Affect labeling enhances exposure effectiveness for public speaking anxiety

Andrea N. Niles[*], Michelle G. Craske, Matthew D. Lieberman, Christopher Hur

University of California, Los Angeles, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563, USA

ARTICLE INFO

ABSTRACT

Exposure is an effective treatment for anxiety but many patients do not respond fully. Affect labeling (labeling emotional experience) attenuates emotional responding. The current project examined whether affect labeling enhances exposure effectiveness in participants with public speaking anxiety. Participants were randomized to exposure with or without affect labeling. Physiological arousal and self-reported fear were assessed before and after exposure and compared between groups. Consistent with hypotheses, participants assigned to Affect Labeling, especially those who used more labels during exposure, showed greater reduction in physiological activation than Control participants. No effect was found for self-report measures. Also, greater emotion regulation deficits at baseline predicted more benefit in physiological arousal from exposure combined with affect labeling than exposure alone. The current research provides evidence that behavioral strategies that target prefrontal-amygdala circuitry can improve treatment effectiveness for anxiety and these effects are particularly pronounced for patients with the greatest deficits in emotion regulation.

© 2015 Elsevier Ltd. All rights reserved.

Although behavioral treatments for anxiety disorders are highly effective in reducing symptoms of anxiety (Butler, Chapman, Forman, & Beck, 2006; Hofmann & Smits, 2008; Tolin, 2010), many patients do not improve, drop out of treatment, or relapse (Arch & Craske, 2009; Clark et al., 2006; Davidson et al., 2004). Given the need to improve treatments, the goal of the current project is to translate neuroscience research to enhancing the effectiveness of exposure therapy for public speaking anxiety. Specifically, this project compared the effectiveness of exposure alone to exposure plus affect labeling (i.e. putting feelings into words) for individuals with public speaking anxiety.

Public speaking anxiety, a form of social phobia, is one of the most common psychological problems in the United States with prevalence estimates ranging from 11% to 30% of the population (Pollard & Henderson, 1988; Stein, Walker, & Forde, 1996; Wittchen, Stein, & Kessler, 1999). Current treatments for public speaking anxiety combine traditional exposure (e.g., repeated trials of public speaking) with cognitive restructuring in which patients are taught to think about the feared situation neutrally or positively rather than negatively (Heimberg, 2002; Hofmann & Smits, 2008; Hope, Heimberg, Juster, & Turk, 2000; Rapee & Heimberg, 1997). Such treatments aim to reduce anticipatory anxiety, anxiety during speaking, and rumination about the speech after it is over (Clark & Wells, 1995). Although exposure alone appears to be an effective treatment for social anxiety disorder (Feske & Chambless, 1995), to our knowledge, no researchers have used laboratory studies to assess whether adding verbalization (such as cognitive restructuring) to exposure enhances its effects on fear reduction.

Neuroscience research can inform our understanding of anxiety and exposure therapy, and studies on fear learning and anxiety pinpoint the amygdala as central to fear acquisition and responding (Davis, 1992). Activation of prefrontal regions and the strength of connectivity between the prefrontal cortex (PFC) and the amygdala, are essential to successful fear extinction. For example, electrical stimulation of the medial PFC led to reductions of conditioned fear responding in rats (Milad & Quirk, 2002). Greater ventromedial PFC activity is associated with better extinction of conditioned fear in humans (Delgado, Nearing, LeDoux, & Phelps, 2008; Milad et al., 2005; Phelps, Delgado, Nearing, & LeDoux, 2004). Assuming that extinction is a central mechanism of exposure therapy (Craske et al., 2008; Craske, Liao, Brown, & Vervliet, 2012), PFC downregulation of amygdala may contribute to successful exposure therapy, and strategies that augment such downregulatory pathways may augment outcomes from exposure therapy. In addition,

* Corresponding author.
E-mail address: aniles@ucla.edu (A.N. Niles).

evidence suggests that patients with social anxiety disorder have weaker connectivity between the medial orbitofrontal PFC and the amygdala compared to healthy controls (Hahn et al., 2011). Therefore, treatments that strengthen connectivity between prefrontal regions and the amygdala may prove particularly beneficial in the treatment of social anxiety.

Disruption theory of language and emotion (Lieberman, 2003, 2011) posits that labeling one's emotional state can disrupt the experience of that emotional state. However, because intent to reduce distress is not explicit, affect labeling has been conceptualized as an incidental emotion regulation strategy (Burklund, Creswell, Irwin, & Lieberman, 2014), which differs from intentional strategies such as cognitive restructuring or emotional suppression. A number of neuroimaging studies have demonstrated that labeling one's emotional experience activates areas of the PFC, and reduces activation in the amygdala (Gorno-Tempini et al., 2001; Hariri, Bookheimer, & Mazziotta, 2000; Hariri, Mattay, Tessitore, Fera, & Weinberger, 2003; Narumoto et al., 2000). The right ventrolateral PFC is consistently activated during affect labeling (Cunningham, Johnson, Chris, Gore, & Banaji, 2003; Lieberman et al., 2007; Narumoto et al., 2000), and it is presumed that this region downregulates amygdala activation. The principle of neural plasticity states that repetition of a process can increase efficiency and efficacy of that process through changes in neuron function, chemical profile, and structure (Anderson, 2010; Kandel & Schwartz, 1982). Therefore, repeated affect labeling may enhance connectivity in PFC-amygdala pathways in turn improving patients' ability to regulate emotional responses. Additionally, there may be a dose response relationship between the quantity of affect labeling trials and the degree of enhanced connectivity.

In accord with this notion, two studies have demonstrated that affect labeling enhances the effectiveness of exposure. Tabibnia, Lieberman, and Craske (2008) examined the effect of repeated exposure to evocative images with and without negative affective labels. In study 1 with healthy controls, repeated presentation of emotionally evocative images paired with an affect label resulted in greater attenuation of skin conductance responding and heart rate deceleration upon re-presentation of the images without a label at one-week re-test. In study 2, the findings were replicated for skin conductance response in spider-fearful subjects who were exposed to spider images paired with negative labels compared to no labels or neutral labels. Kircanski, Lieberman, and Craske (2012) compared the effects of exposure to a live spider without linguistic processing, with affect labeling, with reappraisal, and with distraction, in spider fearful subjects. At one-week re-test the group that completed exposure with affect-labeling had lower skin conductance responses while viewing a spider and moved closer to the spider compared to the reappraisal and exposure alone groups. In addition, those who used the greatest number of anxiety- and fear-related words during affect labeling showed the greatest reductions in skin conductance responding and moved closest to the spider.

Another consideration we examine in the current study is whether the matching of treatments to individuals may improve therapy outcomes. Two possibilities have been evaluated: individuals with a deficit are more likely to benefit from treatments that target that deficit (compensation), and individuals with a strength will benefit most from a treatment that matches that strength (capitalization; Rude & Rehm, 1991). Recent studies on depression and suicidality treatment have found support for both capitalization (Cheavens, Strunk, Lazarus, & Goldstein, 2012) and compensation (Wingate, Van Orden, JoinerJr., Williams, & David, 2005). Studies examining amygdala activation (McClure et al., 2007), emotional reactivity to evocative images (Niles, Mesri, Burklund, Lieberman, & Craske, 2013), and heart rate variability (Davies, Niles, Pittig, Arch, & Craske, 2015) as predictors of treatment outcome for anxiety patients, support a compensation model, with superior outcomes for patients with greater reactivity at baseline. We aimed to evaluate whether affect labeling would most benefit those with a deficit or with a strength in affect labeling at baseline. The extent to which affect labeling at baseline reduces distress serves as an indicator of incidental emotion regulation capacity, and can be used to determine whether participants with strengths (capitalization) or with deficits (compensation) in emotion regulation benefit more from an intervention augmented with implicit emotion regulation training (i.e., affect labeling).

The current study had three aims. The first aim was to assess whether affect labeling enhanced the effectiveness of exposure compared to exposure alone. We hypothesized that participants instructed to use affect labeling during exposure would show greater attenuation of fear in anticipation of and recovery from public speaking compared to those who completed exposure alone. The second aim was to assess whether the number of anxiety- or fear-related words used during affect labeling predicted greater attenuation of fear responding at re-test. We hypothesized that participants who used more anxiety or fear related words compared to other negative emotion words would show the greatest fear reduction at re-test. The third aim was to assess whether individual differences in incidental emotion regulation (i.e., the extent to which affect labeling reduced distress in a pretesting session) moderated response to exposure with affect labeling versus exposure alone. Given mixed findings in the literature, these analyses were mainly exploratory, and we made no a priori predictions.

1. Method

1.1. Design

This study used a 2 (Group) × 3 (Time) mixed design with public speaking fearful participants. Groups included exposure combined with affect labeling (AL), and exposure alone (Control). Time included assessment time-points at baseline (Time 1), following exposure (Time 2) and at 1-week follow-up (Time 3).

1.2. Participants

One hundred two participants (AL = 52; Control = 50) were recruited to participate. Two participants assigned to the Control group were not included in analyses: one participant received the incorrect study protocol due to experimenter error and another fell asleep during the experiment. Therefore, the final sample included in analyses was 100. See Fig. 1 for a consort diagram of flow through study procedures. Participants had a mean age of 25 (SD = 9.1), 80% were female, 92% were students, and 37% spoke English as a second language. The ethnic breakdown of the sample was 55% Asian, 16% Hispanic, 14% Caucasian, 6% African American, and 9% other.

Eligible participants reported a 6 or higher on anxiety and a 5 or higher on avoidance of public speaking on a 0 to 8 scale. The prompts for anxiety and avoidance respectively were "How anxious would you feel giving a formal speech before a live audience?" and "How likely would you be to avoid taking a class that required an oral presentation?" Zero indicated no anxiety/never avoid, and 8 indicated extreme anxiety/always avoid. This two question survey has been used to recruit public speaking fearful participants in previous studies (Culver, Stoyanova, & Craske, 2012; Tsao & Craske, 2000). Participants were over 18 years of age, fluent in English, free of heart, neurological, or respiratory conditions, hearing impairment, physician recommendation to avoid stressful situations, current treatment for public speaking anxiety, or psychotropic medication prescription for an emotional problem. Participants
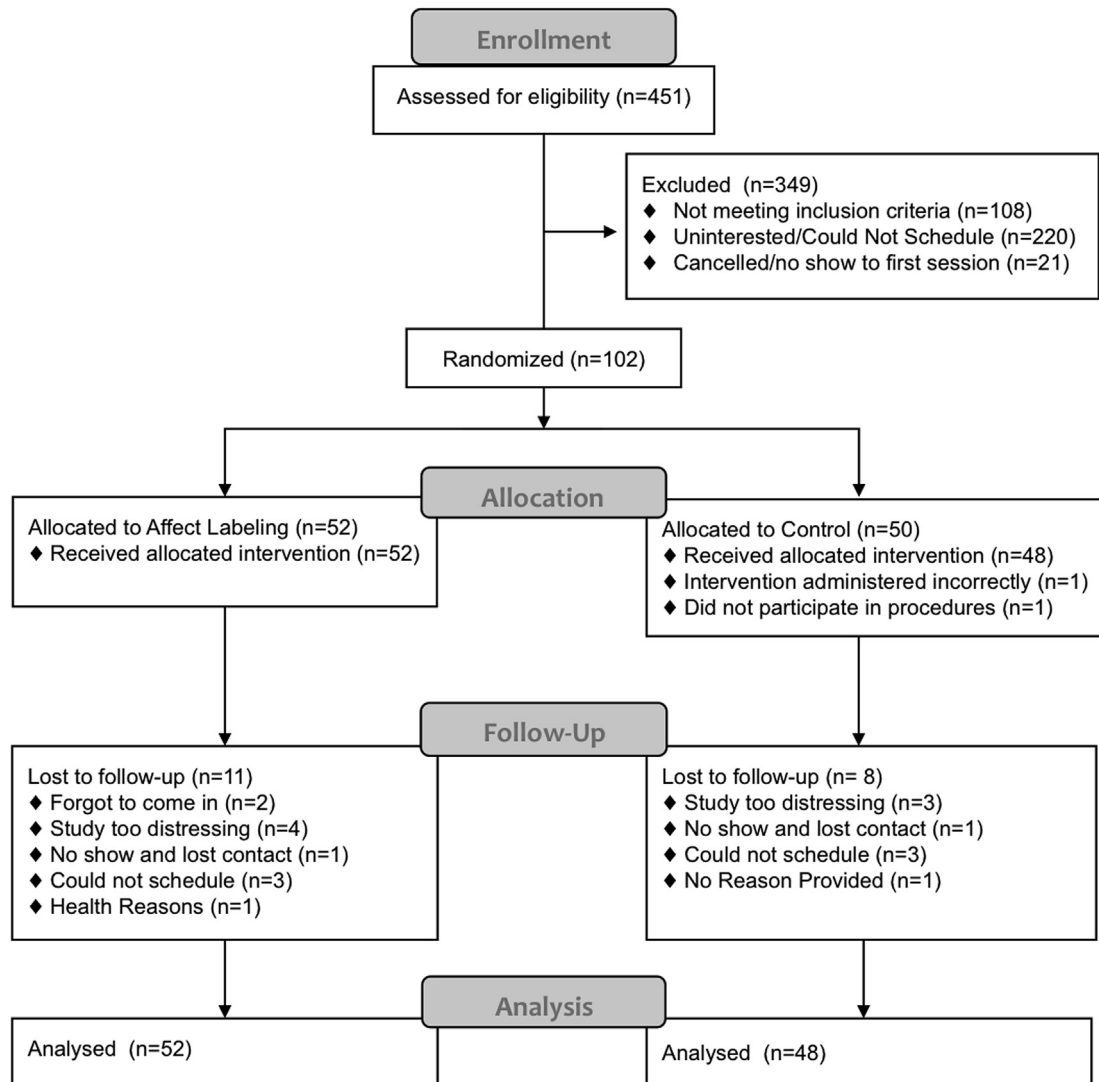
Fig. 1. Participant flow chart.

were recruited from the UCLA Psychology Subject Pool and flyers posted around UCLA campus. Participants were given 1 h of research credit per day or were paid $10.00 per day.

1.3. Materials

1.3.1. Physiological activity

Physiological activity was recorded using a Biopac system, an IBM Pentium II, and AcqKnowledge software (AcqKnowledge 4.1 for Windows; BIOPAC Systems, inc). Non-specific skin conductance responses (SCR-NS) and heart rate (HR) were recorded as measures of fear arousal. All physiological data were first visually inspected to ensure proper measurement. For HR, one participant at Time 1, and two participants at Time 2 were excluded from analysis due to recording error. For SCR-NS, nine participants at Time 1, 10 participants at Time 2, and 6 participants at Time 3 were excluded from analyses because no variations in signal were observed.[1]

To assess HR, electrocardiogram signals were collected from electrodes on the right clavicle and below the bottom left front rib. HR was defined as the number of heart beats per minute. A band pass filter with a low cutoff of 1.00 Hz and a high cutoff of 35.00 Hz was applied prior to analysis to limit the effect of signal noise on the data. SCR-NS was recorded from electrodes attached to the medial phalanges of the second and third fingers of a participant's non-preferred hand. SCR-NS was assessed by calculating the frequency of non-specific skin conductance responses per minute. A skin conductance response was defined by a minimum increase of .02 μs. Data were analyzed using built in analysis tools in AcqKnowledge software for HR and SCR-NS.

1.3.2. Self report

1.3.2.1. Personal report of public speaking anxiety (McCroskey, 1970).
The PRPSA is a 34-item measure that assesses fear of public speaking. Responders rate their degree of agreement with each statement on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). Sample items include "While preparing for giving a speech, I feel tense and nervous," and "My thoughts become confused and jumbled when I am giving a speech." The scale has excellent reliability (α = .90; McCroskey, 1970). In the

---

[1] These numbers are consistent with estimates that approximately 10% (or more in clinical samples) of individuals do not show a reliable GSR response (Braithwaite, Watson, Jones, & Rowe, 2013).

current study, $\alpha = .97$. Scores of 131 or higher indicate high public speaking anxiety, scores of 98–131 indicate moderate anxiety and scores below 98 indicate low public speaking anxiety.

1.3.2.2. Subject units of distress scale (SUDS). SUDS is a single item measure used to assess state anxiety. Participants were shown a zero to eight point Likert scale with zero indicating no anxiety and eight indicating extreme anxiety to rate their SUDS directly prior to and following each speech task. Participants were asked to rate their current anxiety level. SUDS data from before and after the speech are averaged.

1.3.3. Word use during exposure

During exposures, participants assigned to AL chose from a set of emotion words displayed on the computer screen using the keyboard. Their responses were recorded, and the number of anxiety-related words chosen was identified for each participant. The mean percent of anxiety-related words chosen was 36% ($SD = 25\%$; Range = 0–85%).

1.3.4. Incidental emotion regulation

Incidental emotion regulation was defined as the reduction in subjective distress following affect labeling. Incidental emotion regulation was assessed using the Affect Labeling Task (Lieberman, Inagaki, Tabibnia, & Crockett, 2011), and scores were calculated by subtracting average level of distress when labeling negative images from average level of distress when viewing negative images without labeling. Scores ranged from −1.36 to 2.01 ($M = .08$, $SD = .63$) with higher scores indicating more effective emotion regulation. For more details on the Affect Labeling Task, see procedure below.

1.4. Procedure

Participants who were eligible were scheduled for three appointment times. At Time 1, participants completed a behavioral approach test (BAT), and the first exposure session. At Time 2 (3 days following Time 1), participants completed the second exposure and the second BAT, and at Time 3 (8 days after Time 1), participants completed the third BAT. The session timing followed the procedure used by Kircanski et al. (2012)

1.4.1. BAT

After informed consent, electrodes were attached for continuous physiological measurement. Participants reported on demographic characteristics and completed questionnaires before a 1-min baseline recording. Next, participants were trained to use the SUDS and were given instructions for completing the BAT. A 1-min anticipation period was recorded as participants sat behind a screen, following which they provided a SUDS rating and stood in front of an audience of three confederates sitting in chairs. The beginning of the speech task was signaled by a computer-generated tone, and the first speech topic was displayed on a computer screen on a desk situated to the left of the participant. The participant spoke for 1 min then entered a SUDS rating using the computer keyboard before a 1-min recovery period was recorded. Only measurements taken during the 1-min baseline, anticipation, and recovery periods were analyzed. The BAT procedure was repeated at Time 2 and Time 3, but speech topics differed. Speech topics were health care, president Obama, and global warming, and the order was counterbalanced across participants using a Latin square.

1.4.2. Affect labeling task

Following the BAT, physiological equipment was disconnected and the participant was taken to another room to complete the

Affect Labeling Task as described by Lieberman et al. (2011) with slight modifications to reduce the length of the procedure (Figure S1 in the supplemental material available online). Participants viewed negative images from the International Affective Picture System (Lang, Bradley, & Cuthbert, 1999). Images were organized into eight blocks of four images each. Each block contained two moderately negative, and two extremely negative images. Prior to each block, participants were prompted by cues that said either "scene description" (labeling) or "look and let yourself respond naturally" (watching). Four blocks were labeling, and four were watching, and participants were randomly assigned to view the blocks in one of two orders. Pictures appeared for 5 s. For labeling blocks, participants were asked to choose from three labels that appeared at the bottom of the screen (e.g., attack, tornado, sitting). One label was relevant to the image, and the other two were not. Two of the labels were negative, and one was neutral. Participants chose a word by pressing a key on the keyboard that corresponded to the position of the word on the screen. For watching blocks, participants simply viewed each image for 5 s. Following the presentation of each image, regardless of block condition, participants were asked "How distressed did you feel while looking at the picture?" and responded on a 9-point Likert scale with 0 being not distressed, and 8 being very distressed.

1.4.3. Exposure

Participants were randomly assigned to the Control or AL group. Study personnel were blinded to study condition. Exposures were conducted over two days, and the same procedure was used on both days. For a diagram of the exposure procedure, see Figure S2 in the supplemental materials available online. Following the procedure timing used by Kircanski et al. (2012), participants completed 10 1-min speech trials in front of an audience; after each speech, they were prompted to step behind a screen for a 30-s inter-trial interval. Speech topics were different for each day of exposure, but were the same across all participants and were presented to participants in the same order. Examples of speech topics used are favorite movie, favorite sport, favorite food, and Los Angeles weather. The audience was comprised of three confederates (different than the BAT confederates) who were seated facing the participant.

Prior to each speech, participants in AL were prompted by the computer to choose words to label their emotions and words to label their feared outcome from four options presented on the screen (see Figure S2 in online supplemental materials). The inclusion of emotion labeling followed by feared outcome identification was consistent with the approach used by Kircanski et al. (2012). At the top of the screen, the phrase, "I feel _____" was presented followed by three possible emotion words or "other." All three emotions were negative (one anger, one sadness, and one anxiety label). Two independent raters correctly classified 100% of the affect labels into anger-related, sadness-related, and anxiety-related emotions. Examples of anxiety-related emotions include "afraid," "nervous," "anxious," and "jittery." The phrase "The audience will _____" was then presented followed by three possible feared outcomes related to the audience's response to the participant. Examples include "laugh at me," "be disinterested," "notice I'm nervous," and "think I'm stupid." An "other" option was also available. Participants were presented with nine different sets of emotion labels and feared outcome labels for the first nine speeches, and the first set was repeated for the tenth speech.[2]

---

[2] Upon investigation of labels chosen by the first five participants, one anxiety option was never chosen, and was replaced with an option that was more frequently chosen.

Prior to each speech, participants in the Control group completed a shape-matching task; they were presented with a large black shape at the top of the screen and were asked to match the shape with one of three options at the bottom of the screen (see Figure S2 in online supplemental materials). If the shape at the top did not match any of the three shapes at the bottom, the participants were asked to choose "other". Participants repeated the shape matching exercise a second time with blue shapes.

1.5. Data analysis

Analyses were conducted using Stata 12. Dependent measures assessed for all three study aims were physiology (HR, SCR-NS) and self-reported anxiety (PRPSA, SUDS) measured during the BAT at Times 1, 2, and 3. For models including physiology as the dependent variable, baseline HR squared and baseline SCR-NS from Time 1 were included as covariates to control for baseline individual differences in physiological activation. For models including SCR-NS (a count variable) as the dependent measure, Poisson regression was used to account for non-normality. All tests were two-tailed with an α level of .05.

All three study aims were tested using multi-level modeling (MLM), which accounts for within and between participant variance, and effectively handles missing data by including all participants in the model regardless of missing data points. Time was modeled at level 1, and participant level variables (e.g., Group) were modeled at level two. Time was modeled using two segments from Time 1 to Time 2 and from Time 2 to Time 3 based on a pattern of results typically observed in intervention studies characterized by an initial steep change in symptoms from pre to post intervention and a leveling out of change through follow-up. In all models, effects from Time 1 to Time 2, Time 2 to Time 3, and Time 1 to Time 3 are tested. For each dependent variable, significance of random effects was tested using likelihood ratio tests and the best fitting model with the fewest parameters was selected. For aim 1, predictors in the model were Time, Group and the Time × Group interaction. For aim 2, predictors were Time, number of anxiety-related labels chosen, and the Time × number of anxiety-related labels interaction, and the analysis was only conducted within the AL group. For Aim 3, we first tested moderation (i.e. does incidental emotion regulation predict greater fear reduction in one group over the other) and predictors were Time, Group, Incidental Emotion Regulation, Time × Group, Time × Incidental Emotion Regulation, Group × Incidental Emotion Regulation, and Time × Group × Incidental Emotion Regulation. If the three way interaction was not significant, it was dropped from the model and we then tested prediction (i.e. does incidental emotion regulation predict fear reduction regardless of group) by examining the Time × Incidental Emotion Regulation fixed effect for significance.

1.6. Effect size

For study Aim 1, effect sizes reported are Cohen's *d*, and were calculated using an approach described by Feingold (2009) for estimating group differences in randomized clinical trials with repeated measures. For study aims 2 and 3, effect sizes reported are Cohen's f^2 and were estimated using an approach outlined by Selya, Rose, Dierker, Hedeker, and Mermelstein (2012). Cohen's f^2 uses residual variance from the model to estimate effect size. However, for multi-level models, effect sizes calculated using residual variance and proportion of variance explained should be interpreted with caution because the addition of variables to the model can, in some cases, increase residual variance resulting in negative estimates of explained variance and even of effect size (Snijders & Bosker, 1994). In addition, this method cannot be used for non-

continuous dependent measures. As a result, effect sizes are not reported for analyses with SCR-NS (count variable) as the outcome.

1.7. Power analyses

Effect sizes for group comparisons in the study by Kircanski et al. (2012) ranged from .58 to .99 depending on which outcome measure was assessed. We used G*Power to calculate the sample size needed to achieve power of 0.8, for an effect size of .60, and the goal sample size was 72 participants.

2. Results

2.1. Preliminary analyses

2.1.1. Sample characteristics
Groups did not significantly differ on any demographic or clinical characteristics at baseline (*p*s > .05). At baseline, 54% of the sample fell in the "high" public speaking anxiety range on the PRPSA (scores above 131), 46% of the sample fell in the moderate range (scores between 97 and 131), and 0% of the sample fell in the "low" public speaking anxiety range. The current sample had a mean PRPSA score of 133.0, which is approximately one standard deviation above the mean of 114.6 (SD = 17.2) observed in a college sample (McCroskey, 1970). Raw means for all dependent measures by Group over Time are displayed in Table 1.

2.1.2. Comparison of completers vs. dropout
Participants who dropped had significantly higher HR during anticipation of giving a speech, t(97) = −2.1, *p* < .05, marginally significantly higher HR during recovery after giving a speech, t(96) = −1.9, *p* < .10, and marginally significantly higher SCR-NS during anticipation of giving a speech, z = −1.8, *p* < .10.[3] Dropout rates did not differ between AL (N = 11) and Control (N = 8), χ^2 = .33, *p* = .568. Given that significant differences were found between completers and those who dropped from the study, main effects and interactions with dropout were tested in each model, and when significant, dropout was included in the final model.

2.2. Aim 1: does affect labeling enhance exposure effectiveness compared to exposure alone?

For significant Time × Group interactions, the following simple effects were tested. (1) Group differences at Time 2 and 3; (2) Simple slopes (whether slopes differ from zero) from Time 1 to Time 2, Time 2 to Time 3, and/or Time 1 to Time 3 (depending on significant interactions).

2.2.1. Physiology
For HR during recovery, the Time × Group interactions from Time 1 to Time 2 and Time 1 to Time 3 were not significant (*p*s > .153). The Time × Group interaction from Time 2 to Time 3 was significant (b = −3.79, 95% Confidence Interval (CI) = −7.4 to −.2, *p* = .041, d = .33) such that participants in AL showed a steeper decrease in HR from Time 2 to Time 3 than participants in Control (Fig. 2A). Tests of group differences revealed no significant differences at Time 2 (*p* = .419) or at Time 3 (.281). Tests of simple slopes from Time 2 to Time 3 revealed a significant increase in HR in Control (change = 2.77, *p* = .036) and no significant change in AL (*p* = .432). This significant increase in Control can be attributed to an increase in HR during baseline from Time 1 to 3 (*p* = .002) and Time 2 to 3 (*p* = .044) in Control. The baseline at Time 3 functioned

---

[3] Mann–Whitney test was used because variable is non-normal (count).

Table 1
Raw means by group and time for dependent measures.

	AL			Control		
	T1 (n = 52)	T2 (n = 43)	T3 (n = 41)	T1 (n = 48)	T2 (n = 42)	T3 (n = 40)
	Mean (SD)					
HR						
Baseline	77.1 (12.6)	77.3 (11.7)	77.4 (12.5)	72.4 (10.0)	73.9 (11.0)	76.5 (13.6)
Anticipation	81.9 (13.3)	81.3 (12.2)	80.4 (13.3)	78.3 (12.1)	78.1 (11.9)	79.8 (13.1)
Recovery	75.5 (13.1)	75.5 (10.5)	74.7 (12.7)	71.4 (9.7)	71.0 (10.3)	72.8 (12.0)
SCR-NS						
Baseline	.8 (1.3)	.5 (1.2)	.8 (1.3)	.3 (.8)	.2 (.8)	.8 (1.4)
Anticipation	2.5 (1.9)	1.9 (2.0)	1.8 (2.4)	2.5 (1.7)	1.9 (2.2)	2.0 (1.8)
Recovery	.8 (1.3)	.6 (1.3)	.3 (.7)	.6 (.9)	.4 (.7)	.6 (.9)
PRPSA	133.2 (14.7)	126.2 (17.6)	122.7 (17.8)	132.8 (16.1)	121.5 (16.1)	117.3 (15.1)
SUDS	4.9 (1.4)	3.7 (1.8)	3.3 (1.8)	4.7 (1.6)	3.5 (1.5)	3.4 (1.6)

Note. HR = Heart Rate; SCR-NS = Non Specific Skin Conductance Response; PRPSA = Personal Report of Public Speaking Anxiety; SUDS = Subjective Units of Distress Scale.

more as an anticipation period because participants were aware of the upcoming speech task. Because HR baseline was higher at Time 3 than 1 and 2, this increase in HR is not necessarily indicative of increased fear following exposure, but rather increased awareness of the upcoming speech task by Time 3.

For SCR-NS during recovery following the speech, the Time × Group interaction from Time 1 to Time 2 was not significant ($p = .587$), whereas the interactions from Time 2 to Time 3 ($b = -1.14$, CI = -2.1 to $-.2$, $p = .023$, $d = 1.0$) and Time 1 to Time 3

Fig. 2. Effect of group on heart rate (A) and non-specific skin conductance response (B) during recovery following the speech.

($b = -.90$, CI = -1.8 to 0, $p = .047$, $d = .79$) were significant such that participants in AL showed a steeper decrease in SCR-NS from Time 2 to Time 3 and from Time 1 to Time 3 than participants in Control (Fig. 2B). Tests of group differences revealed no significant group difference at Time 2 ($p = .273$) and a marginally significant difference at Time 3 (difference = .29, $p = .100$) such that participants in AL had fewer SCR-NS during recovery than participants in Control. Tests of simple slopes from Time 2 to Time 3 revealed a marginally significant reduction in SCR-NS during recovery in AL (change = $-.61$, $p = .077$), but not in Control ($p = .146$), and no significant changes from Time 1 to Time 3 in either group ($ps > .123$). Again, significant increases in baseline SCR-NS were observed in Control from Time 1 to 3 ($p = .022$) and Time 2 to 3 ($p = .001$), which likely explains the lack of significant reduction in SCR-NS following exposure.

For HR and SCR-NS during anticipation of giving a speech, the Time × Group interactions were not significant ($ps > .152$).

2.2.2. Self-report

For PRPSA and SUDS, the Time × Group interactions were not significant ($ps > .196$).

2.3. Aim 2: does the number of anxiety-related labels used during exposure predict greater attenuation of fear responding at re-test?

For significant Time × Number of Anxiety-Related Labels interactions, simple effects tested were whether participants at one standard deviation below the mean (-1SD) on Number of Anxiety-Related Labels differed from those at one standard deviation above the mean (+1SD) at Time points 1, 2 and 3.

2.3.1. Physiology

For SCR-NS during anticipation of giving a speech, the Time × Number of Anxiety-Related labels interaction from Time 1 to Time 2 was not significant ($p = .402$). The interactions from Time 2 to Time 3 ($b = -1.42$, CI = -2.8 to 0, $p = .045$), and Time 1 to Time 3 ($b = -1.92$, CI = -3.2 to $-.6$, $p = .004$) were significant such that participants who used more anxiety-related labels during exposure had a steeper decline in SCR-NS over time. Tests of the difference between SCR-NS during anticipation for participants at +1SD and -1SD from the mean revealed no significant difference at Time 1 or Time 2 ($ps > .117$), and a significant difference at Time 3 (difference = 1.70, $p = .001$) such that participants at +1SD from the mean on use of anxiety-related labels had fewer SCR-NS during anticipation of giving a speech than participants at -1SD from the mean.

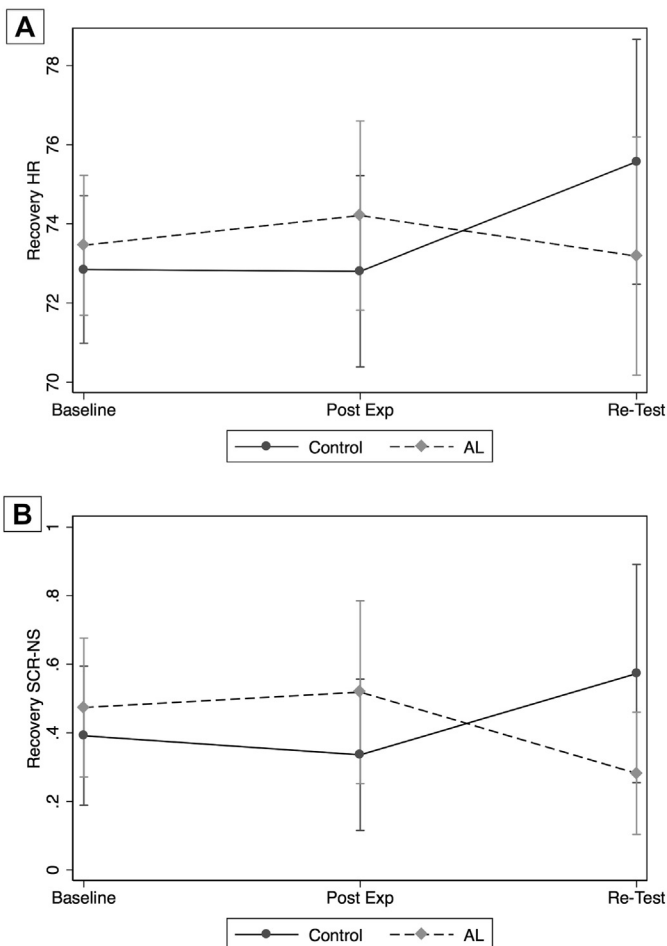For HR during anticipation of and recovery following the speech,

the Time × number of anxiety-related labels interactions were not significant (p > .325). For SCR-NS during recovery following the speech, the Time × number of anxiety-related labels interactions were not significant (ps > .151).

2.3.2. Self-report

For PRPSA, the Time × number of anxiety-related labels interactions from Time 2 to Time 3 and Time 1 to Time 3 were not significant (ps > .095). The interaction from Time 1 to Time 2 (b = 15.76, CI = 1.6 to 29.9, p = .029) was significant such that participants who used fewer anxiety-related labels during exposure had a steeper decline in PRPSA over time (interaction f^2 = .05). Tests of the difference between PRPSA for participants at +1SD and -1SD from the mean on number of anxiety-related labels revealed no significant difference at Time 1 (p = .368), and significant differences at Time 2 (difference = 11.70, p = .011) and Time 3 (difference = 11.52, p = .032) such that participants at +1SD from the mean on use of anxiety-related labels had higher PRPSA scores than participants at -1SD from the mean.

For SUDS, the Time × number of anxiety-related labels interactions from Time 2 to Time 3 and Time 1 to Time 3 were not significant (ps > .058). The interaction from Time 1 to Time 2 was significant (b = 1.72, CI = .3 to 3.2, p = .019) such that participants who used fewer anxiety-related labels during exposure had a steeper decline in SUDS over time (interaction f^2 = .05). Tests of the difference between SUDS for participants at +1SD and -1SD from the mean on number of anxiety-related labels revealed a marginally significant difference at Time 1 (difference = .67, p = .083), and significant differences at Time 2 (difference = 1.52, p < .001) and Time 3 (difference = 1.45, p = .003) such that participants at +1SD from the mean on use of anxiety-related labels had higher SUDS scores than participants at -1SD from the mean.

2.4. Aim 3: does incidental emotion regulation at baseline moderate or predict response to exposure with affect labeling versus exposure alone?

For significant Time × Group × Incidental Emotion Regulation (moderation) effects, the following simple effects were tested: (1) Group mean differences at Time 1, Time 2, and Time 3 at +1SD and -1SD from the mean of Incidental Emotion Regulation. (2) Group slope differences from Time 1 to Time 2, Time 2 to Time 3, and/or Time 1 to Time 3 (depending on significant interactions) at +1SD and -1SD from the mean of Incidental Emotion Regulation. For significant Time × Incidental Emotion Regulation (prediction) effects, we tested whether participants at -1SD from the mean on Incidental Emotion Regulation differed from those at +1SD from the mean at Time points 1, 2 and 3.

2.4.1. Moderation
2.4.1.1. Physiology.
For HR during anticipation, the Time × Group × Incidental Emotion Regulation interactions from Time 1 to Time 2 and Time 2 to Time 3 (ps > .269) were not significant, whereas the interaction from Time 1 to Time 3 was significant (b = 7.56, CI = .3 to 14.9, p = .042) (interaction f^2 = .04). Results are displayed in Fig. 3. Tests of group mean differences at -1SD and +1SD from the mean revealed no differences at Time 1 (ps > .745), or Time 2 (ps > .443). At Time 3, for participants at -1SD from the mean on Incidental Emotion Regulation, participants in AL had significantly lower HR than participants in Control (difference = 6.77, p = .014). No group difference was found for participants at +1SD from the mean on Incidental Emotion Regulation at Time 3 (p = .592). Tests of Group slope differences from Time 1 to Time 3 revealed that for participants at -1SD from the mean on Incidental Emotion Regulation, AL had a significantly

more negative slope than Control (slope difference = 7.57, p = .019). No group slope difference was found for participants at +1SD from the mean on Incidental Emotion Regulation (p = .543).

For HR during recovery following the speech, the Time × Group × Incidental Emotion Regulation interactions from Time 1 to Time 2 and Time 2 to Time 3 were not significant (ps > .063), whereas the interaction from Time 1 to Time 3 was significant (b = 8.48, CI = 2.2 to 14.8, p = .008) (interaction f^2 = .07). Tests of group mean differences at -1SD and +1SD from the mean revealed no differences at Time 1 (ps > .207) or Time 2 (ps > .301). At Time 3, for participants at -1SD from the mean on Incidental Emotion Regulation, participants in AL had marginally significantly lower HR than participants in Control (difference = 4.72, p = .061). No group mean differences were observed for participants at +1SD from the mean on Incidental Emotion Regulation at Time 3 (p = .329). Tests of Group slope differences from Time 1 to Time 3 revealed that for participants at -1SD from the mean on Incidental Emotion Regulation, AL had a significantly more negative slope than Control (slope difference = 7.56, p = .007). No group slope difference was observed for participants at +1SD from the mean on Incidental Emotion Regulation (p = .262).

For SCR-NS during recovery following the speech, the Time × Group × Incidental Emotion Regulation interactions from Time 1 to Time 3 and Time 2 to Time 3 were not significant (ps > .080), whereas the interaction from Time 1 to Time 2 was significant (b = 2.37, CI = .5 to 4.2, p = .012). Tests of group mean differences at -1SD and +1SD revealed no differences at Time 1 (ps > .432) or Time 3 (ps > .200). At Time 2, for participants at +1SD from the mean on Incidental Emotion Regulation, participants in AL had significantly higher SCR-NS than participants in Control (difference = .48, p = .010). At Time 2, no group differences emerged for participants at -1SD from the mean on Incidental Emotion Regulation (p = .487). Tests of Group slope differences from Time 1 to Time 2 revealed that for participants at +1SD from the mean on Incidental Emotion Regulation, AL had a significantly more positive slope than Control (slope difference = .58, p = .027). No group slope difference was observed for participants at -1SD from the mean on Incidental Emotion Regulation (p = .236).

For SCR-NS during anticipation of giving a speech, the Time × Group × Incidental Emotion Regulation interactions were not significant (ps > .267).

2.4.1.2. Self-report.
For PRPSA and SUDS, the Time × Group × Incidental Emotion Regulation interactions were not significant (ps > .085).

2.4.2. Prediction
2.4.2.1. Physiology.
For SCR-NS during anticipation of giving a speech, the Time × Incidental Emotion Regulation interactions from Time 1 to Time 2 and Time 1 to Time 3 were not significant (ps > .125), whereas the Time × Incidental Emotion Regulation interaction from Time 2 to Time 3 was significant (b = .53, CI = .1 to .9, p = .009) such that participants at -1SD from the mean on Incidental Emotion Regulation had more negative slopes in SCR-NS from Time 2 to Time 3 than participants at +1SD from the mean on Incidental Emotion Regulation regardless of group assignment. Tests of the difference between SCR-NS during anticipation for participants at +1SD and -1SD from the mean revealed no significant differences at Time 1 or Time 3 (ps > .331), and a significant difference at Time 2 (difference = .90, p = .032) such that participants at +1SD from the mean on Incidental Emotion Regulation had lower SCR-NS than participants at -1SD from the mean.

2.4.2.2. Self-report.
For PRPSA and SUDS, the Time × Incidental Emotion Regulation interactions were not significant (ps > .061).
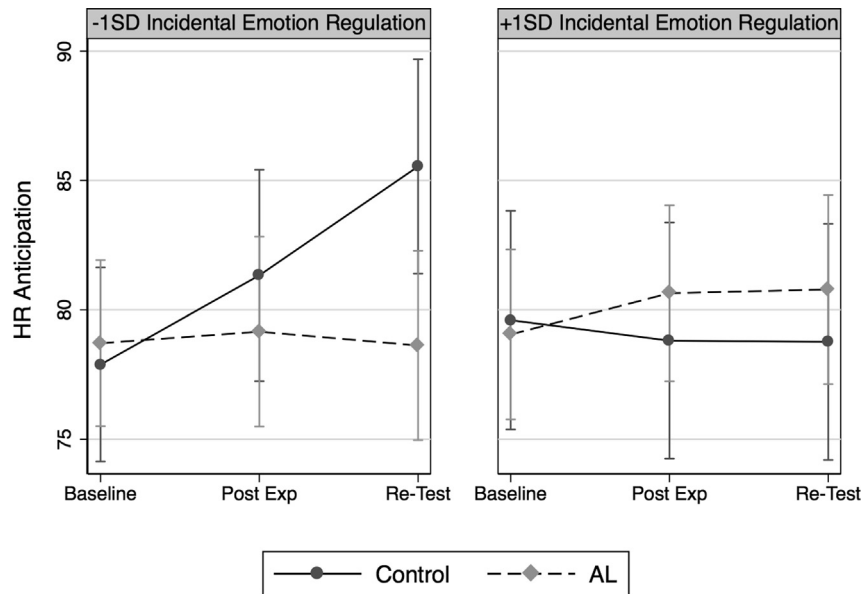
Fig. 3. Moderation of Heart Rate during anticipation over Time by Incidental Emotion Regulation and Group.

3. Discussion

The goals of the current study were to test whether affect labeling enhanced the effectiveness of exposure for public speaking anxiety, whether the use of more anxiety-related labels during exposure predicted better outcome, and to test incidental emotion regulation as a potential moderator of response to exposure plus affect labeling versus exposure alone. We found that affect labeling enhanced the effectiveness of exposure on measures of physiological arousal, particularly for participants who used more anxiety-related labels during exposure. We also found that participants with deficits in incidental emotion regulation benefitted more from exposure plus affect labeling than from exposure alone on measures of physiological arousal.

Consistent with hypotheses, participants in the exposure plus affect-labeling group had a steeper decline in heart rate and in nonspecific skin conductance responses during recovery following the speech than participants in the exposure alone condition. This finding is consistent with previous research showing that exposure combined with affect labeling results in greater reduction in physiological arousal than exposure alone (Kircanski et al., 2012; Tabibnia et al., 2008). It is notable that the effect was found for skin conductance response and heart rate during recovery following the speech only and not in anticipation of the speech. Although research on public speaking tasks generally focuses on anticipation of speaking, post-event rumination is common in individuals with social anxiety, relates to the severity of social anxiety symptoms, and predicts subsequent avoidance of similar social situations (Rachman, Grüter-Andrew, & Shafran, 2000). Therefore, the attenuation of physiological activation following the speech may reflect beneficial effects of affect labeling upon negative post event rumination.

Consistent with previous research (Kircanski et al., 2012), the benefit of affect labeling during exposure was found only for physiological measures and not for self-report measures. One possible explanation is the tendency for people to predict that affect labeling will not effectively reduce distress (Lieberman et al., 2011); As suggested by Kircanski et al. (2012) this bias may reflect in subjective self-report measures, but not in more objective measures of physiological arousal. Another possibility is that by

repeatedly labeling anxiety and feared outcomes during exposure, participants were trained to report anxiety symptoms on self-report measures administered during subsequent assessments and therefore did not show a reduction in self reported anxiety. A third possibility, consistent with mindfulness based approaches to the treatment of anxiety, is that participants who engaged in affect labeling continued to experience and report anxiety symptoms at follow-up assessments, but were less distressed by these symptoms, and therefore less physiologically reactive. A phrase often used in mindfulness practice is that the goal is not to feel better, but to get better at feeling (Hayes, Strosahl, & Wilson, 1999). Finally, it is possible that self-reported anxiety takes more time to modify and that effects would be observed if participants were followed for a longer period of time. Future studies that include training in affect labeling should measure acceptance of anxiety in addition to anxiety symptoms and should extend the follow-up period.

Consistent with hypotheses, the more anxiety labels participants chose during exposure, the fewer non-specific skin conductance responses participants had during anticipation of giving a speech by one-week follow-up. These findings are consistent with previous research in spider fearful participants where participants who used more anxiety words during affect labeling showed a greater reduction in non-specific skin conductance response during anticipation of touching a spider at re-test than participants who used fewer anxiety words (Kircanski et al., 2012). Perhaps more frequent use of affect labels during exposure led to greater activation in PFC-amygdala pathways. Participants with social anxiety show weaker connectivity between areas of the PFC and the amygdala (Hahn et al., 2011), and the number of repetitions of activation in PFC-amygdala pathways during exposure may positively correlate with the strength of PFC-amygdala connectivity following completion of exposure. Consistent with the principle of neural plasticity, which states that repetition of a process can increase efficiency and efficacy of that process through changes in neuron function, chemical profile, and structure (Anderson, 2010; Kandel & Schwartz, 1982), greater activation of PFC-amygdala neural pathways as a result of more frequent labeling may have produced greater neural change and ultimately more effective down regulation of physiological fear responding. It is notable however that a significant effect for the number of anxiety labels

was found for the anticipation period, whereas significant affect labeling versus control group differences were found only during recovery. The effects for the number of anxiety-related labels for the recovery period were in the predicted direction, but the effect was not robust enough to reach significance, and the lack of consistency in whether effects were found for anticipation or recovery across the two study aims is a limitation of the current findings.

Contrary to hypotheses, participants who used fewer affect labels during exposure showed greater improvement in self-reported public speaking anxiety than participants who used more affect labels during exposure. Similar explanations to those provided above regarding the self-report data can be applied to these findings as well. In particular, high labelers may have been even more likely than low labelers in the affect labeling group to experience a training effect during exposure, thereby increasing anxious reporting on self-report measures at subsequent assessments. However, non-significant interactions from the baseline to one-week follow-up suggest that, had participants been followed for a longer period of time, those who labeled more frequently may have shown continued improvement on self-report measures consistent with physiological improvements, while those who labeled less frequently may have shown a return of symptoms.

Finally, participants who had deficits in incidental emotion regulation at baseline benefited more from exposure combined with affect labeling than exposure alone, whereas for participants with strengths in incidental emotion regulation, no group differences were found. This differential moderation effect was found for heart rate during anticipation and recovery from the speech. The results from skin conductance were mixed and transient, not lasting beyond Time 2. Also, participants with deficits in incidental emotion regulation benefited more overall, regardless of group, in terms of skin conductance during recovery. The findings for heart rate during anticipation and recovery and for skin conductance during recovery support the compensation hypothesis, which suggests that those with deficits will benefit more from an intervention that compensates that deficit, and are consistent with previous research showing that physiological activation (Davies et al., 2015; McClure et al., 2007) and emotional reactivity during a behavioral task (Niles et al., 2013) predict greater improvement following treatment. These findings are of particular interest because participants with more difficulty regulating emotion show decreased connectivity between areas of the PFC and the amygdala (Banks, Eddy, Angstadt, Nathan, & Phan, 2007). These findings suggest that targeting patients with the greatest neural deficits using interventions that directly address those deficits could improve treatment outcome for patients with anxiety. Neural activation, however, was not directly assessed in the current project, and future studies that include affect labeling as a strategy for enhancing exposure would benefit from inclusion of measures of neural activation before and after the intervention to determine whether the intervention enhances connectivity.

Although the current study has many strengths, there are a number of notable limitations. First, participants assigned to the affect labeling group were given the option to choose "other" during the exposures rather than an emotion label. Many participants did opt to choose "other," and it is unclear whether these participants received the intervention as intended by the researchers. This allowed us to assess a dose response relationship between affect labeling and outcome, but also introduced heterogeneity within the affect labeling group, potentially dampening group differences. Because many participants did not select any of the available emotions ostensibly because the available options did not match their emotional experience, it is possible that allowing participants to choose their own labels as opposed to choosing from a predetermined list would prove even more powerful for

enhancing exposure effectiveness. Future research should assess differential effects of predetermined and participant identified affect labels. A second limitation is that we were unable to assess true baseline physiological arousal at time points 2 and 3 because participants were no longer naïve to study procedures. Elevated baseline physiological activation may have artificially elevated anticipatory arousal, making interpretation of change over time on physiological measures difficult.

In sum, the current research supports the theory that affect labeling can enhance exposure effectiveness. Patients with anxiety show deficits in prefrontal amygdala connectivity (Hahn et al., 2011; Kim et al., 2011), and repeated activation of prefrontal regions that project to the amygdala through exposure and affect labeling can lead to a reduction in physiological activation in response to an anxiety provoking stimulus. The results of the current study indicate that instruction to label emotional experiences improves physiological attenuation of fear, and that the more a fearful individual labels his or her emotional experience during exposure, the greater the reduction in galvanic skin response (a measure of fear arousal) when they next encounter the feared stimulus. The benefit of affect labeling was not shown for self-reported anxiety, which may require longer-term follow up than was possible in the current study. Finally, adding affect labeling to exposure was particularly beneficial (on heart rate) for individuals who showed deficits in incidental emotion regulation, which is an indicator of poor prefrontal-amygdala connectivity. This finding provides further evidence that targeting prefrontal-amygdala circuitry in anxiety patients using tasks that activate key regions involved in emotion regulation can improve treatment effectiveness, and that such interventions will be particularly effective for patients who show the greatest deficits in this circuit.

Author contributions

A.N.N. developed the study concept, collected and analyzed the data and drafted the manuscript under the supervision in M.G.C. M.G.C. aided in study design, interpretation of results, and critical revision of the manuscript, M.D.L. aided in study design, interpretation of results, and critical revision of the manuscript, and C.H. aided with data collection and oversight of study procedures. All authors approved the final version of the paper for submission.

Acknowledgments

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.brat.2015.03.004.

References

Anderson, M. L. (2010). Neural reuse: a fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 33*(4), 245.

Arch, J. J., & Craske, M. G. (2009). First-line treatment: a critical appraisal of cognitive behavioral therapy developments and alternatives. *The Psychiatric Clinics of North America, 32*(3), 525–547. http://dx.doi.org/10.1016/j.psc.2009.05.001.

Banks, S. J., Eddy, K. T., Angstadt, M., Nathan, P. J., & Phan, K. L. (2007). Amygdala-frontal connectivity during emotion regulation. *Social Cognitive and Affective Neuroscience, 2*(4), 303–312. http://dx.doi.org/10.1093/scan/nsm029.

Braithwaite, J. J., Watson, D. G., Jones, R., & Rowe, M. (2013). *A guide for analysing electrodermal activity (EDA) and skin conductance responses (SCRs) for*

psychophysiological experiments via the MP36R and AcqKnowledge software (No. 1). University of Birmingham. UK: Selective Attention & Awareness Laboratory (SAAL), Behavioral Brain Sciences Centre, School of Psychology.

Burklund, L. J., Creswell, J. D., Irwin, M. R., & Lieberman, M. D. (2014). The common and distinct neural bases of affect labeling and reappraisal in healthy adults. *Frontiers in Psychology, 5*, 221. http://dx.doi.org/10.3389/fpsyg.2014.00221.

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clinical Psychology Review, 26*(1), 17–31. http://dx.doi.org/10.1016/j.cpr.2005.07.003.

Cheavens, J. S., Strunk, D. R., Lazarus, S. A., & Goldstein, L. A. (2012). The compensation and capitalization models: a test of two approaches to individualizing the treatment of depression. *Behaviour Research and Therapy, 50*(11), 699–706. http://dx.doi.org/10.1016/j.brat.2012.08.002.

Clark, D. M., Ehlers, A., Hackmann, A., McManus, F., Fennell, M., Grey, N., et al. (2006). Cognitive therapy versus exposure and applied relaxation in social phobia: a randomized controlled trial. *Journal of Consulting and Clinical Psychology, 74*(3), 568–578. http://dx.doi.org/10.1037/0022-006X.74.3.568.

Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. Hope, & F. R. Schneier (Eds.), *Social phobia* (pp. 69–94). New York, NY: Guilford Press.

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy, 46*(1), 5–27. http://dx.doi.org/10.1016/j.brat.2007.10.003.

Craske, M. G., Liao, B., Brown, L., & Vervliet, B. (2012). Role of inhibition in exposure therapy. *Journal of Experimental Psychopathology, 3*, 322–345.

Culver, N. C., Stoyanova, M., & Craske, M. G. (2012). Emotional variability and sustained arousal during exposure. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(2), 787–793. http://dx.doi.org/10.1016/j.jbtep.2011.10.009.

Cunningham, W. A., Johnson, M. K., Chris, J., Gore, J. C., & Banaji, M. R. (2003). Neural components of social evaluation. *Journal of Personality and Social Psychology, 85*(4), 639–649. http://dx.doi.org/10.1037/0022-3514.85.4.639.

Davidson, J. R., Foa, E. B., Huppert, J. D., Keefe, F. J., Franklin, M. E., Compton, J. S., et al. (2004). Fluoxetine, comprehensive cognitive behavioral therapy, and placebo in generalized social phobia. *Archives of General Psychiatry, 61*(10), 1005.

Davies, C. D., Niles, A. N., Pittig, A., Arch, J. J., & Craske, M. G. (2015). Physiological and behavioral indices of emotion dysregulation as predictors of outcome from cognitive behavioral therapy and acceptance and commitment therapy for anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 46*, 35–43. http://dx.doi.org/10.1016/j.jbtep.2014.08.002.

Davis, M. (1992). The role of the amygdala in fear and anxiety. *Annual Review of Neuroscience, 15*(1), 353–375. http://dx.doi.org/10.1146/annurev.ne.15.030192.002033.

Delgado, M. R., Nearing, K. I., LeDoux, J. E., & Phelps, E. A. (2008). Neural circuitry underlying the regulation of conditioned fear and its relation to extinction. *Neuron, 59*(5), 829–838.

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14*(1), 43–53. http://dx.doi.org/10.1037/a0014699.

Feske, U., & Chambless, D. L. (1995). Cognitive behavioral versus exposure only treatment for social phobia: a meta-analysis. *Behavior Therapy, 26*(4), 695–720. http://dx.doi.org/10.1016/S0005-7894(05)80040-1.

Gorno-Tempini, M. L., Pradelli, S., Serafini, M., Pagnoni, G., Baraldi, P., Porro, C., et al. (2001). Explicit and incidental facial expression processing: an fMRI study. *NeuroImage, 14*(2), 465–473. http://dx.doi.org/10.1006/nimg.2001.0811.

Hahn, A., Stein, P., Windischberger, C., Weissenbacher, A., Spindelegger, C., Moser, E., et al. (2011). Reduced resting-state functional connectivity between amygdala and orbitofrontal cortex in social anxiety disorder. *NeuroImage, 56*(3), 881–889. http://dx.doi.org/10.1016/j.neuroimage.2011.02.064.

Hariri, A. R., Bookheimer, S. Y., & Mazziotta, J. C. (2000). Modulating emotional responses: effects of a neocortical network on the limbic system. *Neuroreport, 11*(1), 43–48.

Hariri, A. R., Mattay, V. S., Tessitore, A., Fera, F., & Weinberger, D. R. (2003). Neocortical modulation of the amygdala response to fearful stimuli. *Biological Psychiatry, 53*(6), 494–501.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: The Guilford Press.

Heimberg, R. G. (2002). Cognitive-behavioral therapy for social anxiety disorder: current status and future directions. *Biological Psychiatry, 51*(1), 101–108.

Hofmann, S. G., & Smits, J. A. J. (2008). Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. *The Journal of Clinical Psychiatry, 69*(4), 621–632.

Hope, D. A., Heimberg, R. G., Juster, H. A., & Turk, C. L. (2000). *Managing social anxiety: A cognitive-behavioral therapy approach client workbook*. New York, NY: Oxford University Press.

Kandel, E. R., & Schwartz, J. H. (1982). Molecular biology of learning: modulation of transmitter release. *Science, 218*(4571), 433–443.

Kim, M. J., Loucks, R. A., Palmer, A. L., Brown, A. C., Solomon, K. M., Marchante, A. N., et al. (2011). The structural and functional connectivity of the amygdala: from

normal emotion to pathological anxiety. *Behavioural Brain Research, 223*(2), 403–410. http://dx.doi.org/10.1016/j.bbr.2011.04.025.

Kircanski, K., Lieberman, M. D., & Craske, M. G. (2012). Feelings into words: contributions of language to exposure therapy. *Psychological Science*. http://dx.doi.org/10.1177/0956797612443830.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1999). *International affective picture system (IAPS): Technical manual and affective ratings*. Gainesville, FL: The Center for Research in Psychophysiology, University of Florida.

Lieberman, M. D. (2003). Reflexive and reflective judgment processes: a social cognitive neuroscience approach. In J. P. Forgas, K. D. Williams, & W. von (Eds.), *Social judgments: Implicit and explicit processes* (pp. 44–67). New York, NY: Cambridge University Press.

Lieberman, M. D. (2011). Why symbolic processing of affect can disrupt negative affect: social cognitive and affective neuroscience investigations. *Social Neuroscience: Toward Understanding the Underpinnings of the Social Mind*, 188–209.

Lieberman, M. D., Eisenberger, N. I., Crockett, M. J., Tom, S. M., Pfeifer, J. H., & Way, B. M. (2007). Putting feelings into words: affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science, 18*(5), 421–428. http://dx.doi.org/10.1111/j.1467-9280.2007.01916.x.

Lieberman, M. D., Inagaki, T. K., Tabibnia, G., & Crockett, M. J. (2011). Subjective responses to emotional stimuli during labeling, reappraisal, and distraction. *Emotion, 11*(3), 468–480. http://dx.doi.org/10.1037/a0023503.

McClure, E. B., Adler, A., Monk, C. S., Cameron, J., Smith, S., Nelson, E. E., et al. (2007). fMRI predictors of treatment outcome in pediatric anxiety disorders. *Psychopharmacology, 191*(1), 97–105. http://dx.doi.org/10.1007/s00213-006-0542-9.

McCroskey, J. C. (1970). Measures of communication-bound anxiety. *Speech Monographs, 37*(4), 269–277. http://dx.doi.org/10.1080/03637757009375677.

Milad, M. R., Quinn, B. T., Pitman, R. K., Orr, S. P., Fischl, B., & Rauch, S. L. (2005). Thickness of ventromedial prefrontal cortex in humans is correlated with extinction memory. *Proceedings of the National Academy of Sciences of the United States of America, 102*(30), 10706–10711.

Milad, M. R., & Quirk, G. J. (2002). Neurons in medial prefrontal cortex signal memory for fear extinction. *Nature, 420*(6911), 70–74. http://dx.doi.org/10.1038/nature01138.

Narumoto, J., Yamada, H., Iidaka, T., Sadato, N., Fukui, K., Itoh, H., et al. (2000). Brain regions involved in verbal or non-verbal aspects of facial emotion recognition. *Neuroreport, 11*(11), 2571–2576.

Niles, A. N., Mesri, B., Burklund, L. J., Lieberman, M. D., & Craske, M. G. (2013). Attentional bias and emotional reactivity as predictors and moderators of behavioral treatment for social phobia. *Behaviour Research and Therapy, 51*(10), 669–679. http://dx.doi.org/10.1016/j.brat.2013.06.005.

Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron, 43*(6), 897–905. http://dx.doi.org/10.1016/j.neuron.2004.08.042.

Pollard, C. A., & Henderson, J. G. (1988). Four types of social phobia in a community sample. *The Journal of Nervous and Mental Disease, 176*(7), 440–445.

Rachman, S., Grüter-Andrew, J., & Shafran, R. (2000). Post-event processing in social anxiety. *Behaviour Research and Therapy, 38*(6), 611–617. http://dx.doi.org/10.1016/S0005-7967(99)00089-3.

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy, 35*(8), 741–756.

Rude, S. S., & Rehm, L. P. (1991). Response to treatments for depression: the role of initial status on targeted cognitive and behavioral skills. *Clinical Psychology Review, 11*(5), 493–514. http://dx.doi.org/10.1016/0272-7358(91)90001-B.

Selya, A. S., Rose, J. S., Dierker, L. C., Hedeker, D., & Mermelstein, R. J. (2012). A practical guide to calculating cohen's f(2), a measure of local effect size, from PROC MIXED. *Frontiers in Psychology, 3*, 111. http://dx.doi.org/10.3389/fpsyg.2012.00111.

Snijders, T. A. B., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods & Research, 22*(3), 342–363. http://dx.doi.org/10.1177/0049124194022003004.

Stein, M. B., Walker, J. R., & Forde, D. R. (1996). Public-speaking fears in a community sample. Prevalence, impact on functioning, and diagnostic classification. *Archives of General Psychiatry, 53*(2), 169–174.

Tabibnia, G., Lieberman, M. D., & Craske, M. G. (2008). The lasting effect of words on feelings: words may facilitate exposure effects to threatening images. *Emotion, 8*(3), 307.

Tolin, D. F. (2010). Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review, 30*(6), 710–720. http://dx.doi.org/10.1016/j.cpr.2010.05.003.

Tsao, J. C. I., & Craske, M. G. (2000). Timing of treatment and return of fear: effects of massed, uniform-, and expanding-spaced exposure schedules. *Behavior Therapy, 31*(3), 479–497. http://dx.doi.org/10.1016/S0005-7894(00)80026-X.

Wingate, L. R., Van Orden, K. A., Joiner, T. E., Jr., Williams, F. M., & David, M. (2005). Comparison of compensation and capitalization models when treating suicidality in young adults. *Journal of Consulting and Clinical Psychology, 73*(4), 756–762. http://dx.doi.org/10.1037/0022-006X.73.4.756.

Wittchen, H. U., Stein, M. B., & Kessler, R. C. (1999). Social fears and social phobia in a community sample of adolescents and young adults: prevalence, risk factors and co-morbidity. *Psychological Medicine, 29*(2), 309–323.