Data cleaning procedures resulted in the elimination of median *Md* = 7.0, mean ± standard deviation, range, *M*= 12.4 ± 12.9, 1 - 42, trials per subject out of *Md* = 61.5, *M* = 69.2 ± 32.1, 29-124, trials presented. Expressed as a proportion, cleaning resulted in discarding *Md* = 8.9%, *M* =14.7% ± 10.6%, 3.4% - 36.5%, of total trials. These statistics exclude participants with missing data from the Reappraisal Block (*n* = 2). Data were smoothed by applying a 3-point unweighted average filter twice. Pupil diameter, measured at time of participant's button press, was subtracted from subsequent trial-related pupil diameter to produce pupil motility indices.

Within each subject, mean pupil dilation for each block (Naturalistic Block or Reappraisal Block) was computed by averaging only the trials that lasted to a given sample or longer. Thus, a condition-related mean for the first sample would be calculated from 100% of the trials – since all trials were at least one sample long – whereas the condition-related mean of a later sample would be calculated from only those trials that lasted at least to that point. To provide adequate signal-to-noise ratio, we chose to end a condition-related mean when it would be comprised of fewer than four trials.

3.6 COGNITIVE REACTIVITY

Cognitive reactivity was employed as an analog for both effectiveness of coping and relapse-recurrence risk status. This construct is conceptualized as the ease to which a dysfunctional thinking style can be made accessible when provoked in some way (Z. V. Segal et al., 2006). It is calculated by subtracting an initial score on the Dysfunctional Attitudes Scale (DAS; Weissman & A. T. Beck, 1978) to a score of the alternative-form DAS delivered following a stress provocation such as a sad mood induction. Segal et al. (1999, 2006) have demonstrated that change in dysfunctional attitude endorsement is positively associated with subsequent relapse-recurrence. Thus, this measure may be used as a potential marker of relapse-recurrence risk. In the current context, cognitive reactivity may also tap effectiveness of challenging, since it was measured following cognitive responses to negative thoughts.

3.7 DATA ANALYSIS

To assess frequency of CT skill use years after CT completion, descriptive statistics of current SoCT-P scores were examined and summarized. To compare CT skill use years after treatment to previous CT skill use (Figure 3; path a), we tested the effect of time on SoCT-P score with a mixed effects analysis with Time (mid-acute, post-acute, 3-month follow-up, current) as an ordinal repeated fixed effect with AR1 autocorrelation structure, subject as a random effect, and SoCT-P as the dependent variable.

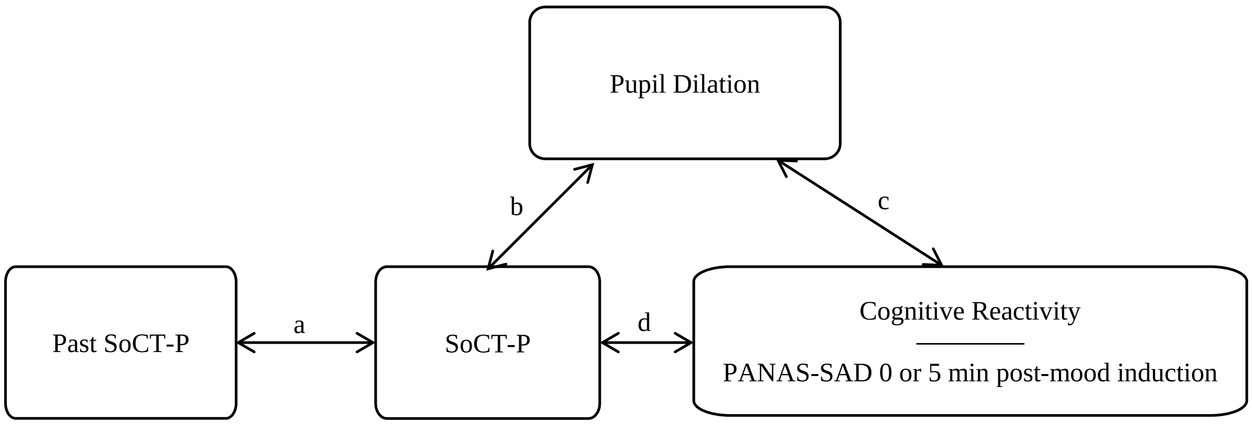
To assess the relationship between CT skill use and effort and emotional engagement in response to negative sentences (path b), we tested the correlation between SoCT-P and mean

Naturalistic Block pupil dilation at every sample. We used Guthrie and Buchwald's (1991) technique for controlling Type 1 Error, as it has been adapted for pupil dilation data (Siegle, Steinhauer, & Thase, 2004; Siegle, Steinhauer, Carter, et al., 2003; Siegle, Steinhauer, Stenger, et al., 2003). In particular, for each sample, we calculated the correlation between mean pupil dilation and SoCT-P. A correlation at any particular sample was judged to be statistically significant if and only if data from several consecutive samples were also significantly correlated with SoCT-P. We used simulations to determine the number of significant consecutive tests required to judge a region as statistically significant. We simulated 1000 sets of 18 randomly generated waveforms (corresponding to the current sample size of $n = 18$). To better represent current data, waveforms were constrained to have an autocorrelation equivalent to that found in the current data. Autocorrelation of the current data was defined using Guthrie and Buchwald's (1991) method, with principal components set *a priori* at 3, yielding $r_{xx} = 0.96$. After generating the random waveforms, we calculated a temporal contiguity threshold such that, in less than 5% of simulations, there would be a time period longer than the threshold during which waveforms were significantly correlated with arbitrary values (i.e., random numbers). Simulation results indicated that 483 ms of contiguous tests significant at $p < 0.10$ would be needed to judge a given window significant at $p < 0.05$. In an exploratory fashion, we examined what results would be obtained when setting principal components to be 4 or 5, resulting in $r_{xx} = 0.95$ or 0.94 and temporal thresholds of 417 ms or 350 ms.

To assess the extent to which these results were likely consistent with reappraisal processes, time intervals with significant correlations during the Naturalistic Block were examined in the Reappraise Block for a significant correlation in the same direction.

Finally, to assess the effectiveness of coping among CT skill users (path d), we assessed the correlations between SoCT-P scores with two indices of coping success: PANAS-SAD score after sentence viewing ("Residual Sadness") and cognitive reactivity. To examine robustness of effects, PANAS-SAD score from immediately following the mood induction was stepped into the model first to test whether pupil dilation (path c) and SoCT-P explained significantly more variance in residual sadness. Then interaction terms were entered as a third step. To assess generalizability, significant findings from paths b and c were re-examined in Reappraisal Block.

In all analyses, variables were Winsorized to minimize the influence of potential outliers. In particular, values greater than the 75th percentile + 1.5 * the interquartile range, were reduced to the nearest valid value, and values lesser than the 25th percentile – 1.5 * the interquartile range, were increased to the nearest valid value.



Path a represents the trajectory of CT skill use over time, leading to its current value 2-6 years following CT. Path b represents level of effort and emotional engagement involved in how CT skill users respond to negative thoughts. Path c represents the extent to which level of effort and emotional engagement is linked to adaptive emotional outcomes, such as lack of increase in dysfunctional attitudes following a mood induction ("cognitive reactivity"), and residual sadness five minutes after a mood induction. Path d represents whether CT skill users are more likely to experience these adaptive emotional outcomes compared to CT skill non-users.

Figure 3. Relationships examined in the present study

4.0 RESULTS

4.1 TO WHAT EXTENT ARE CT SKILLS STILL USED YEARS AFTER TREATMENT? (FIGURE 3, PATH A)

SoCT-P responses indicated that, on average, participants reported using CT skills between "half" and "most" of the time in the past month (Table 2). Mixed effects analysis revealed a significant effect of Time (mid-acute, post-acute, 3-month follow-up, current) on SoCT-P (Figure 4), p = 0.008, η^2 = 0.19. Post-hoc comparisons indicated that current SoCT-P was not significantly different from mid-therapy scores, p = 0.69, D = -0.10, significantly lower than post-acute SoCT-P, p = 0.04, D =.49, and not significantly different from 4-month follow-up scores, p = 0.82, D = 0.05.

This pattern of results suggested that CT skills may have increased from mid- to post-acute before returning to mid-acute levels and remaining at that value. To test this interpretation, we conducted post-hoc tests comparing SoCT-P at each time period with SoCT-P at the subsequent time period. Mid-acute SoCT-P was significantly lower than post-acute SoCT-P, p = 0.0001, D =.58, which was, in turn, significantly higher than 4-month follow-up SoCT-P, p = 0.003, D = 0.43.

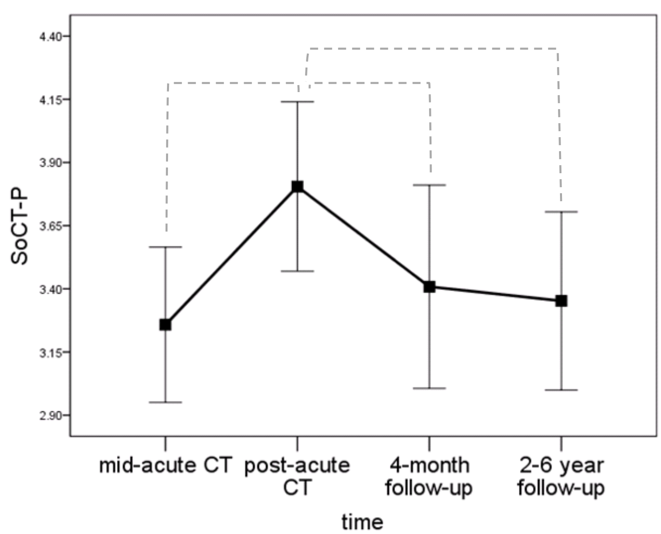
Table 2. Change in self-reported use of CT skills over time

	Mid-acute	Post-acute	4 month follow-up	2-6 year follow-up
n	16	16	15	22
M	3.26_a	3.80_b	3.41_a	3.35_a
SD	0.58	0.63	0.73	0.79
Minimum	2.25	2.50	1.88	1.88
Maximum	4.25	4.75	4.38	4.88

"Mid-acute" refers to the first time the measure was administered, during the midpoint of 16-20 week Cognitive Therapy. "Post-acute" refers to the end of the acute therapy regimen. "4 month follow-up" is self-explanatory. "2-6 year follow-up" refers to scores from the present study.

Means with different subscripts are significantly different, $p < 0.05$.

SoCT-P: Skills of Cognitive Therapy questionnaire.



In the current sample, self-reported use of CT skills, as indexed by the SoCT-P increased from mid-acute CT to post-acute CT before returning to mid-acute level for the 4-month follow-up and for the current 2-6 year follow-up. Error bars represent 95% confidence intervals. Starred comparisons are significant at $p < 0.05$.

Figure 4. Change in self-reported use of CT skills over time

4.2 TASK MANIPULATION CHECKS

Analyses of usefulness and effort of CT skill depend on the success of laboratory manipulations. Thus, we first sought to verify the effectiveness of our manipulations.

4.2.1 Were participants sad after the mood inductions?

After the first mood induction, eighteen out of nineteen participants reported sadness on the PANAS-SAD. The one participant who reported no sadness was excluded from further analyses on this block. In the remaining sample, mean response on PANAS-SAD items fell between 2 (i.e., "a little" sad, blue, downhearted, alone, and lonely) and 3 ("moderately" sad, etc.), suggestive of mild sadness (total PANAS-SAD $M = 11.00$, $SD = 3.73$).

After the second negative mood induction, sixteen participants reported sadness on the PANAS-SAD. The three participants who reported no sadness were excluded from further analyses of this block. In the remaining sample, mean response on PANAS-SAD items was 2 ("a little" sad, etc.), again suggestive of mild sadness (total PANAS-SAD $M = 10.19$, $SD = 3.13$).

A paired t-test comparing PANAS-SAD following the first mood induction versus PANAS-SAD following the second mood induction did not reveal a statistically significant difference, $t(15) = 0.89$, $p = 0.39$, $D = 0.44$, $d = 0.22$.

We concluded that there was no apparent difference between sadness elicited in the first induction versus the second induction, and that participants were mildly sad following both inductions. This level of sadness was of the same magnitude of that found in other mood induction studies of recovered depressed individuals. After normalizing sadness ratings to 0-100%, sadness in the current study corresponded to 28%, whereas other studies have measured mean ratings of 6% (Gemar, Z. V. Segal, Sagrati, & S. J. Kennedy, 2001), 30% (Z. V. Segal et al., 2006), and 17% (Gilboa & Gotlib, 1997). The mood inductions appeared to be successful.

4.2.2 Was trait CT skill use related to CT skill use during the experimental task?

To guide interpretation of task-related results, we tested whether those who scored higher on the SoCT-P also reported greater challenging of negative thoughts during the laboratory assessment.

In the Naturalistic Block, there was a non-significant association of moderate effect size between SoCT-P and Task-related Challenging, $F(1,17) = 4.27$, $r = 0.46$, $p = 0.055$.

In contrast, for the Reappraisal Block, Task-related Challenging was minimally and not significantly correlated with SoCT-P scores, $F(1,14) = 0.12$, $r = 0.09$, $p = 0.74$. Since the Reappraisal Block was preceded by a refresher CT training course, the training may have caused individuals to calibrate their frequency of challenging negative thoughts to the same level. On this interpretation, one would predict a negative association between SoCT-P and increase in Task-related Challenging from the Naturalistic Block to the Reappraisal Block, such that low-SoCT-P individuals would increase challenging more than high-SoCT-P individuals. More power is needed to support or disconfirm this claim: the correlation between SoCT-P and change in Challenging was of moderate magnitude but not statistically significant, $F(1,14) = 2.36$, $r = -0.39$, $p = 0.15$.

In summary, during the Naturalistic Block there was a moderate relationship between trait and state CT skill use, while during the Reappraisal Block this relationship was not detected. Results from the Naturalistic Block may be more interpretable than those from the Reappraisal Block.

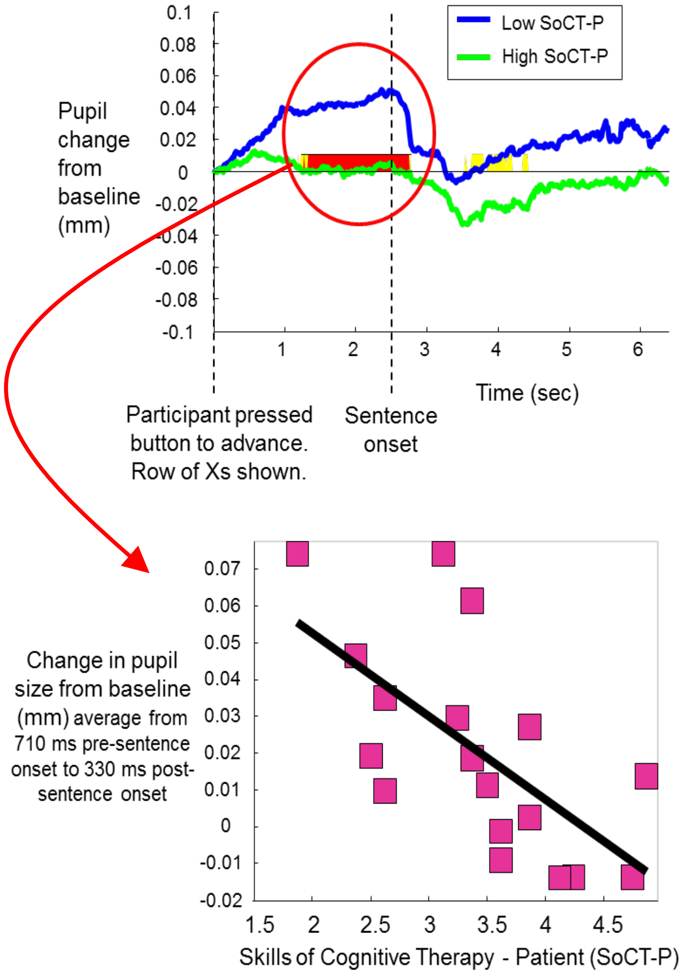
4.3 DO CT SKILL USERS EXPEND SIGNIFICANT EFFORT AND EMOTIONAL ENGAGEMENT WHILE VIEWING NEGATIVE THOUGHTS? (FIGURE 3, PATH B)

In the Naturalistic Block, significant negative correlations (*p*s < 0.1) emerged between SoCT-P and pupil dilation at several time regions. The first region was 1230 – 1540 ms following trial onset, $F(1,17) = 4.01$, $r = -0.45$, $p = 0.06$. However, the length of significant contiguous correlations (310 ms) did not surpass the temporal contiguity threshold needed to control Type 1 error at $p < 0.05$ (483 ms, or in an exploratory fashion, 450 ms or 317 ms). Therefore, this effect was not investigated further.

Significant negative correlations between SoCT-P and pupil dilation (*p*s < *0*.1) were also present 1790 – 2830 ms following trial onset (Figure 5), $F(1,17) = 11.40$, $r = -0.64$, $p = 0.004$. This region spanned 710 ms prior to sentence onset through 330 ms following sentence onset. In the Reappraisal Block, there was not a statistically significant relationship between SoCT-P and pupil dilation during this time interval, $F(1,15) = 1.19$, $r = -0.28$, $p = 0.29$. Throughout the entire Reappraisal Block, pupil dilation was not significantly correlated with SoCT-P.

We conducted follow-up tests to determine whether lack of statistical significance in Reappraisal Block might be attributed to task differences as opposed to a lower signal-to-noise ratio. First, we examined whether pupillary responses to the Reappraisal Block during the previously identified region, 1790 – 2830 ms, were significantly different from those during Naturalistic Block. If so, this could indicate the presence of different cognitive and emotional processes involved in the two tasks rather than greater noise. A paired t-test revealed that Reappraisal pupil dilation was significantly lower than Naturalistic pupil dilation during this time period, $t(15) = -3.05$, $p = 0.01$, $D = -0.02$, $d = -0.76$. This result could indicate that relationships between SoCT-P and pupil were not evident in the Reappraisal Block because in that block low-

SoCT-P individuals reduced their pupillary responses, equaling the pupillary responses of high-SoCT-P individuals. However, again, more power is needed to test this hypothesis: SoCT-P score was moderately but nonsignificantly correlated with degree of pupil change from the Naturalistic Block to the Reappraisal Block, $F(1,15) = 2.74$, $r = 0.40$, $p = 0.12$.



For illustrative purposes, the top panel shows participants as part of two groups separated by a median split based on SoCT-P. It is evident in the top panel that high CT skill users have significantly lower pupil responses than low CT skill users, from 710 ms pre-sentence to 330 ms post-sentence. In the bottom figure, pupil values are averaged across the region of contiguous significance tests and plotted against SoCT-P, resulting in a significant negative correlation between SoCT-P and pupil dilation during this time interval, $p < 0.004$.

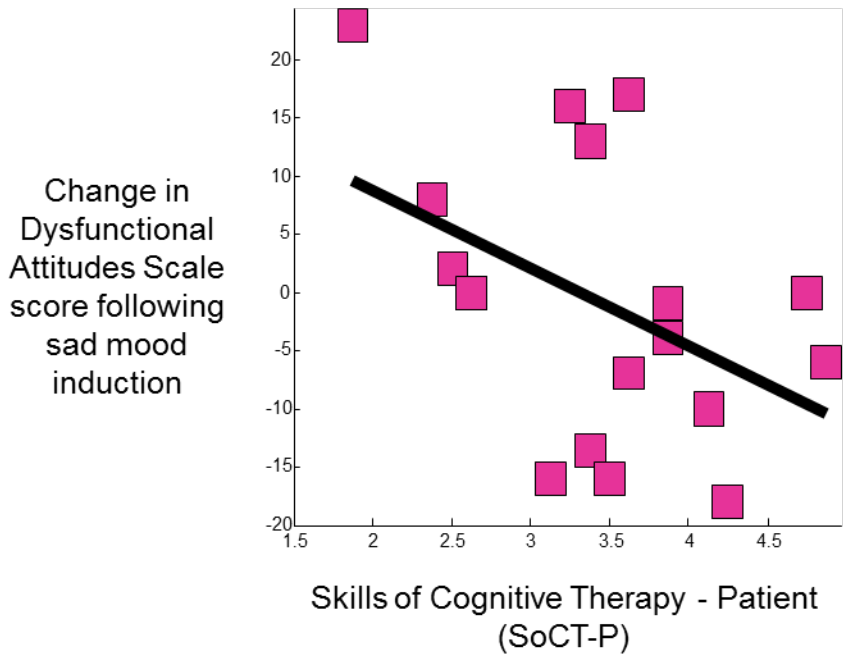
Figure 5. Relationship of pupil motility to self-reported use of CT skills

4.4 DO CT SKILL USERS DEMONSTRATE MORE ADAPTIVE EMOTIONAL OUTCOMES? (FIGURE 3, PATH D)

In the Naturalistic Block, SoCT-P was non-significantly negatively associated with sadness remaining after the sentences task with a moderate effect size, $F(1,17) = 4.07$, $r = -0.45$, $p = 0.06$.

However, SoCT-P was significantly negatively associated with sadness immediately following the mood induction, $F(1,17) = 4.53$, $p = 0.049$, $r = -0.47$, which in turn was significantly correlated with sadness remaining after the sentences task, $F(1,17) = 11.29$, $r = 0.64$, $p = 0.004$. We tested whether SoCT-P was correlated with sadness-remaining above and beyond that contributed by the first sadness rating. The addition of SoCT-P to the model did not significantly increase the F value, $\Delta F(1,15) = 0.76$, $p = 0.40$, $\Delta R^2 = 0.03$. Thus, the association between SoCT-P and sadness after the task could have been due to an effect of SoCT-P on the effectiveness of the mood induction itself or to a relationship of SoCT-P with baseline level of sadness.

SoCT-P scores were non-significantly correlated with cognitive reactivity with a moderate effect size, $F(1,17) = 4.06$, $r = -0.45$, $p = 0.06$ (Figure 6). In this case, SoCT-P does not appear to be a proxy for mood, as mood itself was minimally and not significantly associated with cognitive reactivity, $F(1,17) = 0.13$, $r = 0.09$, $p = 0.73$. After including sadness immediately following the mood induction as a first step, the addition of SoCT-P into the model significantly increases the variance explained in cognitive reactivity, $\Delta F(1,15) = 5.71$, $p = 0.03$, $\Delta R^2 = 0.28$. Adding an interaction term of mood and SoCT-P did not explain significantly more variance, $\Delta F(1,14) = 0.19$, $p = 0.67$, $\Delta R^2 = 0.01$. Furthermore, SoCT-P was unrelated to baseline level of dysfunctional attitudes, $F(1,17) = 0.65$, $r = -0.20$, $p = 0.43$. Thus, SoCT-P appears to be a reasonable albeit not statistically significant predictor of cognitive reactivity.



Although not statistically significant, a moderate effect size supported the notion that individuals who used CT skills more frequently had lower cognitive reactivity, i.e., they were less likely to increase endorsement of dysfunctional attitudes following a sad mood induction.

Figure 6. Relationship between self-reported use of CT skills and cognitive reactivity

4.5 IS EFFORT AND EMOTIONAL ENGAGEMENT RELEVANT TO COPING OUTCOMES? (FIGURE 3, PATH C)

In the Naturalistic Block, significant correlations emerged between sadness remaining after the sentences task and pupil dilation at 1180 – 1710 ms following trial onset, $F(1,17) = 4.80$, $r = 0.48$, $p = 0.04$. To test whether pupil dilation here was more than a proxy for mood, we entered pre-task sadness in the first step of a linear regression and then entered in a term representing mean pupil dilation during this time interval. The addition of pupil dilation to the model nonsignificantly explained more variance with a moderate effect size, $\Delta F(1,15) = 2.74$, $p = 0.12$, $\Delta R^2 = 0.13$. The further addition of a pupil x initial mood interaction term to the model did not explain more variance, $\Delta F(1,15) = 0.08$, $p = 0.79$, $\Delta R^2 = 0.00$. Since results suggest that pupil dilation may be a proxy for mood here, we tested correlations between pre-task sadness and pupil dilation. Results suggested significant positive correlations of large magnitude, 2180 ms to 3940 ms, interval = 1760 ms, $F(1,17) = 9.08$, $p = 0.008$, $r = 0.60$. It is concluded that the relationship of pupil dilation to sadness after the sentences task may be at least partially explained by the relationship of pupil dilation to sadness before the sentences task.

The second coping outcome measure, cognitive reactivity, was significantly correlated with pupil dilation 1260-1740 ms following trial onset, $F(1,17) = 4.57$, $p = 0.048$, $r = 0.47$, and at 1950-2220 ms, $F(1,17) = 4.68$, $p = 0.046$, $r = 0.48$. The first temporal region was long enough only to pass the exploratory contiguity threshold, while the second temporal region was not long enough to survive thresholding. Results from these two time regions did not generalize to the Reappraisal Block, $F(1,17) = 0.94$, $p = 0.35$, $r = 0.23$, and $F(1,17) = 0.004$, $p = 0.95$, $r = 0.02$, respectively.

Post-sentences sadness was moderately but non-significantly correlated with pupil dilation during the period in which SoCT-P and pupil dilation are correlated, $F(1,17) = 3.77$, $r = 0.44$, $p = 0.07$, as was cognitive reactivity, $F(1,17) = 2.95$, $r = 0.39$, $p = 0.11$.

We tested for outliers in all analyses resulting in borderline significance. Outliers emerged in the correlation between pupil dilation and cognitive reactivity: two individuals met criteria for both high influence (Cook's D > 4/(n – k – 1); Fox, 1991) and high leverage (leverage

> 2p/n; Hoaglin & Welsch, 1978), suggesting that they may be considered outliers. After excluding these two individuals, there was a significant positive correlation between pupil dilation and cognitive reactivity, $F(1,16) = 10.66$, $r = .51$, $p = 0.04$.

4.6 SENSITIVITY ANALYSES

Using only the two SoCT-P items that assess challenging of negative thoughts, a significant correlation of moderate effect size still remained between the SoCT-P and Naturalistic Block pupil dilation from 1790 – 2830 ms following trial onset, $F(1,17) = 10.123$, $r = -0.62$, $p = 0.006$.

We also tested whether pupil dilation during this time period was also related to Task-related Challenging. We observed a significant negative correlation of moderate effect size between pupil dilation and Task-related Challenging, $F(1,17) = 5.075$, $r = -0.49$, $p = 0.04$.

Excluding individuals taking psychiatric medication did not change eliminate the correlation between SoCT-P and Naturalistic Block pupil, 1790 – 2830 ms following trial onset, $F(1,14) = 5.89$, $p = 0.03$, $r = -0.56$. Excluding those who had received Continuation Cognitive Therapy did not change this result either, $F(1,12) = 8.58$, $p = 0.01$, $r = -0.66$.

We confirmed that pupil results were not an index of current depressive symptomatology as measured by the QIDS. Nowhere during the Naturalistic Block pupil dilation time-series was pupil dilation significantly correlated with QIDS, all ps > 0.2, all R^2s < 0.10.

4.7 EXPLORATORY ANALYSES

Since, on average, current SoCT-P was not significantly greater than scores at any other time period, we sought to test whether current SoCT-P represents any skill use or, alternatively, whether current SoCT-P represents the lowest score that participants generally provide. Since no baseline measure of pre-treatment SoCT was collected, we chose a comparison point of individuals with the lowest mean: individuals who do not respond to acute CT. At mid-therapy, those who go on to not respond to CT demonstrate lower SoCT-P than those who do eventually respond to CT (R. B. Jarrett et al., 2011). Thus, to confirm that CT skill use in the present sample did not represent the lowest scores measurable, we compared SoCT-P in the current sample and with mid-acute SoCT-P in non-responders within the Pittsburgh arm of the relapse prevention study. Current SoCT-P was nonsignificantly greater than mid-acute SoCT-P in non-responders, with a small-to-medium effect size, $t(52) = 1.46$, $p = 0.15$, $D = 0.29$, $d = 0.40$. At mid-acute, non-responders ($n = 32$) on average endorsed using CT skills "half the time" over the last month ($M = 3.06$, $SD = 0.66$, min = 1.58, max = 4.38), while the current sample ($n = 22$) on average endorsed using CT skills between “half” and “most” of the time in the past month ($M = 3.35$, $SD = 0.79$, min = 1.88, max = 4.88).

Exploratory analyses indicated no significant relationship between SoCT-P and mean sentence viewing time during Naturalistic Block, $F(1,17) = 0.39$, $r = -0.15$, $p = 0.54$, or Reappraisal Block, $F(1,15) = 0.01$, $r = 0.03$, $p = 0.91$.

SoCT-P scores were not significantly related to post-acute relapse as measured by the LIFE, $t(20) = 0.84$, $p = 0.41$, $D = 0.29$, $d = 0.36$. Pupil dilation during the SoCT-P-relevant time interval was also not related to post-acute relapse, $t(16) = -1.41$, $p = 0.18$, $D = 0.02$, $d = 0.67$.

Pupil dilation from 2550 to 3420 ms, was significantly associated with post-acute relapse, $t(16) = 2.13$, $p = 0.049$, $D = 0.03$, $d = 1.01$, such that individuals with lower pupil dilation during that time period were less likely to have had a relapse following acute CT.

We conducted additional exploratory analyses to investigate whether anticipatory anxiety might be related to the unexpected timing of the current results, i.e., SoCT-P was related to pupillary dilation *before* each sentence was displayed. Although we did not have measures of anxiety, we tested whether history of anxiety disorder accounted for these results. History of anxiety disorder was not significantly associated with pupil dilation during the SoCT-P-relevant time interval, $t(16) = 0.73$, $p = .48$, $D = 0.01$, $d = 0.35$, although there was a significant association between past anxiety and pupil dilation during an earlier time period (100-350 ms) that did not pass temporal contiguity threshold, $t(16) = 2.36$, $p = 0.03$, $D = 0.01$, $d = 1.14$.

We further sought to examine whether the pre-stimulus timing of the current results could reflect carryover processing from the previous trial. For each trial, we first expressed pupillary dilation in terms of previous trial activity, by subtracting pupil dilation from the sentence onset of the *previous* trial from the pupil dilation during each sample of the current trial. After recomputing mean pupil dilation at each sample, mean pupil dilation remained significantly correlated with SoCT-P during the same time period as found in the primary analysis, 710 ms prior to sentence onset through 330 ms following sentence onset, $F(1,17) = 6.78$, $p = 0.02$, $r = -0.55$.

We investigated late pupil dilation, defined as 4000-6000 ms after trial onset (based on visual inspection of mean pupil dilation curves). Current SoCT-P was nonsignificantly negatively associated with pupil dilation during this period, $F(1,15) = 0.621$, $p = 0.44$, $r = -0.21$.

To better understand the unexpected association between SoCT-P and PANAS-SAD immediately post-mood induction, we investigated other variables that may be associated with PANAS-SAD immediately post-mood induction. Scores on this measure were not significantly related to QIDS scores, $F(1,19) = 0.80$, $r = -0.21$, $p = 0.38$, to cognitive reactivity, $F(1,17) = 0.64$, $r = 0.20$, $p = 0.44$, to last known 17-item Hamilton Rating Scale for Depression from their previous relapse prevention study, $F(1,17) = 0.46$, $r = -0.17$, $p = 0.51$, or to last known BDI from the previous study, $F(1,17) = 1.0$, $r = -0.24$, $p = 0.33$. PANAS-SAD did not appear to be related to the quantity of current SCID MDE symptoms, as only one participant in the current sample met criteria for any MDE symptoms.

Finally, we analyzed to what extent marginally nonsignificant results could be related to low statistical power. As shown in Table 3, at least 16 more participants would be necessary to be adequately powered to detect these effects.

Table 3. Post-hoc power analyses for correlations of marginal significance

Analysis type	Variable 1	Variable 2	Effect size	p	Current power	Current n	*n* needed for 0.80 power
Manipulation check	SoCT-P	Naturalistic Challenging	*r* = 0.46	0.06	0.51	18	34
Primary	SoCT-P	PANAS-SAD 5 min post-induction	*r* = -0.45	0.06	0.49	18	36
Primary	SoCT-P	Cognitive reactivity	*r* = -0.45	0.06	0.49	18	36
Primary	Naturalistic pupil dilation during SoCT-P-relevant time	PANAS-SAD 5 min post-induction	*r* = 0.44	0.07	0.47	18	38
Primary	Naturalistic pupil dilation during SoCT-P-relevant time	Cognitive reactivity	*r* = 0.39	0.11	0.37	18	49
Sensitivity	Naturalistic pupil dilation	PANAS-SAD 5 min post-induction after accounting for PANAS-SAD 0 min post-induction	$\Delta R^2 = 0.13$	0.12	0.30	16	55
Exploratory	SoCT-P	ΔChallenging from Naturalistic to Reappraisal	*r* = -0.39	0.15	0.31	15	49
Exploratory	SoCT-P	ΔPupil dilation from Naturalistic to Reappraisal	*r* = 0.40	0.12	0.35	16	46
Exploratory	SoCT-P	Mid-acute SoCT-P of non-responders	*d* = .40	0.15	0.29	54	328
Exploratory	SoCT-P	Late pupil dilation (4-6 sec after trial onset)	*r* = -.21	0.44	0.12	16	289

SoCT-P: Skills of Cognitive Therapy – Patient questionnaire.

PANAS-SAD: Sadness subscale of the Positive and Negative Affect Schedule - Extended

5.0 DISCUSSION

Here we investigated the use, usefulness, and effort involved in deploying a core CT skill, cognitive reappraisal of negative thoughts, years after CT is completed. To assess frequency of CT skill use, we asked individuals who had responded to research-protocol CT 2-6 years ago to complete the SoCT-P self-report measure. Results suggested that participants generally continue to make use of CT skills around half the time. Although this frequency was somewhat reduced compared to frequency immediately post-therapy, it was noteworthy that skill use did not decline from its level at 4 months post-therapy. This is consistent with the notion that CT teaches individuals tools that they can continue to use independently, without the guidance of an expert (Glasman et al., 2004). Similar effects, in which some level of therapy skill use is maintained at follow-up, have been found in investigations of CT for depression in older adults (Powers et al., 2008), Mindfulness-based Stress Reduction (Baer, 2003), and Mindfulness-based Cognitive Therapy for depressive relapse prevention (Bondolfi et al., 2010). What is particularly important in this study is the length of the follow-up. It is remarkable that participants continue to use CT-related skill up to six years later, suggesting that participants continue to find CT skills relevant long after therapy completion.

Still, caution must be warranted, as there was no baseline pre-therapy measure of CT skill use. It cannot be ruled out that participants are simply using CT-consistent skills to the same extent as they were pre-therapy. Anecdotally, most but not all participants attributed their use of reappraisal to CT, but empirically validating these reports (e.g., by including a CT skills use assessment at pre-treatment) will be important for future research.

We examined the potential usefulness of CT skills by testing the correlation between SoCT-P and two adaptive emotional outcomes: greater recovery from a sad mood and lower cognitive reactivity. With respect to the former, we observed a nonsignificant negative association of moderate magnitude between use of CT skills and prolonged sadness following a sad mood induction. However, interpretation of this result is ambiguous because SoCT-P was also associated with sad mood directly following a mood induction. In other words, individuals who use CT skills more frequently tended to have less post-mood induction sadness, which in turn predicts more prolonged sadness.

The evidence is consistent with several non-exclusive potential interpretations. CT skill users may have a lower baseline level of sadness, less reactivity to sad mood inductions, and/or greater recovery from sad mood inductions. Future studies may take several steps to disentangle these possible interpretations. First, by assessing pre-mood induction levels of sadness, the relative contributions of baseline mood and mood reactivity may be better understood. Second, with a larger sample size it may be useful to remove from analysis CT skill users who do not achieve a high level of sadness following a mood induction. In this way, level of sadness post mood-induction could be held constant and the relationship between CT skill use and mood recovery could be more clearly understood.

In contrast, results related to cognitive reactivity were not easily attributable to mood differences. CT skills use moderately though non-significantly predicted less cognitive reactivity, and it did not predict potential confounding variables (baseline level of dysfunctional attitudes or post-mood induction sadness). This suggests that individuals who continue to make use of CT skills are not as vulnerable to activation of mood-congruent dysfunctional thinking. Lower cognitive reactivity has itself predicted lower rates of relapse (Z. V. Segal et al., 1999, 2006), suggesting that CT skill use may similarly predict lower relapse rates. This is consistent with hypotheses that continued CT skill use underlies CT's enduring relapse-prevention effects (Hollon, 2003; Hollon et al., 2006; R. B. Jarrett & Kraft, 1997).

On this basis it is possible that CT skill users experience no increase in dysfunctional thoughts when primed with a negative mood, or it is also possible that CT skill users may automatically or voluntarily challenge dysfunctional thinking patterns that are activated by a mood induction. Another possibility is that CT skill use is a proxy for another variable, such as general life stress, that could be related to cognitive reactivity.

Taken together, results of the "effectiveness" analyses suggest that CT skill use is associated with adaptive emotional outcomes. This suggests that skill use may be useful to deploy even in a euthymic state, and therapists may wish to devise methods to keep patients motivated to deploy CT skills even when no longer feeling depressed.

At the same time, with a ~50% relapse rate even in CT responders (Vittengl, L. A. Clark, Dunn, & R. B. Jarrett, 2007), it is important to ask whether post-CT coping techniques can be improved. One potential area for improvement is that some CT skills, such as challenging dysfunctional negative thoughts, depend on the occurrence of those thoughts in the first place. If recovered depressed individuals experience only normative levels of dysfunctional thoughts (Z. V. Segal et al., 1999), and if CT skill practice is necessary for successful use (R. B. Jarrett et al., 2011), then post-CT individuals may not have sufficient opportunities to practice reappraising negative thoughts. If this is the case, different techniques that work without negative thoughts could be promoted for relapse prevention this time period, such as Mindfulness-Based Cognitive

Therapy (Teasdale et al., 2000). This form of therapy uses everyday experience as a target for intervention, circumventing the relative lack of dysfunctional thoughts in recovered depressed individuals.

Finally, we indexed the effort and emotional engagement involved in CT skill users' responses to negative thoughts, by measuring pupil dilation during presentation of dysfunctional self-relevant thoughts. Pupil results were robust, indicating that those who use more CT skills tended to allocate less effort and emotional engagement in response to negative thoughts. Moreover, less allocation of effort and emotional engagement was itself related to adaptive emotional outcomes, suggesting that low pupil dilation is relevant for these outcomes. This result argues against the view that former CT patients voluntarily deploy effortful compensatory strategies to counteract automatic negative information processing biases (Barber & R. J. DeRubeis, 1989; R. J. DeRubeis et al., 2008; Hollon et al., 2006). In fact, those CT patients who continue to use CT skills generally appear to expend less effort and emotional engagement when coping with negative stimuli. Such former patients do not face a life of deploying arduous CT skills, but rather are (a) less emotionally affected by negative stimuli, and/or (b) deploy less effort to cope with negative thoughts, at least relative to former CT patients who do not use CT skills as frequently.

Given that reduced effort and emotional involvement is associated with adaptive emotional outcomes, therapists may find it useful to delay treatment termination until use of CT skills become automatized. Additionally, effortful emotion regulatory activities such as completing dysfunctional thought records may be thought of as scaffolding that allows the patient to reach the point where disputation of negative thoughts requires little to no effort.

One surprising finding was that CT skill use was negatively correlated with pupil dilation before a negative sentence even appeared, during the time in which participants were aware that it would appear shortly within the next second. Here we speculate on what might account for the early timing of this effect. One possibility is that pupil differences result from anticipatory reappraisal. This would be consistent with SoCT-P and Task-related Challenging responses suggesting high use of reappraisal at both a trait and state level was associated with the decreased anticipatory pupil response. One participant, a former attorney, compared this lack of anticipatory anxiety to the feeling of calm that occurs when one knows that the opposing attorney is about to make an unconvincing case that barely warrants rebuttal. In some sense, this could be construed as anticipatory automatic emotion regulation, a construct that appears to be unexplored in the extant literature.

However, since pupillary response is sensitive to both emotion regulation and emotion reactivity. a more parsimonious explanation may be that lower pupil dilation in high-SoCT-P individuals reflects their lower level of sadness during the task. Since pupil measurements have been associated with emotional intensity (e.g., Steinhauer et al., 1983), it is possible that the pupil differences simply index these mood differences.

Another possibility is that the early pupil reaction in skill non-users represents anxious vigilance toward threat. Although the current participants did not demonstrate clinical anxiety disorders on the SCID and we did not assess non-clinical anxiety, the high comorbidity between anxiety and depression (Kaufman & Charney, 2000) and high incidence of past anxiety disorders in the current sample (45%) suggest that anxiety may be present. Koster et al. (2005) found that high trait anxiety even in the absence of a clinical disorder predicted heightened early attention to threatening stimuli, suggesting that vigilance or anticipation of threatening stimuli may be prominent in such individuals. Bradley et al. (1997) found that dysphoric individuals demonstrated early attention to depression-related words, and this tendency was associated with individual differences in anxiety, suggesting that vigilance or anticipation of self-relevant negative stimuli is present in depression-vulnerable individuals, but that this relates to level of anxiety rather than depression. It is possible, then, that recovered depressed individuals who make less use of CT skills are more vulnerable to anxiety-like reactions, despite the fact that the primary focus of the CT skills they learned was depression.

However, it is also possible that early pupil dilation in CT skill non-users represents extended processing of the previous negative statement. This interpretation was bolstered by an exploratory analysis finding that the correlation between CT skill use and pupil dilation remains significant even when baseline-correcting pupil dilation to a baseline from the *previous* trial. This interpretation would be in line with a large body of research associating depression vulnerability to sustained processing of self-relevant negative information as opposed to early attention towards stimuli (for review, see Mathews & MacLeod, 2005).

Results suggest several important future lines of research, in addition to those previously mentioned. First, future CT studies may benefit from including the current laboratory task in several assessments during treatment. Such a study could fully test the theory that CT invokes a transient increase in controlled processing that then becomes automatic. In addition, it will be useful to assess whether purported CT skills are in fact specific to CT. It is possible that recovery from depression by any means would involve an increase in CT-consistent emotion regulatory strategies such as reappraising negative thoughts. This can be tested by adding the current task to studies comparing recovery with CT and in wait-list controls, or comparing CT to antidepressant medication. Finally, perhaps the most important future direction is to prospectively test whether use, usefulness, and/or effort involved in CT skills can predict future relapse. If so, then promoting these elements may be important, and if not, it may be the case that those skills that aid recovery depression are different from those skills that may aid maintenance of recovery.

Several limitations temper the conclusions of this study. The study was limited by low statistical power. Many effects were near statistical significance, such that the addition of further participants may change conclusions toward significance or non-significance. Given the observed moderate effect sizes, however, we find it likely that many of these effects are truly non-null.

Another limitation was that SoCT-P related differences in pupil dilation during the Naturalistic Block were not replicated in the Reappraisal Block. Consequently, it is unclear whether such differences can be ascribed to differential reappraisal. The largely uninstructed nature of the Naturalistic Block, with few ways to probe into the cognitive processes of participants, imposes difficulty on any such interpretation.

Moreover, although the current evidence suggests that CT skill users have more adaptive emotional outcomes and lower effort and emotional engagement in response to negative thoughts, it cannot be concluded that these differences are in virtue of the skill use. CT skill use may be associated with any number of potentially confounding individual difference variables, e.g., life stress, socioeconomic status, intelligence, etc.

Finally, only 22% of the potential subject pool went on to complete the study, limiting the generalizability of effects. The current sample may represent only the most highly functioning CT responders, and relationships of variables examined here may be different in a more complete sample.

Strengths of the study include its use of laboratory measures of CT skill use and effectiveness, as well as the development of a laboratory induction of self-relevant negative thoughts. In addition, the study implemented both trait and task-related measures of CT skills so as to increase confidence that the validated trait measure was relevant to the experimental situation. Also, rather than following up patients of community therapists, the study made use of a subject pool that had mostly received high-fidelity research-protocol CT.

Perhaps most importantly, the use of pupil dilation in the current study allowed us to discriminate among competing hypotheses concerning the level of effort and emotional engagement involved in CT skill users. It appears that, although acquiring CT skills may not be easy, implementing them after years of practice may become second nature.

APPENDIX A

SUPPLEMENTARY MATERIAL

A.1 INSTRUCTIONS TO GENERATE MEMORIES FOR SAD MOOD INDUCTIONS

In the coming experiment we are interested in how mood affects responses on tasks. To do this research, we will ask you to allow yourself to experience a temporary sad mood.

To help with this, we ask that in the space below, you describe one of the worst times in your life, when you felt very sad or depressed. **We are specifically looking for a situation in which you felt really bad about yourself, for example, because you could have been at least partially at fault, or in which you felt guilty, inadequate, worthless, hopeless, or like a bad person.** Please do not choose an experience in which you felt guilty for someone's death or attempted suicide. On a scale of 1 (neutral) to 9 (extremely sad), we ask that you try to pick an experience that you would rate as at least a 7.

We will ask you to read this during the experiment and to try to feel as you felt at that time, so please try to pick an experience that is vivid enough that you will be able to re-create your mood. Please describe it in a way that will allow you to recreate this mood during the study. Describe the event clearly and in detail. Please describe your thoughts and feelings at the time of the event.

A.2 MUSICAL SELECTIONS FOR SAD MOOD INDUCTION

"Adagio" (Albinoni)

"Drive Home (from Field of Dreams Soundtrack)" (Horner)

"Russia Under the Mongolian Sky" (Prokofiev)

"Adagio" (Barber)

A.3 MUSICAL SELECTIONS FOR HAPPY MOOD INDUCTION

"Baroque and Blue" (Bolling & Rampal)

"This Old Train" (Cushnie)

"Nuns for Nixon" (Fleck)

"Agolo" (Kidjo)

"Kumnandi" (Lasser)

"Asturian Way" (Flook)

"Plekete" (Zap Mama)

"Signe" (Clapton)

A.4 INSTRUCTIONS TO SELECT SELF-RELEVANT DYSFUNCTIONAL THOUGHTS

Listed below are a variety of thoughts that pop into people's heads. Please read each thought and indicate, **when you are in a sad mood, how sad would that thought would make you feel?** Think about recent sad moods, not necessarily when you were depressed. On the right hand side of the page, write the appropriate answers in the following fashion: 1 = "no more sad", 2 = "a little more sad", 3 = "moderately more sad", 4 = "quite a bit more sad", and 5 = "extremely more sad"

A.5 INSTRUCTIONS TO GENERATE MEMORY FOR HAPPY MOOD INDUCTION

In the coming experiment we are interested in how mood affects responses to different tasks. To do this research we will ask you to allow yourself to experience a temporary happy mood.

To help with this, we ask that in the space below, you describe an extremely happy personal experience. That is, we are asking you to describe one of the best times in your life, when you felt happy or exuberant. On a scale of 1 (neutral) to 9 (extremely happy) we ask that you try to pick an experience you would rate as at least a 7.

During the study we will ask you to read this script and try to feel as you felt at that time, so please try to pick an experience that is vivid enough that you will be able to re-create your mood. Please describe it in a way that will allow you to re-create this mood during the study. Please describe the event clearly and in detail. Please describe your thoughts and feelings at the time of the event.

A.6 INSTRUCTIONS FOR LABORATORY TASK, NATURALISTIC BLOCK

When you see each sentence, think about the sentence, as if it were a thought that popped into your head. So, in your mind, deal with the thought in whatever way YOU NORMALLY WOULD if it popped into your head.

Now, whenever you find yourself getting off-task or distracted, press the space bar to continue to the next sentence.

Remember, deal with each thought in whatever way YOU NORMALLY WOULD if it popped into your head. There is no right or wrong way.

Sentences will be repeated. And when this happens, please deal with them again as new thoughts.

The task will end after 5 minutes.

A.7 INSTRUCTIONS FOR LABORATORY TASK, REAPPRAISAL BLOCK

When you see each sentence, think of a REFRAME to each sentence you see, using the skills you just learned. Please do the reframe in your head, do NOT type it. It is important you think of a reframe for every sentence.

Once you have thought of a reframe, press space bar to continue to the next sentence.

Sentences will be repeated. When this happens, make A NEW, DIFFERENT REFRAME. To remind you of these instructions, you will see the word REFRAME before each sentence.

The task will end after 5 minutes.

Again, sentences will be repeated. Please do ANOTHER DIFFERENT REFRAME when this happens.

BIBLIOGRAPHY

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review*, *30*(2), 217-237.

Baer, R. A. (2003). Mindfulness training as a clinical intervention: A conceptual and empirical review. *Clinical psychology: Science and practice*, *10*(2), 125-143.

Barber, J. P., & DeRubeis, R. J. (1989). On second thought: Where the action is in cognitive therapy for depression. *Cognitive Therapy and Research*, *13*(5), 441-457.

Beck, A. T. (1979). *Cognitive therapy of depression*. The Guilford Press.

Beck, A. T., Brown, G., Steer, R. A., Eidelson, J. I., & Riskind, J. H. (1987). Differentiating anxiety and depression utilizing the Cognition Checklist. *Journal of Abnormal Psychology*, *96*, 179-183.

Bondolfi, G., Jermann, F., Der Linden, M. V., Gex-Fabry, M., Bizzini, L., Rouget, B. W., Myers-Arrazola, L., et al. (2010). Depression relapse prophylaxis with Mindfulness-Based Cognitive Therapy: Replication and extension in the Swiss health care system. *Journal of affective disorders*, *122*(3), 224-231.

Bradley, B. P., Mogg, K., & Lee, S. C. (1997). Attentional biases for negative information in induced and naturally occurring dysphoria. *Behaviour Research and Therapy*, *35*(10), 911-927.

Clark, D. M., & Teasdale, J. D. (1985). Constraints on the effects of mood on memory. *Journal of Personality and Social Psychology*, *48*(6), 1595.

Crandell, C. J., & Chambless, D. L. (1986). The validation of an inventory for measuring depressive thoughts: The Crandell Cognitions Inventory. *Behaviour research and therapy*, *24*(4), 403-411.

DeRubeis, R. J., Evans, M. D., Hollon, S. D., Garvey, M. J., Grove, W. M., & Tuason, V. B. (1990). How does cognitive therapy work? Cognitive change and symptom change in cognitive therapy and pharmacotherapy for depression. *Journal of Consulting and Clinical Psychology*, *58*(6), 862.

DeRubeis, R. J., Siegle, G. J., & Hollon, S. D. (2008). Cognitive therapy vs. medications for depression: Treatment outcomes and neural mechanisms. *Nature reviews. Neuroscience*, *9*(10), 788.

Eich, E., Ng, J. T. W., Macaulay, D., Percy, A. D., & Grebneva, I. (2007). Combining music with thought to change mood. *The handbook of emotion elicitation and assessment*, 124-136.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). Structured clinical interview for DSM-IV Axis I disorders. *Structured Clinical interview for DSM-IV Axis I Disorders*.

Fox, J. (1991). *Regression diagnostics* (Vol. 79). Sage Publications, Inc.

Garratt, G., Ingram, R. E., Rand, K. L., & Sawalani, G. (2007). Cognitive processes in cognitive therapy: Evaluation of the mechanisms of change in the treatment of depression. *Clinical Psychology: Science and Practice*, *14*(3), 224-239.

Gemar, M. C., Kapur, S., Segal, Z. V., Brown, G. M., & Houle, S. (1996). Effects of self generated sad mood on regional cerebral activity: A PET study in normal subjects. *Depression*, *4*(2), 81-88.

Gemar, M. C., Segal, Z. V., Sagrati, S., & Kennedy, S. J. (2001). Mood-induced changes on the Implicit Association Test in recovered depressed patients. *Journal of Abnormal Psychology*, *110*(2), 282.

Gibbons, M. B. C., Crits-Christoph, P., Barber, J. P., Wiltsey Stirman, S., Gallop, R., Goldstein, L. A., Temes, C. M., et al. (2009). Unique and common mechanisms of change across cognitive and dynamic psychotherapies. *Journal of consulting and clinical psychology*, *77*(5), 801.

Gilboa, E., & Gotlib, I. H. (1997). Cognitive biases and affect persistence in previously dysphoric and never-dysphoric individuals. *The cognitive psychology of depression*, *5*(6), 517-538.

Glasman, D., Finlay, W. M. L., & Brock, D. (2004). Becoming a self therapist: Using cognitivebehavioural therapy for recurrent depression and/or dysthymia after completing therapy. *Psychology and Psychotherapy: Theory, Research and Practice*, *77*(3), 335-351.

Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: Current status and future directions. *Annual review of clinical psychology*, *6*, 285.

Greenberger, D., & Padesky, C. A. (1995). *Mind over mood: Change how you feel by changing the way you think*. The Guilford Press.

Gross, J. J. (1998). Antecedent-and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of personality and social psychology*, *74*(1), 224.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of personality and social psychology*, *85*(2), 348.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*(2), 240-244.

Harrell, T. H., & Ryon, N. B. (1983). Cognitive-behavioral assessment of depression: Clinical validation of the Automatic Thoughts Questionnaire. *Journal of Consulting and Clinical Psychology*, *51*(5), 721.

Hoaglin, D. C., & Welsch, R. E. (1978). The hat matrix in regression and ANOVA. *American Statistician*, 17-22.

Hofmann, S. G., Heering, S., Sawyer, A. T., & Asnaani, A. (2009). How to handle anxiety: The effects of reappraisal, acceptance, and suppression strategies on anxious arousal. *Behaviour research and therapy*, *47*(5), 389-394.

Hollon, S. D. (2003). Does cognitive therapy have an enduring effect? *Cognitive Therapy and Research*, *27*(1), 71-75.

Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research*, *4*(4), 383-395.

Hollon, S. D., Kendall, P. C., & Lumry, A. (1986). Specificity of depressotypic cognitions in clinical depression. *Journal of Abnormal Psychology*, *95*(1), 52.

Hollon, S. D., Stewart, M. O., & Strunk, D. R. (2006). Enduring effects for cognitive behavior therapy in the treatment of depression and anxiety. *Annu. Rev. Psychol.*, *57*, 285-315.

Jarrett, R. B., & Kraft, D. (1997). Prophylactic cognitive therapy for major depressive disorder. *In Session: Psychotherapy in Practice*, *3*(3), 65-79.

Jarrett, R. B., & Thase, M. E. (2010). Comparative efficacy and durability of continuation phase cognitive therapy for preventing recurrent depression: Design of a double-blinded, fluoxetine- and pill placebo-controlled, randomized trial with 2-year follow-up. *Contemporary Clinical Trials*, *31*(4), 355-377. doi:10.1016/j.cct.2010.04.004

Jarrett, R. B., Vittengl, J. R., Clark, L. A., & Thase, M. E. (2011). Skills of Cognitive Therapy (SoCT): A new measure of patients' comprehension and use.

Johnstone, T., van Reekum, C. M., Urry, H. L., Kalin, N. H., & Davidson, R. J. (2007). Failure to regulate: counterproductive recruitment of top-down prefrontal-subcortical circuitry in major depression. *The Journal of Neuroscience*, *27*(33), 8877.

Kaufman, J., & Charney, D. (2000). Comorbidity of mood and anxiety disorders. *Depression and Anxiety*, *12*(S1), 69-76.

Keller, M. B., Lavori, P. W., Friedman, B., Nielsen, E., Endicott, J., McDonald-Scott, P., & Andreasen, N. C. (1987). The Longitudinal Interval Follow-up Evaluation. A comprehensive method for assessing outcome in prospective longitudinal studies. *Archives of General Psychiatry, 44*(6), 540.

Koster, E. H. W., Verschuere, B., Crombez, G., & Van Damme, S. (2005). Time-course of attention for threatening pictures in high and low trait anxiety. *Behaviour Research and Therapy, 43*(8), 1087-1098.

Kreutz, G., Ott, U., Teichmann, D., Osawa, P., & Vaitl, D. (2008). Using music to induce emotions: Influences of musical preference and absorption. *Psychology of music, 36*(1), 101.

Mathews, A., & MacLeod, C. (2005). Cognitive vulnerability to emotional disorders. *Annu. Rev. Clin. Psychol.*, *1*, 167-195.

Mauss, I. B., Cook, C. L., & Gross, J. J. (2007). Automatic emotion regulation during anger provocation. *Journal of Experimental Social Psychology*, *43*(5), 698-711.

Mauss, I. B., Evers, C., Wilhelm, F. H., & Gross, J. J. (2006). How to bite your tongue without blowing your top: Implicit evaluation of emotion regulation predicts affective responding to anger provocation. *Personality and Social Psychology Bulletin*, *32*(5), 589.

Miller, J. J., Fletcher, K., & Kabat-Zinn, J. (1995). Three-year follow-up and clinical implications of a mindfulness meditation-based stress reduction intervention in the treatment of anxiety disorders* 1. *General Hospital Psychiatry*, *17*(3), 192-200.

Miller, P. R., Dasher, R., Collins, R., Griffiths, P., & Brown, F. (2001). Inpatient diagnostic assessments: 1. Accuracy of structured vs. unstructured interviews* 1. *Psychiatry Research*, *105*(3), 255-264.

Moore, S. A., Zoellner, L. A., & Mollenholt, N. (2008). Are expressive suppression and cognitive reappraisal associated with stress-related symptoms? *Behaviour research and therapy*, *46*(9), 993-1000.

Nolen-Hoeksema, S., & Morrow, J. (1993). Effects of rumination and distraction on naturally occurring depressed mood. *Cognition & Emotion*, *7*(6), 561-570.

Oliveira, I. R. (2008). Trial-Based Thought Record (TBTR): preliminary data on a strategy to deal with core beliefs by combining sentence reversion and the use of analogy with a judicial process. *Revista Brasileira de Psiquiatria*, *30*(1), 12-18.

Persons, J. B. (1993). The process of change in cognitive therapy: Schema change or acquisition of compensatory skills? *Cognitive therapy and research*, *17*(2), 123-137.

Powers, D. V., Thompson, L. W., & Gallagher-Thompson, D. (2008). The Benefits of Using Psychotherapy Skills Following Treatment for Depression: An Examination of. *Cognitive and Behavioral Practice*, *15*(2), 194-202.

van Reekum, C. M., Johnstone, T., Urry, H. L., Thurow, M. E., Schaefer, H. S., Alexander, A. L., & Davidson, R. J. (2007). Gaze fixations predict brain activation during the voluntary regulation of picture-induced negative affect. *NeuroImage*, *36*(3), 1041-1055. doi:10.1016/j.neuroimage.2007.03.052

Reibel, D. K., Greeson, J. M., Brainard, G. C., & Rosenzweig, S. (2001). Mindfulness-based stress reduction and health-related quality of life in a heterogeneous patient population* 1. *General Hospital Psychiatry*, *23*(4), 183-192.

Reynolds, M., & Brewin, C. R. (1999). Intrusive memories in depression and posttraumatic stress disorder. *Behaviour Research and Therapy*.

Rush, A. J., Bernstein, I. H., Trivedi, M. H., Carmody, T. J., Wisniewski, S., Mundt, J. C., Shores-Wilson, K., et al. (2006). An evaluation of the quick inventory of depressive symptomatology and the hamilton rating scale for depression: a sequenced treatment alternatives to relieve depression trial report. *Biological psychiatry*, *59*(6), 493-501.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime*. Pittsburgh, PA: Psychology Software Tools.

Segal, Z. V., & Ingram, R. E. (1994). Mood priming and construct activation in tests of cognitive vulnerability to unipolar depression. *Clinical Psychology Review, 14*(7), 663-695.

Segal, Z. V., Gemar, M. C., & Williams, S. (1999). Differential cognitive response to a mood challenge following successful cognitive therapy or pharmacotherapy for unipolar depression. *Journal of Abnormal Psychology*, *108*(1), 3.

Segal, Z. V., Kennedy, S., Gemar, M. C., Hood, K., Pedersen, R., & Buis, T. (2006). Cognitive reactivity to sad mood provocation and the prediction of depressive relapse. *Archives of General Psychiatry*.

Siegle, G. J., Granholm, E., Ingram, R. E., & Matt, G. E. (2001). Pupillary and reaction time measures of sustained processing of negative information in depression. *Biological psychiatry*, *49*(7), 624-636.

Siegle, G. J., Ichikawa, N., & Steinhauer, S. R. (2008). Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology*, *45*(5), 679-687.

Siegle, G. J., Moore, P. M., & Thase, M. E. (2004). Rumination: One construct, many features in healthy individuals, depressed individuals, and individuals with lupus. *Cognitive Therapy and Research*, *28*(5), 645-668.

Siegle, G. J., Steinhauer, S. R., & Thase, M. E. (2004). Pupillary assessment and computational modeling of the Stroop task in depression. *International Journal of Psychophysiology*, *52*(1), 63-76.

Siegle, G. J., Steinhauer, S. R., Carter, C. S., Ramel, W., & Thase, M. E. (2003). Do the seconds turn into hours? Relationships between sustained pupil dilation in response to emotional information and self-reported rumination. *Cognitive Therapy and Research*, *27*(3), 365-382.

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., & Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *Neuroimage*, *20*(1), 114-124.

Siegle, G. J., Steinhauer, S. R., Thase, M. E., Stenger, V. A., & Carter, C. S. (2002). Can't shake that feeling: event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biological Psychiatry*, *51*(9), 693-707.

Simons, A. D., Garfield, S. L., & Murphy, G. E. (1984). The process of change in cognitive therapy and pharmacotherapy for depression. Changes in mood and cognition. *Archives of General Psychiatry*, *41*(1), 45.

Steer, R. A., Beck, A. T., Clark, D. A., & Beck, J. S. (1994). Psychometric properties of the Cognition Checklist with psychiatric outpatients and university students. *Psychological Assessment*, *6*(1), 67.

Steidtmann, D., Ingram, R. E., & Siegle, G. J. (2010). Pupil response to negative emotional information in individuals at risk for depression. *Cognition & Emotion*, *24*(3), 480-496.

Steinhauer, S. R., & Hakerem, G. (1992). The Pupillary Response in Cognitive Psychophysiology and Schizophreniaa. *Annals of the New York Academy of Sciences*, *658*(1), 182-204.

Steinhauer, S. R., Boller, F., Zubin, J., & Pearlman, S. (1983). Pupillary dilation to emotional visual stimuli revisited. *Psychophysiology*, *20*, 472.

Strunk, D. R., Brotman, M. A., & DeRubeis, R. J. (2010). The process of change in cognitive therapy for depression: Predictors of early inter-session symptom gains. *Behaviour Research and Therapy*, *48*(7), 599-606. doi:10.1016/j.brat.2010.03.011

Strunk, D. R., DeRubeis, R. J., Chiu, A. W., & Alvarez, J. (2007). Patients' competence in and performance of cognitive therapy skills: Relation to the reduction of relapse risk following treatment for depression. *Journal of Consulting and Clinical Psychology*, *75*(4), 523.

Teasdale, J. D., & Fennell, M. J. V. (1982). Immediate effects on depression of cognitive therapy interventions. *Cognitive Therapy and Research*, *6*(3), 343-351.

Teasdale, J. D., Segal, Z. V., Williams, J. M. G., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, *68*(4), 615.

Urry, H. L., van Reekum, C. M., Johnstone, T., Kalin, N. H., Thurow, M. E., Schaefer, H. S., Jackson, C. A., et al. (2006). Amygdala and ventromedial prefrontal cortex are inversely coupled during regulation of negative affect and predict the diurnal pattern of cortisol secretion among older adults. *The Journal of Neuroscience*, *26*(16), 4415.

Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive-behavioral therapy's effects. *Journal of consulting and clinical psychology*, *75*(3), 475.

Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive-behavioral therapy's effects. *Journal of consulting and clinical psychology*, *75*(3), 475.

Warshaw, M. G., Dyck, I., Allsworth, J., Stout, R. L., & Keller, M. B. (2001). Maintaining reliability in a long-term psychiatric study: an ongoing inter-rater reliability monitoring program using the longitudinal interval follow-up evaluation. *Journal of Psychiatric Research*, *35*(5), 297-305. doi:10.1016/S0022-3956(01)00030-9

Watson, D., & Clark, L. A. (1999). The PANAS-X: Manual for the positive and negative affect schedule-expanded form.

Weissman, A. N., & Beck, A. T. (1978). Development and Validation of the Dysfunctional Attitude Scale: A Preliminary Investigation.

Williams, L. E., Bargh, J. A., Nocera, C. C., & Gray, J. R. (2009). The unconscious regulation of emotion: Nonconscious reappraisal goals modulate emotional reactivity. *Emotion*, *9*(6), 847.

Wright, J. H., Wright, A. S., Albano, A. M., Basco, M. R., Goldsmith, L. J., Raffield, T., & Otto, M. W. (2005). Computer-assisted cognitive therapy for depression: maintaining efficacy while reducing therapist time. *American Journal of Psychiatry*, *162*(6), 1158.

Young, J., & Beck, A. T. (1980). Cognitive therapy scale: Rating manual. *Unpublished manuscript*.

Zanarini, M. C., Skodol, A. E., Bender, D., Dolan, R., Sanislow, C., Schaefer, E., Morey, L. C., et al. (2000). The collaborative longitudinal personality disorders study: Reliability of axis I and II diagnoses. *Journal of Personality Disorders*, *14*(4), 291-299.