The Effects of Systematic Self-Monitoring, Feedback, and Collaborative Assessment on Stress Reduction

Jessica Ann Chen

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Ronald E. Smith, Chair

Yuichi Shoda

Peter Vitaliano

Program Authorized to Offer Degree:

Psychology

© Copyright 2016
Jessica Ann Chen

University of Washington

Abstract

The Effects of Systematic Self-Monitoring, Feedback, and Collaborative Assessment on Stress Reduction

Jessica Ann Chen

Chair of Supervisory Committee:
Ronald E. Smith, Ph.D.
Department of Psychology

The goal of the present study was to test the potential efficacy of an online stress management intervention, the daily diary; to develop and test a personalized feedback intervention for the diary; and to test the effectiveness of an empirically-supported, cognitive behavioral stress management program when augmented with collaborative assessment procedures. Based on the assumptions of a stepped-care model, I hypothesized that each more intensive intervention would lead to added benefits in terms of stress reduction, increased positive affect, and decreased negative affect. 111 undergraduate students who scored 20 or higher on the Perceived Stress Scale participated in either the diary + feedback intervention (n = 33), the CA-enhanced stress management training (n = 43), or the control condition (n = 35). Ninety participants (81.1%) completed their allocated intervention and 75 (67.6%) completed the post-study assessment. Our results pointed to a significant association between diary adherence and stress reduction (r = .31), and suggested that, when compared to a control group, CA-enhanced SMT produced moderate decreases in stress (d = .45). To our knowledge, this is the first study to use an online daily diary as a stress management intervention and the first to test ongoing collaborative assessment as an adjunct to evidence-based psychotherapy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

The Effects of Systematic Self-Monitoring, Feedback, and Collaborative Assessment on Stress Reduction

Introduction

The most recent *Stress in America*™ survey, conducted annually by the American Psychological Association, showed once again that Americans' stress levels are higher now than they were a generation ago (American Psychological Association, 2015). Younger people, particularly those born after 1980, known as "Millenials," reported the highest stress levels of any adult generation. More than one-third of Millenials stated they "watch TV or movies for more than two hours per day" and "surf the Internet" to relieve stress (American Psychological Association, 2015). High stress levels combined with a reliance on sedentary coping strategies makes this generation of younger Americans a prime population for targeted interventions that teach self-regulatory strategies for handling everyday, chronic stressors.

Additionally, among Americans of all ages who reported significant financial worries, more than half relied on sedentary coping strategies (American Psychological Association, 2015). One in five surveyed Americans reported considering delaying health care (e.g., skipping doctors' visits, not seeking out care) because of financial concerns. Given the well-known negative impact stress has on physical and psychological wellbeing (Schneiderman, Ironson, & Siegel, 2005), there is a public health imperative to identify low-cost interventions that reduce stress, particularly for at-risk populations, which may include Millenials and those with more financial worries.

Internet-based interventions may be ideal for both of these populations because they are cost-effective and appeal to younger people (Rice et al., 2014). Systematic reviews and meta-analyses of randomized clinical trials (RCTs) have reached inconsistent conclusions for technology-delivered interventions, which include both online/Internet-based interventions and offline computer programs. Technology-delivered interventions have been shown to be

comparable to face-to-face psychotherapy for some anxiety disorders (Christensen, Batterham, & Calear, 2014; Dèttore, Pozza, & Andersson, 2015) and effective for depression and PTSD compared to controls (Arnberg, Linton, Hultcrantz, Heintz, & Jonsson, 2014; Cowpertwait & Clarke, 2013; Sloan, Gallagher, Feinstein, Lee, & Pruneau, 2011). However, another meta-analysis found that technology-delivered interventions for university students did not result in significant benefits for anxiety, depression, or distress compared to controls (Davies, Morriss, & Glazebrook, 2014).

Stress Management and College Students

Among college students, technology-delivered interventions for stress have shown more consistent positive results than interventions for diagnosable mental health conditions, but the results are still mixed. One of the earliest studies tested the efficacy of an online, personalized text and graphical feedback intervention. In this study, college students with above-average stress received psychoeducation, peer stories, and stress management techniques based on their answers to a stress questionnaire (Chiauzzi, Brevard, Thurn, Decembrele, & Lord, 2008). The authors found a significant decline in stress for the feedback group post-intervention, but this decline was matched by the control groups by the six-month follow-up (Chiauzzi et al., 2008). Another RCT found that a six-session, computerized stress intervention, consisting of psychoeducation, cognitive behavioral activities, and journaling, resulted in a significant decline in stress among graduate students compared to a stress informational video group (Rose et al., 2013). However, this computerized intervention was administered in-person, in a clinic, which would not necessarily increase accessibility compared to face-to-face psychoatherapy.

More recently, two studies have tested the efficacy of online interventions that specifically target either mindfulness or present control as mechanisms for stress reduction. A self-guided, online mindfulness intervention was found to reduce stress among college students (Cavanagh et al., 2013), but the intervention was not directly compared to a control group. A more recently published study found that a four-session, online intervention consisting of psychoeducation,

three self-reflective stress logs, and exercises to improve present control (i.e., exercising control over how one deals with a given event in the present), led to a significant decline in stress compared to an information-only control (Hintz, Frazier, & Meredith, 2015). As this study's primary goal was to assess the intervention's effects on present control, participants were not recruited for particular stress characteristics. Thus, it is unclear whether the intervention effects would be the same for a sample with above average stress levels.

In summary, a number of technology-delivered interventions have shown promise for reducing stress among college students, but few have tested the efficacy of an online intervention for individuals reporting high stress levels.

Stepped Care and Collaborative Assessment

A final consideration with technology-delivered interventions is that online and face-to-face interventions need not be an either/or proposition. At the population level, online interventions may provide an excellent, low-intensity health care option, particularly for individuals dealing with adjustment issues or less severe problems; however, online interventions could also be used to supplement or enhance face-to-face psychotherapy, possibly increasing efficiency and cost-effectiveness. That is, online interventions might address the problem of limited treatment resources, namely the time and availability of mental health professionals, without the expectation that they replace face-to-face psychotherapy.

In fact, stepped care models suggest that the problem of limited treatment resources be addressed in exactly this way. In a stepped-care model, low-intensity treatments, such as self-help materials and technology-delivered programs, are made available initially to a wide array of individuals. Systematic monitoring of treatment outcomes then identifies those who do not benefit from lower-intensity treatment, triggering stepped up care to higher-intensity treatments with more specialized providers (Bower & Gilbody, 2005). To date, few studies have tested the possibility of using an online intervention within a stepped care model.

Self-Monitoring Reactivity

One of the lowest intensity interventions available is allowing individuals to merely monitor their own behavior. Self-monitoring has been shown to activate behavior change (Febbraro & Clum, 1998), and different theorists have proposed a variety of mechanisms to explain this effect. Self-regulation theory posits that self-monitoring triggers awareness and evaluation of one's performance, which can lead to behavior change when there is a discrepancy between current performance and future goals (Bandura, 1991; Carver & Scheier, 1998). Self-monitoring has also been described as a "disengagement from automaticity, or a transition from 'mindlessness' to 'mindfulness' " (Karoly, 1993, p. 33) because it requires attention to the antecedents and consequences of behavior. Neuroimaging research suggests mindfulness enhances self-regulation of attention and emotion (Hölzel et al., 2011; Vago & Silbersweig, 2012) and mindfulness has been found to mediate changes in depression and anxiety (Gu, Strauss, Bond, & Cavanagh, 2015). Consequently, one may infer that self-monitoring may work by influencing attentional processes that are necessary for the regulation of distressing states, e.g., anxiety, depression, or stress. Finally, the narrative theory of Pennebaker and colleagues (1990) (Pennebaker, Colder, & Sharp, 1990) suggests that reflecting on one's emotionally significant experiences enables an integration of thoughts, feelings, and behaviors, which facilitates self-efficacy and a sense of control over one's life (Pennebaker & Seagal, 1999). Self-efficacy has been found to mediate changes in anxiety and stress (Clarke et al., 2014). Across multiple theories, the core function of self-monitoring seems to be increasing awareness, attention, and/or insight, which promotes self-regulation and improves self-perception of one's ability to handle stress.

Febbraro and Clum (1998) conducted a meta-analysis on the effectiveness of self-monitoring (SM). Compared to no intervention at all, SM yielded a small but significant effect size (d = .29). When interventions utilizing other self-regulation techniques (e.g., self-reinforcement) were added, the comparison of SM-plus to waitlist/minimal contact control groups yielded a medium effect size (d = .45), suggesting that self-regulation interventions are effective. The somewhat

modest effect sizes must be counterbalanced with the fact that SM is often far less time intensive than face-to-face psychotherapy.

In summary, self-monitoring is a promising low-intensity intervention that may be easily adapted to an online format. To our knowledge, a self-monitoring intervention has not been tested for stress management.

Personalized Feedback

Self-monitoring provides informal feedback to the individual about their own behavior. Feedback may also constitute a stand-alone intervention when delivered by a professional in a systematic fashion. Within the stepped care model, systematic feedback from a professional may be slightly more intensive than self-monitoring alone, but it is still a relatively hands-off intervention compared to face-to-face psychotherapy.

A recent review of 41 studies concluded that personalized feedback is an effective intervention for problematic college student drinking whether delivered in-person, by mail, or online (Miller et al., 2013). Personalized feedback interventions, based on self-regulation theory and motivational interviewing, are thought to work by providing individuals with discrepant information that prompts behavior change (Walters & Neighbors, 2005). The most common types of feedback include normative feedback (e.g., a comparison of self-reported drinking with norms for one's peer group), personal profiles (e.g., self-reported drinking frequency and peak consumption), consequences (e.g., estimated financial cost of alcohol use), and strategies (e.g., ways to moderate drinking) (Miller et al., 2013).

To our knowledge, only one study (Chiauzzi et al., 2008) has used personalized feedback as a stress management intervention, and it did not find feedback to be superior to a no-treatment control. This study provided text and graphical feedback about individuals' responses to stress questionnaires and linked the feedback to related coping strategies. However, the study did not test feedback as a stand-alone intervention, but instead provided a wide array of interventions (peer stories, frequently asked questions, interactive relaxation tools, time management tips,

sleep hygiene, and communication skills), raising the question of whether the personalized feedback was lost among the other components. Given the strong evidence in favor of personalized feedback interventions for problematic drinking among college students, it seems worthwhile to explore the potential efficacy of feedback for stress management in this population.

Collaborative Assessment in Psychotherapy

Several studies have tested the efficacy of providing individualized feedback as part of psychotherapy, specifically within the context of a framework called collaborative assessment (CA; Fischer, 2000) or therapeutic assessment (TA; Finn & Tonsager, 1997). CA/TA refers to the use of assessment procedures as therapeutic interventions, rather than as merely information-gathering endeavors. CA/TA is distinguished by collaborative goal setting between the client and the assessor and detailed, individualized feedback sessions in which the client's questions about the assessment and their results are answered. The active, systematic use of assessment feedback as therapy may provide a framework for how online interventions could be used to enhance face-to-face therapy. For example, assessment and feedback may be automated and conducted online, while the results may be incorporated into therapy planning and in-person sessions.

A meta-analysis found a significant overall effect (d = 0.40) for the therapeutic effects of CA over waitlist controls and assessment as usual (Hanson & Poston, 2011). A recent randomized controlled trial found that, among patients receiving specialized treatment for personality disorders, pre-treatment CA yielded medium effect sizes for improved outcome expectancies, perceptions of greater treatment progress, and higher treatment satisfaction, ds = 0.56 – 0.68 (De Saeger et al., 2014). Thus, within the context of an already-effective psychotherapy, CA may enhance therapeutic effects by improving process variables (engagement with treatment) or outcome variables (faster rate of symptom reduction). Within the stepped care model, CA may

improve the efficiency of the highest intensity interventions, existing face-to-face therapies, particularly to the degree that CA procedures are automated and web-based.

Present Study

The goal of the present study was to test the efficacy of an online stress intervention within the context of a stepped care model. More specifically, three interventions were tested, ranging from low-intensity/no therapist contact to higher-intensity: 1) an online self-monitoring intervention, consisting of a stress and coping daily diary; 2) the self-monitoring intervention combined with personalized weekly feedback; and 3) the self-monitoring intervention plus feedback combined with an in-person group stress management training. Additionally, data were collected from a matched control group for comparison. Each intervention is discussed in detail in the Methods section.

As data collection is ongoing for diary-only intervention group, only three of the four conditions were compared in the present manuscript: control, diary + feedback, and the in-person stress management training,

Hypotheses

I assessed intervention outcome in three primary domains: perceived stress, positive affect, and negative affect. For all outcome domains, I hypothesized that the two intervention conditions would show a significant pre-post decline in stress, and that this decline would be significant relative to the matched controls. I also hypothesized that the in-person stress management training would yield better outcomes than the diary + feedback group on all domains.

Method

Participants

Participant flow through the study is presented in the CONSORT diagrams in Figures 1a and b. Participants included 111 undergraduate students (71.2% female) at a large, public West Coast university who were at least 18 years old (*M* age = 19.24 years, *SD* = 2.07 years), fluent in

English, and scored 20 or higher on the Perceived Stress Scale (PSS; [1]. The majority of participants identified as either Asian/Asian-American (43.2%) or White/Caucasian (36.9%), similar to the demographic distribution of the overall undergraduate population at this university; 8.1% identified their race/ethnicity as Other, 4.5% identified as Black/African-American, 4.5% as Latino/Hispanic, and 0.9% as Native American/American Indian (see Table 1). Less than one-third of participants (29.7%) reported ever receiving psychotherapy in the past.

Participant recruitment and screening. All participants were undergraduate students enrolled in introductory level psychology classes who were earning extra credit through research participation. Students were recruited through the Psychology Subject Pool (PSP). Based on a PSS cutoff of 20 or above, approximately 45% of PSP participants were eligible for the study. The PSS cutoff was initially intended to capture the top 33% of the undergraduate population, based on data from a pilot sample of 74 undergraduates. However, the distribution of PSS scores in subsequent quarters, with larger samples, was skewed higher.

Recruitment and screening occurred in one of two ways, depending on condition and cohort. Participants were recruited over five academic quarters; each quarter represented a cohort.

Matched controls. Students in the PSP could sign up for a study on "college stress and coping," in which they completed questionnaires at three time points throughout the quarter. Students were not allowed to sign up if they were currently or had ever been enrolled in the intervention arm of the study (described below). Only participants with PSS scores ≥ 20 were included in the data analysis. Recruitment of the matched controls occurred primarily over the final two quarters of data collection.

Intervention conditions. For the intervention arm, recruitment methods changed mid-way through data collection due to policy changes in the PSP. In the initial two cohorts, students enrolled in the PSP could sign up for a study on "stress and coping in college life." They were emailed an online consent form (Information Statement), and, if they agreed to participate, were

redirected to the online screening questionnaires, which included demographics and the PSS. Eligible participants were invited to participate in the five-week-long intervention study; ineligible participants were given referrals for local counseling resources.

For the final three cohorts, participants were recruited through a large, online screening mechanism administered by the PSP. All introductory psychology students were given two-and-a-half weeks to complete a screening survey for multiple Psychology department studies; the survey included the PSS. Eligible participants were emailed an online consent form (Information Statement) and an invitation code for a five-week-long study on "stress and coping in college life." Interested participants were given one week to enroll. After enrollment, participants were asked to complete the remaining baseline questionnaires before beginning any of the interventions.

Across all four conditions, 1,917 students were screened and 855 were eligible (44.6%). Of those eligible, 117 (13.7%) enrolled and were randomized (if they were in the intervention arm of the study). In the present study, 111 enrolled for the three conditions that are being analyzed. Of the 111 enrolled, 90 (81.1%) received their allocated intervention and 75 (67.6%) completed the post-study assessment. The CONSORT diagrams (Figures 1a and b) shows participant flow through the matched control and intervention arms of the study.

Design, randomization, and power. Upon completion of the baseline questionnaires, the principal investigator randomized participants to a study condition using a random number generator (the Rv.Chisq(df) function in the Statistical Package for the Social Sciences, version 19 [SPSS]). Due to the increased time requirements of the SMT condition, participants were more heavily allocated to the SMT condition than the diary or feedback conditions, at a ratio of approximately 4:3:3, to ensure a sufficient number of participants completed the entire SMT intervention.

A priori power analyses were conducted using the GPOWER software application and were calculated for the potential to detect differences between group means on primary outcomes. On

the basis of previous collaborative assessment studies, I estimated that a sample of 160 for all four conditions, with maximum overall attrition of 30%, would yield adequate power (0.80) to detect effect sizes in the medium to large range (Cohen, 1992; f = .25-.40).

Procedures

Prior to completing baseline questionnaires, participants read an online consent form (Information Statement) detailing the study procedures, timeline, treatment interventions (if they were enrolled in the intervention arm), compensation, and risks/benefits. Those who agreed to participate were led to a secure online interface to complete a computerized battery of questionnaires. Participants were given 1 week to complete the questionnaires. Mid-study, post-study, and 3-month follow-up questionnaires were administered using the same online interface.

Once our study staff received the baseline questionnaires, participants in the intervention arm were notified via email of the condition to which they were randomized. The email detailed study procedures for that condition. Participants received a separate email with their daily diary login information and brief instructions on how to complete the first diary entry. Subsequently, all intervention participants were emailed a daily reminder with a link to the diary website. Participants who elected to be removed from this email list (n = 1) were not sent any reminders. All procedures were reviewed and approved by the local institutional review board.

Interventions

Daily diary. The present study's online daily diary has been developed and refined in previous studies on personality, stress, and coping (Shoda, Wilson, Chen, Gilmore, & Smith, 2013; Smith et al., 2011; Wilson, 2008). The daily diary queried participants about their general stress level and positive and negative emotions (using the PANAS-short form; Mackinnon et al., 1999) since their last diary entry. Participants provided a qualitative description of a stressful experience that occurred since their last diary entry; they also provided qualitative descriptions their stress-producing and stress-reducing thoughts. They rated the degree to which the

experience fit certain "psychologically active" situational features (Wilson, 2008) and reported on the use and helpfulness of a variety of coping behaviors, adapted from the Ways of Coping Checklist (Vitaliano, Russo, Carr, Maiuro, & Becker, 1985).

It was not uncommon for participants to report they had not experienced a stressful experience since their last diary entry. In these cases, participants were asked to report on any mildly negative situation. Alternatively, if they had no negative experience to report, or if they had a strongly positive experience that was emotionally significant, participants were invited to respond to the diary questions using this positive experience, with the goal of better understanding the thoughts and behaviors that might have made the experience positive. Study staff routinely checked participants' diary entries for data entry accuracy but did not read the content of the entries. Consequently, participants were informed they should contact study staff directly if they had any issues, questions, or concerns. Additionally, after each entry's submission, participants received a list of local clinics and counseling centers, allowing them to have information about mental health resources without being required to disclose anything.

Feedback. The content of the feedback was developed in pilot studies involving graduate-level therapists, who utilized the diary and constructed their own feedback for individual stress management clients (Shoda et al., 2013; R. E. Smith et al., 2011; Wilson, 2008). The principal investigator met with the therapists weekly during the course of treatment and afterwards compiled their analyses to identify commonalities in the results. From this information, a subset of feedback topics was developed to fit within the time frame of the study.

Feedback was sent on a weekly basis via a personalized email, beginning with the second week of the study. Each week, a trained research assistant exported data from the diary website and, following a written protocol book, used SPSS to analyze the data and export the results. The principal investigator checked a random subset of analyses for accuracy; no protocol deviations were found. Each individual's results were plugged into an email template and sent directly to the participant. The timeline for the feedback was as follows:

1. Week 2: Stressful situations
    a. Participants were told which three situational features were most strongly associated with stress based on their ideographic data.
2. Week 3: Normative feedback and emotions associated with stress
    a. Participants were shown a graph that compared their pre-study stress level to the average for their undergraduate peers.
    b. Participants were told which three emotions were most strongly associated with stress based on their ideographic data
3. Week 4: Most and least helpful/used coping strategies
    a. Participants were told which coping strategies were the most and least helpful to them, based on their ideographic data.
    b. Participants were told which coping strategies were the most and the least used by them, based on their ideographic data.
4. Week 5: General stress and life satisfaction over time
    a. Participants were shown a graph depicting their self-rated general stress level and life satisfaction over time

Stress management training. Cognitive Affective Stress Management Training (CASMT) is an empirically supported, cognitive behavioral program for stress reduction (Chen et al., under review; Crocker, Alderman, Murray, & Smith R., 1988; Holtzworth-Munroe, Munroe, & Smith, 1985; Rohsenow, Smith, & Johnson, 1985; Shoda et al., 2013; R. E. Smith & Rohsenow, 1987; R. E. Smith & Ascough, in press; R. E. Smith & Nye, 1989; R. E. Smith et al., 2011). CASMT focuses on increasing self- and emotion regulation skills through explicit teaching of cognitive and behavioral coping skills and in-session skills rehearsal. CASMT consists of six, one-hour sessions. In this study, CASMT sessions were led by either a master's-level clinician (the principal investigator) or a trained, post-baccalaureate clinician with extensive teaching experience. The topics covered in each of the six sessions were as follows:

1. Introduce cognitive-affective model of stress and teach somatic relaxation skills (deep breathing, progressive muscle relaxation).
2. Illustrate how cognitive appraisals mediate the relationship between situations and stress responses. Teach cognitive coping skills (self-instructional training, finding alternative or balanced thoughts).
3. Review cognitive coping skills and rehearse somatic coping skills through induced affect.
4. Teach integrated coping response (combining somatic and cognitive skills) and rehearse through induced affect; introduce the Benson meditation technique
5. Demonstrate cognitive defusion techniques and rehearse with induced affect.
6. Rehearse coping skills through induced affect; relapse prevention

While each session consisted of approximately 30-40 minutes of didactics and 20-30 minutes of skills rehearsal, a significant amount of coping skills acquisition was expected to occur between sessions. To that end, participants were given handouts that summarized each skill, and a copy was emailed to any participants who missed the session. Participants were allowed to miss no more than one session and remain in the study. A complete treatment manual is available upon request.

In this study, CASMT was adapted to include discussion of the daily diary and feedback, based on a collaborative assessment framework (S. E. Finn & Tonsager, 1997; Poston & Hanson, 2010). Adaptations included: providing participants with printouts of their individualized feedback emails; asking participants to comment on the accuracy and implications of the feedback; using stressful situations from the diary for induced affect; encouraging participants to volunteer their stress-producing thoughts from the diary as examples when teaching cognitive coping skills; and encouraging participants to share their effective stress-reducing thoughts from the diary with the group. The collaborative assessment goals were to promote each client's

understanding of their unique stress and coping patterns and to promote the trainer's understanding of each client's subjective experience.

Measures

Treatment acceptability.

Modified Assessment Questionnaire (MAQ; (S. E. Finn, Schroeder, & Tonsager, 1995). The AQ is a 48-item self-report questionnaire that assesses participants' reaction to collaborative assessment procedures; it has been used in a number of studies (Allen, Montgomery, Tubman, Frazier, & Escovar, 2003; Newman & Greenway, 1997). The AQ consists of four subscales: new self-awareness/understanding (e.g., "I gained a new understanding of myself"), positive accurate mirroring (e.g., "The assessment captured the 'real' me"), positive relationship with assessor (e.g., "I felt the assessor respected me"), and negative feelings about the assessment (e.g., "The assessment made me feel that my life is nothing but problems"). An overall satisfaction score is also obtained. Participants rate the degree to which they agree with each item on a 1 (strongly disagree) to 5 (strongly agree) scale. The AQ was found to have good internal consistency (α = .79 –.90) and two-week test-retest reliability (r =.75 –.81) (S. E. Finn et al., 1995).

For the purposes of this study, the AQ was modified to assess participants' reactions to the daily diary procedure. All items in subscale 3 (positive relationship with assessor) were removed because the daily diary does not involve one-on-one interactions with an assessor. The modified scale consisted of 36 items and showed good internal consistency (α = .86).

Treatment Credibility. Smith and Nye (1989) developed a treatment credibility measure for CASMT, which evaluates the program on four criteria: logic, confidence in the program, usefulness, and willingness to recommend it to a friend.

Stress and distress.

Perceived Stress Scale – 10 item version (PSS-10; Cohen et al., 1983). The PSS is a 10-item self-report questionnaire that measures the degree to which individuals felt overwhelmed and

that their lives were uncontrollable and unpredictable in the past month. The PSS has been shown to be a reliable measure of stress (α = .84-.86) that correlates with self-report measures of depression and anxiety (Cohen et al., 1983; Lee, 2012).

Positive and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). The PANAS is a 20-item self-report questionnaire that measures positive and negative affect. Negative affect is related to self-reported stress and poor coping (Bolger & Zuckerman, 1995; Watson, 1988; Zautra, Affleck, Tennen, Reich, & Davis, 2005), while positive affect is related to social activity and engagement (Clark & Watson, 1988). The PANAS shows good reliability (α =.84-.90; Watson et al., 1988).

Data Analytic Strategy

Data screening. Exploratory data analyses assessed whether the data satisfied the assumptions of normality and homogeneity of variance. When assumptions were violated, I either transformed the data or used an alternate method of analysis. Outliers were identified through visual inspection of boxplots and by examining studentized residuals, where outliers were defined by values ≥ +3 or -3. When significant outliers were found, the scores were corrected or removed if they were due to error; if the scores appeared to be valid, analyses were conducted with and without the outliers to examine their influence. If it was determined that the outliers were influential, both analyses were reported.

Intent-to-treat and per-protocol analyses. Consistent with established guidelines and recommendations for reporting clinical trials (Atkins, 2009; Des Jarlais, Lyles, & Crepaz, 2004), intent-to-treat (ITT) analyses were conducted for all primary outcome measures, meaning all randomized participants were included in the analyses regardless of adherence or attrition. ITT analyses attempt to correct for biases inherent in completer-only analyses, which may overestimate an intervention's effectiveness by limiting analyses to individuals who were willing to finish (Lachin, 2000; Salim, Mackinnon, Christensen, & Griffiths, 2008). ITT analyses

address the question of intervention effectiveness (as opposed to efficacy), that is, how does the intervention perform in routine practice, without perfect adherence?

However, ITT analyses perform best when complete data are available for all randomized participants (Hollis & Campbell, 1999), which is rarely, if ever, the case. In the present study, individuals who explicitly stated they wished to withdraw from the study were not contacted for follow-up, per Human Subjects protocol. Thus, I also conducted subset analyses using only individuals who completed the intervention. Such per-protocol (PP) analyses have merit in that they answer a different question from ITT analyses, namely, how does the intervention perform when administered as intended (Ten Have et al., 2008)?

Missing data. Missing data were handled using multiple imputation (MI; n = 10 imputations) in SPSS 19. Compared to missing data methods such as deletion, mean substitution, or the last observation carried forward, which can unduly influence the probability of Type I error, MI has been shown to produce fairly unbiased parameter estimates (Baraldi & Enders, 2010). MI assumes that data are missing at random (MAR), which means that missingness may be related to study variables but is not related to the individual's "true" value on the missing variable, if it had been collected. MAR is not a testable assumption, but some have suggested that, with a large enough sample size, MI may be fairly robust to departures from MAR (Schafer & Graham, 2002).

In the present study, MI was used to estimate data for participants who dropped out using all available data for perceived stress, positive and negative affect, baseline demographic characteristics, study condition, and treatment adherence (i.e., number of diary entries completed). MI creates multiple, complete datasets, each containing different estimates for the missing data, which are derived from regression analyses and random error terms. Analyses are run on all the imputed datasets, and then the results are pooled. Standard error estimates take into account between-imputation variance to reflect the uncertainty about the true value of the estimates.

Hypothesis testing. Mixed 3x3 (time x condition) analyses of variance (ANOVAs) were used to test the hypothesis that there would a significant difference between conditions in the mean change from baseline to mid-study to post-study on various outcome measures (perceived stress, negative affect, and positive affect). Stated another way, I expected the effect of time on these variables to depend upon condition. Follow-up pre-post *t*-tests were used to test the significance of within-group change for each condition.

Clinically significant change. In addition to the ANOVAs, clinical significance analyses were conducted for the primary outcome measure of perceived stress (PSS). As defined by Jacobson and Truax (1991), clinically significant change is both reliable and meaningful. Reliable change refers to a change from baseline to post-treatment that is unlikely to be explained by measurement error. It is assessed using the reliable change index (RCI), which is calculated using the reported test-retest reliability for the measure and the *SD* of the sample. Meaningful change is described as movement from the clinical to the non-clinical distribution on some measure of functioning. It can be operationalized as having a post-study score that falls 2 *SD* or more from the baseline mean in the direction of improved functionality (e.g., lower stress). Clinical significance analyses examine the percentage of the sample who meet criteria for reliable and meaningful change. The purpose of such analyses is to go beyond hypothesis testing and examine within-sample variability in intervention response and to assess whether statistically significant change is likely to be meaningful for the individual (Jacobson & Truax, 1991).

Results

Descriptive Information and Baseline Differences

Preliminary analyses revealed no significant differences in the outcome variables at baseline or post-study based on demographic characteristics (gender, treatment history, race/ethnicity), but did reveal a significant difference in baseline stress scores between cohorts, $F(3, 107) = 2.71$, $p = .05$, with post-hoc analyses revealing the significant difference was driven by participants

recruited prior to autumn quarter 2014 (n = 8), who reported lower baseline stress scores and less variability in stress (see Table 2). These participants were recruited differently from subsequent cohorts; thus, it is unclear whether the first cohort's participants are drawn from the same population as the rest of the study sample. Additionally, because of the difference in recruitment methods, these participants were able to participate in the intervention for one week longer than subsequent cohorts, thereby receiving a larger "dose" of the intervention. Given these differences, I ran analyses with and without these eight subjects. For simplicity, this section reports only results without these subjects; a footnote is inserted whenever a significant difference was found as a result of removing these subjects.

Table 4 provides descriptive statistics for the outcome measures of stress and distress by condition. No differences between conditions were detected in terms of initial stress severity, Welch's $F(2, 68.51) = 0.05$, $p = .95$, or negative affect, $F(2, 105) = 1.05$, $p = .35$. Differences between conditions in baseline positive affect approached significance, $F(2, 102) = 2.91$, $p = .06$, with post-hoc contrasts revealing the main source of difference was the SMT group reporting lower mean levels of baseline positive affect than the feedback group (M = -3.10, 95% CI, 0.16 to -6.35, $p = .07$).

Attrition and Missing Data

Of the 111 enrolled participants, 90 (81.1%) received their allocated intervention. There was no difference in baseline stress severity between completers and non-completers, $t(109) = 0.35$, $p = .73$. Of the 21 participants who did not receive their allocated intervention, three were in the feedback condition and 18 were in the SMT condition. Reasons for non-completion included: did not complete enough diary entries to receive feedback (n = 3, feedback condition), missed more than one stress management session (n = 12, SMT condition), reported a scheduling conflict or lack of time to attend sessions (n = 2, SMT condition), stated a desire to withdraw without elaborating a reason (n = 3, SMT condition), and withdrawing because the participant had already received the maximum extra credit allowed (n = 1, SMT condition).

As expected, attrition differed significantly by condition, $\chi^2(2) = 10.67$, $p = .005$, with the SMT condition showing higher rates of non-completion and the feedback condition showing lower rates of non-completion. This difference was anticipated as an artifact of the requirements of each condition, i.e., SMT participants had to attend at least five hours of in-person sessions to complete the study, whereas feedback participants needed to complete only three diary entries (approximately 30 minutes of participation) to receive feedback. Since participants did not earn extra credit for attending the SMT sessions, many who enrolled prior to randomization may not have intended to attend any in-person sessions. Consistent with this notion, 32.6% of participants randomized to the SMT condition never attended a session (Table 5).

Of the 90 who received their allocated intervention, 15 (16.7%) did not complete their post-treatment assessment despite several email reminders (control, $n = 8$; feedback, $n = 5$; SMT, $n = 2$). Several participants ($n = 11$) completed the post-treatment assessment but had incomplete data on one of the outcome measures due to item non-response. The rate of missingness for each outcome variable at each time point is shown in Table 3; the percentages are based on the intent-to-treat (ITT) sample.

Compliance

Daily diary. Table 5 provides means and standard deviations for the number diary entries completed by condition for both the intent-to-treat (ITT) and per-protocol (PP) samples. Across the feedback and SMT conditions, the mean number of diary entries was 14.87 ($SD = 10.54$) for the ITT sample, indicating that participants were, on average, completing about 1 diary entry every 2.15 days. There was wide variability in diary compliance (range = 0-34 entries). There was not a significant difference between feedback and SMT conditions in terms of diary entries completed, $t(74) = 0.86$, $p = .39$; feedback $M = 16.06$, $SD = 9.61$; SMT $M = 13.95$, $SD = 11.23$.

Given the large number of participants who dropped out of the SMT condition ($n = 18$, 41.9%), I also analyzed diary compliance for the per-protocol (PP) sample ($n = 55$; feedback $n = 30$; SMT $n = 25$). The mean number of diary entries for the PP sample was 18.98 ($SD = 9.11$),

and the range was 2-34 entries. There still was not a significant difference between feedback and SMT conditions in terms of diary entries completed, $t(53) = -1.30$, $p = .20$; feedback $M = 17.53$, $SD = 8.78$; SMT $M = 20.72$, $SD = 9.38$).

CASMT sessions. As shown in Table 5, 14 (32.6%) of participants randomized to SMT never attended a session, two (4.7%) participants attended one session, two (4.7%) participants attended two sessions, 15 (34.9%) attended five sessions, and 10 (23.3%) attended all six sessions.

Outcome Measures

Acceptability. Table 6 provides descriptive statistics for the outcome measures of acceptability by condition, as well as hypothesis testing results.

Diary acceptability (MAQ). MAQ scores range from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating greater agreement with the question content. Participants, on average, reported neutral to slightly positive opinions about the overall daily diary experience ($M = 3.26$, $SD = .33$) and about whether the diary promoted new self-awareness and understanding ($M = 3.25$, $SD = .52$). Participants were, on average, neutral to slightly negative about whether the daily diary verified their self-views and reflected their strengths ($M = 2.90$, $SD = .47$). At the same time, they reported few negative feelings about the diary procedure ($M = 2.38$, $SD = .70$). The only significant difference between the feedback and the SMT groups was in their reported overall satisfaction with the diary procedures, $t(41) = -2.13$, $p = .04$, $d = .73$, with SMT participants reporting a more positive experience overall (mean difference = 0.23, *SE* difference = .11). These results suggest that the daily diary was an acceptable intervention for a group of individuals reporting high stress levels. In addition, it appears that the in-person stress management intervention improved participants' experience of the diary assessment.

Treatment credibility. The treatment credibility measure ranged from 1 (not at all) to 10 (extremely), such that higher scores indicate greater agreement with the question content. SMT participants reported they found the treatment logical ($M = 6.95$, $SD = 1.81$) and useful ($M =$

6.47, *SD* = 1.92), had confidence in it (*M* = 6.89, *SD* = 1.76), and would be confident recommending it to a friend (*M* = 6.74, *SD* = 1.97). These results suggest that augmenting the existing evidence-based stress management program with collaborative assessment procedures did not negatively impact the treatment's credibility.

Feedback usefulness. Participants were asked to report how frequently they viewed the feedback, how much they remembered of the feedback, and how useful they found each type of feedback. The questions focused on only the first three weeks of feedback, as the last week of feedback coincided with the post-study questionnaires. Table 7 displays the average results by condition and for the total sample.

Stress and distress. Table 4 provides descriptive statistics for the primary outcome measures by condition.

Perceived stress (PSS). Mixed 3x3 ANOVAs were conducted to test the hypothesis that there would be a significant time x condition interaction when comparing baseline, mid-study, and post-study stress scores for the three conditions. Three outliers were identified and the data violated assumptions of normality; when these extreme scores were removed, the data met normality assumptions. Analyses with and without the outliers are reported below.

Observed data. When examining only the observed data (i.e., participants who completed the intervention and post-study assessment), the time x condition interaction was not significant, F (3.93, 121.75) = 0.54, p = .70, η^2 = .02. With the outliers removed, the F-statistic increased but was not significant, F (3.93, 117.87) = 0.85, p = .50, η^2 = .03.

ITT analysis. When all randomized participants were included in the analysis and missing data were addressed using multiple imputation, the time x condition interaction was significant, mean (over 10 imputations) F (3.56, 177.74) = 2.92, p = .05, η^2 = .05; F range = 1.74 – 5.94, p range = .001 – .15, η^2 range = .03 – .11. With the outliers removed, the significance of the interaction was more stable across imputation sets, mean F (3.50, 171.36) = 3.28, p = .03, η^2 =

.06; *F* range = 2.09 – 6.05, *p* range = .001 – .09, η^2 range = .04 – .11. Table 8 provides the results from all 10 imputations after the outliers were removed.

Post-hoc tests with the imputed datasets did not reveal any stable significant differences between conditions at a single time point, which may suggest that the conditions differed in their slopes, or trajectories of change, rather than in mean scores at a given assessment point. Additionally, the study may have been underpowered to detect between-group differences at a single time point. Figures 2a and b, which plot the observed and estimated marginal means of the PSS at each time point, suggest the SMT group may have experienced a more rapid decline in stress than the other two conditions. Table 9, which provides the effect sizes for all pairwise comparisons at mid-study and post-study, provides some support this hypothesis, with the SMT comparisons showing the largest effect sizes.

PP analysis. Results with the per-protocol (PP) sample led to similar conclusions. The 3x3 mixed ANOVA was not significant when the outliers were included in the model, mean *F* (3.90, 142.34) = 2.03, *p* = .10, η^2 = .05; *F* range = 1.59 – 2.39, *p* range = .05 – .18, η^2 range = .04 – .06. With the outliers removed, the *F*-statistic was significant, mean *F* (3.87, 137.48) = 2.56, *p* = .05, η^2 = .07; *F* range = 1.98 – 3.01, *p* range = .02 – .10, η^2 range = .05 – .08.[1] The direction of the interaction effects was the same.

Follow-up pre-post tests. I conducted follow-up pre-post *t*-tests using the observed data to examine the effects of each intervention when received according to protocol. All three conditions showed a significant decline in stress from baseline to post-study, with the SMT condition showing the largest effect size and, surprisingly, the feedback condition showing the smallest effect size (Table 10).

Daily diary adherence. I also examined the influence of treatment adherence on change in stress for the two intervention conditions, as participants varied widely in how often they used

[1] PP analyses that included the first cohort of eight participants were not significant: mean *F* (3.84, 151.49) = 1.37, *p* = .26, η^2 = .03; *F* range = 0.89 – 1.79, *p* range = .14 – .46, η^2 range = .02 – .04. With the outliers removed: mean *F* (3.87, 146.98) = 2.05, *p* = .10, η^2 = .05; *F* range = 1.60 – 2.60, *p* range = .04 – .18, η^2 range = .04 – .06.

the daily diary and such individual variation can obscure treatment effects. A 2 (condition) x 3 (time) mixed ANOVA with the total number of completed diary entries entered as a covariate revealed both significant time x diary, F (1.84, 122.99) = 12.35, p = .04, η^2 = .14; F range = 1.79 – 39.75, p range = <.0001 – .17, η^2 range = .03 – .37, and time x condition interactions, F (1.84, 122.99) = 5.81, p = .01, η^2 = .08; F range = 2.89 – 11.37, p range = .0002 – .06, η^2 range = .04 – .15. The results suggest that change in stress was influenced by diary adherence and that the intervention effects of feedback and SMT remained even after controlling for diary adherence.

I then examined the strength of association between diary adherence and change in stress separately for each condition. The correlation was not significant for the feedback condition, r = .15, p = .49, but was significant for the SMT condition, r = .52, p = .02, providing support for the hypothesis that the diary facilitated skills acquisition in the condition that was explicitly taught coping skills.

Figure 3 shows a scatterplot of the observed data, separated by condition, with change in PSS score (baseline minus post-study) plotted against the number of diary entries completed. Figures 4a and b plot the observed and estimated marginal means of the PSS, respectively, separated by time point and condition at the mean level of diary completion (18.97 entries).

Negative affect (PANAS). A mixed 3 (condition) x 3 (time) ANOVA conducted with the negative affect subscale of the PANAS did not indicate a significant main effect of time, F (1.86, 186.06) = 5.01, p = .16; F range = 0.03 – 13.49, p range = <.001 – .97, or a significant time x condition interaction effect, F (3.72, 186.06) = 1.16, p = .39; F range = 0.29 – 2.08, p range = .10 – .88. The results were similar when examining only the PP sample.

Positive affect (PANAS): ITT analysis. A mixed 3 (condition) x 3 (time) ANOVA conducted with the positive affect subscale of the PANAS revealed a main effect of time that approached significance, F (1.89, 189.29) = 7.90, p = .06; F range = 1.11 – 23.14, p range = <.001 – .33, but was highly variable across imputation sets. The model did not indicate a significant

time x condition interaction effect, F (3.79, 189.29) = 1.72, p = .30; F range = 0.47 – 3.23, p range = .02 – .76. The results were similar when examining only the PP sample.

Clinically significant change. Based on the formulas provided by Jacobson and Truax (1991), the cutoff score for meaningful change for this sample was < 17.25 on the PSS and the reliable change index (RCI) cutoff was >5.01. (See Figures 8a and b for the equations and calculations.) Of the 66 participants who completed the post-study assessment, 33 (50.0%) showed declines in perceived stress that constituted meaningful change (control: n=2 [8.3%]; feedback: n=3 [13.0%]; SMT: n=8 [42.1%]). While a chi-square goodness-of-fit test could not be conducted because some expected cell frequencies were <5, it seems that a much larger proportion of the SMT group showed meaningful change. Figure 5 displays the number of improved and not improved by condition.

Of the 66 participants who completed the post-study assessment, 17 (25.8%) showed reliable improvement (control: n=6 [25.0%]; feedback: n=4 [17.4%]; SMT: n=7 [36.8%]); no significant difference was found between groups, $\chi^2(2) = 0.82$, p = .82. Five participants (6.6%) showed reliable deterioration, defined as a change in the opposite direction towards dysfunctionality, i.e., increased stress (control: n=2 [8.3%]; feedback: n=2 [8.7%]; SMT: n=1 [5.3%]); a chi-square test could not be conducted because the expected cell frequencies were <5. Figure 6 displays the number of participants who showed reliable improvement, reliable deterioration, or no reliable change by condition.

Eleven participants (16.7%) showed clinically significant improvement, defined as meeting the criteria for both meaningful and reliable improvement (control: n = 2 [8.3%]; feedback: n=3 [13.0%]; SMT: n=6 [31.6%]); a chi-square test could not be conducted because some expected cell frequencies were <5. Given the correlation I found between change in stress and diary adherence, I conducted an independent samples Welch's t-test to examine whether those who showed clinically significant improvement (CSI) differed from those who did not (NCSI) in the number of diary entries completed. When examining only the intervention conditions, the

between-group difference approached significance, $t(31.3) = -1.92$, $p = .06$; CSI $M = 22.22$, $SD = 3.90$; NSCI $M = 18.33$, $SD = 8.94$. Those who reported clinically significant improvement completed, on average, four more diary entries and also showed less variability in diary adherence (CSI range: 17 – 28 entries; NCSI range: 2 – 34 entries). When including the control condition, where all participants completed zero diary entries, the groups differed significantly, $t(16.12) = -2.19$, $p = .04$; CSI $M = 18.18$, $SD = 9.64$; NSCI $M = 11.00$, $SD = 11.38$, providing support for the assumption that no diary access was equivalent to a placebo.

Discussion

The goal of the present study was to test the potential efficacy of an online stress management intervention, the daily diary; to develop and test a personalized feedback intervention for the diary; and to test the effectiveness of an empirically-supported, cognitive behavioral stress management program when augmented with collaborative assessment procedures. Based on the assumptions of a stepped-care model, I hypothesized that each more intensive intervention would lead to added benefits in terms of stress reduction, increased positive affect, and decreased negative affect.

Specifically, I hypothesized that merely monitoring one's stress and coping behaviors through the daily diary would increase self-awareness and promote behavior change, as evidenced by decreased stress and negative affect and increased positive affect. As data collection is ongoing for the daily diary condition, no firm conclusions can be drawn at this point about the potential efficacy of mere self-monitoring. Regarding the diary plus feedback condition, I hypothesized that personalized feedback would provide discrepant information in a more systematic and convincing form than mere self-monitoring and would lead to greater treatment effects compared to the control or diary conditions. Finally, I predicted that the CA-enhanced SMT group would show the largest treatment effects, with individuals benefitting not only from monitoring and feedback, but also from the didactic presentation and rehearsal of specific stress coping skills.

I began by examining treatment acceptability, as it is the logical cornerstone of any disseminable intervention. Both the diary plus feedback intervention and the CA-enhanced stress management program were acceptable interventions, which is consistent with previous studies that have successfully used these interventions in their protocols (e.g., (Rohsenow et al., 1985; Smith & Nye, 1989; Wilson, 2008). In-person SMT may have made the daily diary a more enjoyable and meaningful experience, which mirrors the results of existing studies of collaborative assessment (Smith, Eichler, Norman, & Smith, 2015). CA/TA theory predicts that clients benefit more from assessment when the results are discussed in a cooperative, reciprocal manner with an engaged provider; the in-person SMT was necessarily more reciprocal than the feedback condition and thus may have led to a better experience with the daily diary.

In terms of treatment outcome, both the diary plus feedback intervention and the CA-enhanced stress management program led to significant decreases in stress. Thirteen percent of the diary plus feedback condition and nearly one in three individuals in the SMT condition reported decreases in stress that met criteria for clinically significant change. When compared to the control condition, the SMT group showed a decrease in stress equivalent to a medium effect size (Cohen, 1992), and almost half of the SMT participants reported clinically meaningful stress reduction. There is strong evidence to suggest that the CA-enhanced group SMT program is an effective intervention for college students experiencing high levels of stress.

Additionally, diary adherence, as measured by the number of diary entries completed, was associated with change in stress from baseline to post-study, with more diary entries being correlated with greater decreases in stress. Individuals who achieved clinically significant change reported a minimum of 17 diary entries, or an average of one diary entry every 1.88 days. This suggests that diary adherence may have influenced individual variability in outcome within the intervention conditions. When examining the two intervention conditions separately, the association between diary adherence and change in stress was not significant for the feedback condition but was highly significant for the SMT condition. Combined with the result of higher

diary satisfaction among the SMT condition, the present study lends further support for the theory underlying CA, which is that assessment and feedback may enhance treatment outcome through improved therapeutic process, such as treatment engagement and buy-in.

However, the results also indicated that all three conditions showed a significant reduction in stress over the six-week study period, including the control group. Contrary to our hypothesis, the feedback intervention showed the smallest intervention effect size, even smaller than the control group, contradicting the anticipated positive effects of personalized feedback. Several RCTs of Internet-based stress management interventions have also found spontaneous improvement in the control group (Brown, Vanable, Carey, & Elin, 2011; Chiauzzi et al., 2008; Rose et al., 2013; Zetterqvist, Maanmies, Ström, & Andersson, 2003). Stress is, by definition, sensitive to changing environmental demands (Lazarus, 1991), thus fluctuations and regression to the mean may constitute the "normal" trajectory of change (Drozd, Raeder, Kraft, & Bjørkli, 2013). Consistent with this idea, the six-week test-retest reliability estimate for the PSS is .55 (Cohen et al., 1983). Given the possibility that all participants would have shown reductions in stress without any intervention, it is still notable that the SMT group showed lower stress scores at post-study than the other two conditions, with effect sizes in the medium (d = .41 – .45) range.

Also contrary to our hypotheses, there were no significant changes found for positive or negative affect. This is particularly surprising because previous studies have shown a correlation between negative affect and perceived stress (Watson, 1988). However, to the degree that positive and negative affect reflect individual differences in either personality (Letzring & Adamcik, 2015) or vulnerability to psychopathology (Watson, Clark, & Carey, 1988), it may not be surprising that a targeted stress intervention would not lead to significant changes on these outcomes within the time frame of the present study. Although I attempted to measure state, rather than trait, affect by asking participants how they have felt "in the past few days," it may be difficult to disentangle the two in practice.

Limitations

Given the large difference between the *p*-values for the ITT and PP perceived stress analyses, it seems clear that the present study was underpowered to detect between-condition differences, even those of a medium effect size. The fact that the two sets of analyses showed treatment effects in the same direction provided support that, with more participants in each cell, reliable significant time x condition interactions could have been obtained on a complete dataset.

An additional limitation of the present study, which could have influenced statistical power, was the brief duration of intervention period. In the dissertation proposal, the proposed study design was eight weeks long, with six weeks allotted to the interventions and one week each of baseline and follow-up data collection. Unfortunately, due to system-wide changes in the Subject Pool recruitment methods, which were not within our control, the study had to be condensed into six weeks. The interventions were compressed to five weeks and the post-study assessment period coincided with the final week of the diary collection. It is possible that abridging these interventions diluted their effects and hindered our ability to detect differences that would have developed over time.

One final, important limitation is the nature of this sample and their reasons for participation. Participants received extra credit for participation and there was a great deal of variability in individual engagement in the interventions. Even for individuals who appeared to have high adherence because they completed many diary entries, it is unknown if they did so because they found it helpful or because they were motivated to earn the maximum amount of extra credit. These limitations may seriously impair our ability to generalize these findings to a treatment-seeking population. On the other hand, the Psychology Subject Pool offers, on average, 50 or more studies each quarter. Participants had many research study options to choose from to earn their extra credit; the 13.7% of eligible participants who chose the present study might have been more strongly interested in or more willing to learn stress reduction techniques.

Future Directions

In future analyses, I plan to examine the relationship between daily coping and daily life stress. For example, papers published by our research group have computed adaptive and maladaptive coping indices, generated from the daily diary's Ways of Coping Checklist (Vitaliano et al., 1985), and correlated these indices with various outcomes for individual clients. For this study, I could compute a correlation between the coping indices and daily stress for each individual in the intervention conditions, then create estimates for the sampling distribution to test the hypothesis that the average correlation for the sample is significantly different from zero. Such analyses would elucidate how coping processes unfold in real-time, while taking into account individual variability and day-to-day fluctuations.

Another avenue for future analyses would be to examine potential moderators of treatment. A review by Taylor and Stanton (2005) identified stable individual differences, such as optimism, self-efficacy, and social support, as predictors of more effective coping. Baseline levels of optimism and social support, as measured by the College Adjustment Test (CAT; (Pennebaker, Colder, & Sharp, 1990), and of self-efficacy, as measured by the General Self-Efficacy Scale (GSES; (Schwarzer, Mueller, & Greenglass, 1999), could be examined as covariates in our outcome analyses.

One of the most compelling results was the association found between diary adherence and change in stress. Follow-up analyses could attempt to rule out the possibility that diary adherence was a proxy for other variables, such as personality variables (e.g., conscientiousness).

Finally, although the present study wan an efficient test of the stepped-care model of health care delivery, future dismantling studies could attempt to identify which intervention components and dosages produce the maximum effects. For example, SMT groups with and without CA could be compared, and the number and timing of sessions could be varied to examine dosage effects.

Conclusions

To our knowledge, this is the first study to use an online daily diary as a stress management intervention and the first to test ongoing collaborative assessment as an adjunct to evidence-based psychotherapy. Our results pointed to a significant association between diary adherence and stress reduction, and suggested that, when compared to a control group, CA-enhanced SMT produced moderate decreases in stress.

References

Allen, A., Montgomery, M., Tubman, J., Frazier, L., & Escovar, L. (2003). The effects of assessment feedback on rapport-building and self-enhancement processes. *Journal of Mental Health Counseling, 25*(3), 165-182.

American Psychological Association. (2015). *Stress in America: Paying with our health*. Washington, DC.

Arnberg, F. K., Linton, S. J., Hultcrantz, M., Heintz, E., & Jonsson, U. (2014). Internet-delivered psychological treatments for mood and anxiety disorders: A systematic review of their efficacy, safety, and cost-effectiveness. *Plos One, 9*(5)

Atkins, D. C. (2009). Clinical trials methodology: Randomization, intent ‐ to ‐ treat, and random ‐ effects regression. *Depression and Anxiety, 2*6(8), 697-700. doi:10.1002/da.20594

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior & Human Decision Processes, 50*(2), 248-287.

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 4*8(1), 5-37. doi:10.1016/j.jsp.2009.10.001

Bolger, N., & Zuckerman, A. (1995). A framework for studying personality in the stress process. *Journal of Personality & Social Psychology, 6*9(5), 890-902.

Bower, P., & Gilbody, S. (2005). Stepped care in psychological therapies: Access, effectiveness and efficiency. Narrative literature review. *The British Journal of Psychiatry, 18*6(1), 11-17. doi:10.1192/bjp.186.1.11

Brown, J., Vanable, P., Carey, M., & Elin, L. (2011). Computerized stress management training for HIV+ women: A pilot intervention study. *AIDS Care, 2*3(12), 1525-1532. doi:10.1080/09540121.2011.569699

Carver, C. S., & Scheier, M. (1998). *On the self-regulation of behavior*. Cambridge, UK: Cambridge University Press.

Cavanagh, K., Strauss, C., Cicconi, F., Griffiths, N., Wyper, A., & Jones, F. (2013). A randomised controlled trial of a brief online mindfulness-based intervention. *Behaviour Research and Therapy, 5*1(9), 573-578. doi:10.1016/j.brat.2013.06.003

Chen, J. A., Gilmore, A. K., Wilson, N. L., Smith, R. E., Quinn, K., Peterson, A. P., . . . Shoda, Y. (under review). Enhancing stress management coping skills using induced affect and collaborative daily assessment.

Chiauzzi, E., Brevard, J., Thurn, C., Decembrele, S., & Lord, S. (2008). MyStudentBody–Stress: An online stress management intervention for college students. *Journal of Health Communication, 1*3(6), 555-572. doi:10.1080/10810730802281668

Christensen, H., Batterham, P., & Calear, A. (2014). Online interventions for anxiety disorders. *Current Opinion in Psychiatry, 2*7(1), 7-13. doi:10.1097/YCO.0000000000000019

Clark, L. A., & Watson, D. (1988). Mood and the mundane: Relations between daily life events and self-reported mood. *Journal of Personality and Social Psychology Journal of Personality and Social Psychology, 5*4(2), 296-308.

Clarke, J., Proudfoot, J., Birch, M., Whitton, A. E., Parker, G., Manicavasagar, V., . . . Hadzi-Pavlovic, D. (2014). Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: Secondary analysis of a randomised controlled trial. *BMC Psychiatry, 14*

Cohen, J. (1992). A power primer. *Psychological Bulletin, 11*2(1), 155-159. doi:10.1037/0033-2909.112.1.155

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 2*4(4), 385-396.

Cowpertwait, L., & Clarke, D. (2013). Effectiveness of web-based psychological interventions for depression: A meta-analysis. *International Journal of Mental Health and Addiction, 1*1(2), 247-268. doi:10.1007/s11469-012-9416-z

Crocker, P. R. E., Alderman, R. B., Murray, F., & Smith R. (1988). Cognitive-affective stress management training with high performance youth volleyball players: Effects on affect, cognition, and performance. *Journal of Sport & Exercise Psychology, 1*0, 448-460.

Davies, E. B., Morriss, R., & Glazebrook, C. (2014). Computer-delivered and web-based interventions to improve depression, anxiety, and psychological well-being of university students: A systematic review and meta-analysis. *Journal of Medical Internet Research, 1*6(5), 18-39. doi:10.2196/jmir.3142

De Saeger, H., Kamphuis, J. H., Finn, S. E., Smith, J. D., Verheul, R., van Busschbach, J. J., . . . Horn, E. K. (2014). Therapeutic assessment promotes treatment readiness but does not affect symptom change in patients with personality disorders: Findings from a randomized clinical trial. *Psychological Assessment, 2*6(2), 474-483. doi:10.1037/a0035667

Des Jarlais, D., Lyles, C., & Crepaz, N. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health, 9*4(3), 361-366.

Dèttore, D., Pozza, A., & Andersson, G. (2015). Efficacy of technology-delivered cognitive behavioural therapy for OCD versus control conditions, and in comparison with therapist-administered CBT: Meta-analysis of randomized controlled trials. *Cognitive Behaviour Therapy, 4*4(3), 190-211. doi:10.1080/16506073.2015.1005660

Drozd, F., Raeder, S., Kraft, P., & Bjørkli, C. A. (2013). Multilevel growth curve analyses of treatment effects of a web-based intervention for stress reduction: Randomized controlled trial. *Journal of Medical Internet Research, 1*5(4), e84. doi:10.2196/jmir.2570

Febbraro, G. A. R., & Clum, G. A. (1998). Meta-analytic investigation of the effectiveness of self-regulatory components in the treatment of adult problem behaviors. *Clinical Psychology Review, 1*8(2), 143-161. doi:10.1016/S0272-7358(97)00008-1

Finn, S. E., Schroeder, D. G., & Tonsager, M. E. (1995). *The assessment questionnaire-2 (AQ-2): A measure of clients' experiences with psychological assessment. Unpublished Manuscript. Center for Therapeutic Assessment, Austin, TX.*

Finn, S. E., & Tonsager, M. E. (1997). Information-gathering and therapeutic models of assessment: Complementary paradigms. *Psychological Assessment,* 9(4), 374-385.

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clinical Psychology Review, 37*, 1-12. doi:10.1016/j.cpr.2015.01.006

Hanson, W. E., & Poston, J. M. (2011). Building confidence in psychological assessment as a therapeutic intervention: An empirically based reply to lilienfeld, garb, and wood (2011). *Psychological Assessment, 2*3(4), 1056-1062. doi:10.1037/a0025656

Hintz, S., Frazier, P. A., & Meredith, L. (2015). Evaluating an online stress management intervention for college students. *Journal of Counseling Psychology, 6*2(2), 137-147. doi:10.1037/cou0000014

Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials.(statistical data included). *British Medical Journal, 31*9(7211), 670.

Holtzworth-Munroe, A., Munroe, M. S., & Smith, R. E. (1985). Effects of a stress-management training program on first- and second-year medical students. *Journal of Medical Education, 6*0(5), 417-419.

Hölzel, B. K., Lazar, S. W., Gard, T., Schuman-Olivier, Z., Vago, D. R., & Ott, U. (2011). How does mindfulness meditation work? proposing mechanisms of action from a conceptual and neural perspective. *Perspectives on Psychological Science,* 6(6), 537-559. doi:10.1177/1745691611419671

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 5*9(1), 12-19. doi:10.1037/0022-006X.59.1.12

Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology, 4*4(1).

Lachin, J. M. (2000). Statistical considerations in the intent-to-treat principle. *Controlled Clinical Trials, 2*1(3), 167-189. doi:10.1016/S0197-2456(00)00046-5

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist, 4*6(8), 819-834. doi:10.1037/0003-066X.46.8.819

Lee, E. (2012). Review of the psychometric evidence of the perceived stress scale. *Asian Nursing Research,* 6(4), 121-127. doi:10.1016/j.anr.2012.08.004

Letzring, T. D., & Adamcik, L. A. (2015). Personality traits and affective states: Relationships with and without affect induction. *Personality and Individual Differences, 7*5, 114-120. doi:10.1016/j.paid.2014.11.011

Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Rodgers, B. (1999). A short form of the positive and negative affect schedule: Evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences, 2*7(3), 405-416. doi:10.1016/S0191-8869(98)00251-7

Miller, M. B., Leffingwell, T., Claborn, K., Meier, E., Walters, S., & Neighbors, C. (2013). Personalized feedback interventions for college alcohol misuse: An update of walters & neighbors (2005). *Psychology of Addictive Behaviors, 2*7(4), 909-920. doi:10.1037/a0031174

Newman, M. L., & Greenway, P. (1997). Therapeutic effects of providing MMPI-2 test feedback to clients at a university counseling service: A collaborative approach. *Psychological Assessment,* 9(2), 122-131. doi:10.1037/1040-3590.9.2.122

Pennebaker, J. W., Colder, M., & Sharp, L. K. (1990). Accelerating the coping process. *Journal of Personality and Social Psychology, 5*8(3), 528-537.

Pennebaker, J. W., & Seagal, J. D. (1999). Forming a story: The health benefits of narrative. *Journal of Clinical Psychology, 55*(10), 1243-1254.

Poston, J. M., & Hanson, W. E. (2010). Meta-analysis of psychological assessment as a therapeutic intervention. *Psychological Assessment, 2*2(2), 203-212.

Rice, S. M., Goodall, J., Hetrick, S. E., Parker, A. G., Gilbertson, T., Amminger, G. P., . . . Alvarez-Jimenez, M. (2014). Online and social networking interventions for the treatment of depression in young people: A systematic review. *Journal of Medical Internet Research, o 1*6(9), 190-200.

Rohsenow, D. J., Smith, R. E., & Johnson, S. (1985). Stress management training as a prevention program for heavy social drinkers: Cognitions, affect, drinking, and individual differences. *Addictive Behaviors, 1*0(1), 45-54. doi:10.1016/0306-4603(85)90052-8

Rose, R. D., Buckey, J. C. J., Zbozinek, T. D., Motivala, S. J., Glenn, D. E., Cartreine, J. A., & Craske, M. G. (2013). A randomized controlled trial of a self-guided, multimedia, stress management and resilience training program. *Behaviour Research and Therapy, 5*1(2), 106-112. doi:10.1016/j.brat.2012.11.003

Salim, A., Mackinnon, A., Christensen, H., & Griffiths, K. (2008). Comparison of data analysis strategies for intent-to-treat analysis in pre-test–post-test designs with substantial dropout rates. *Psychiatry Research, 16*0(3), 335-345. doi:10.1016/j.psychres.2007.08.005

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods,* 7(2), 147-177. doi:10.1037/1082-989X.7.2.147

Schneiderman, N., Ironson, G., & Siegel, S. D. (2005). Stress and health: Psychological, behavioral, and biological determinants. *Annual Review of Clinical Psychology,* 1(1), 607-628. doi:10.1146/annurev.clinpsy.1.102803.144141

Schwarzer, R., Mueller, J., & Greenglass, E. (1999). Assessment of perceived general self-efficacy on the internet: Data collection in cyberspace. *Anxiety, Stress & Coping, 1*2(2), 145-161.

Shoda, Y., Wilson, N. L., Chen, J., Gilmore, A. K., & Smith, R. E. (2013). Cognitive-affective processing system analysis of intra-individual dynamics in collaborative therapeutic assessment: Translating basic theory and research into clinical applications. *Journal of Personality, 8*1(6), 554-568.

Sloan, D. M., Gallagher, M. W., Feinstein, B. A., Lee, D. J., & Pruneau, G. M. (2011). Efficacy of telehealth treatments for posttraumatic stress-related symptoms: A meta-analysis. *Cognitive Behaviour Therapy, 4*0(2), 111-125. doi:10.1080/16506073.2010.550058

Smith, R. E., & Ascough, J. C. (Eds.). (in press). *Affect elicitation in psychotherapy and stress management training*. New York: Guilford Press.

Smith, R. E., & Rohsenow, D. J. (1987). Cognitive-affective stress management training: A treatment and resource manual. . *Social and Behavioral Sciences Documents, 17, Document #2829*

Smith, J. D., Eichler, W. C., Norman, K. R., & Smith, S. R. (2015). The effectiveness of collaborative/therapeutic assessment for psychotherapy consultation: A pragmatic replicated single-case study. *Journal of Personality Assessment, 97*(3), 261-270. doi:10.1080/00223891.2014.955917

Smith, R. E., Fagan, C., Wilson, N. L., Chen, J., Corona, M., Nguyen, H., . . . Shoda, Y. (2011). Internet-based approaches to collaborative therapeutic assessment: New opportunities for professional psychologists. *Professional Psychology: Research and Practice, 42*(6), 494-504. doi:10.1037/a0025392

Smith, R. E., & Nye, S. L. (1989). Comparison of induced affect and covert rehearsal in the acquisition of stress management coping skills. *Journal of Counseling Psychology, 36*(1), 17-23.

Ten Have, T. R., Normand, S. T., Marcus, S. M., Brown, C. H., Lavori, P., & Duan, N. (2008). Intent-to-treat vs. non-intent-to-treat analyses under treatment non-adherence in mental health randomized trials. *Psychiatric Annals, 38*(12), 772-783. doi:10.3928/00485713-20081201-10

Vago, D. R., & Silbersweig, D. A. (2012). Self-awareness, self-regulation, and self-transcendence (S-ART): A framework for understanding the neurobiological mechanisms of mindfulness. *Frontiers in Human Neuroscience,* 6 doi:10.3389/fnhum.2012.00296

Vitaliano, P. P., Russo, J., Carr, J. E., Maiuro, R. D., & Becker, J. (1985). The ways of coping checklist: Revision and psychometric properties. *Multivariate Behavioral Research, 20*(1), 3-26.

Walters, S. T., & Neighbors, C. (2005). Feedback interventions for college alcohol misuse: What, why and for whom*? Addictive Behaviors, 3*0(6), 1168-1182. doi:10.1016/j.addbeh.2004.12.005

Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affect: Their relation to health complaints, perceived stress, and daily activities. *Journal of Personality and Social Psychology, 54*(6), 1020-1030. doi:10.1037/0022-3514.54.6.1020

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology, 97*(3), 346-353. doi:10.1037/0021-843X.97.3.346

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology Journal of Personality and Social Psychology, 54*(6), 1063-1070.

Wilson, N. (2008). *Identifying the features of stressful situations.* Doctoral dissertation.

Zautra, A. J., Affleck, G. G., Tennen, H., Reich, J. W., & Davis, M. C. (2005). Dynamic approaches to emotions and stress in everyday life: Bolger and zuckerman reloaded with positive as well as negative affects. *Journal of Personality, 73*(6), 1511-1538.

Zetterqvist, K., Maanmies, J., Ström, L., & Andersson, G. (2003). Randomized controlled trial of internet-based stress management. *Cognitive Behaviour Therapy, 32*(3), 151-160. doi:10.1080/16506070302316

Table 1.
Participants' Sociodemographic Characteristics and Treatment History at Baseline: Intent-to-Treat (As-Randomized) Sample.

Variable	Control (n = 35)	Diary + Feedback (n = 33)	Stress Management (n = 43)	Full Sample (N = 111)
Gender, *n* (% female)	29 (82.9)	19 (57.6)	31 (72.1)	79 (71.2)
Age, *M* (*SD*), years	19.46 (2.97)	18.85 (1.56)	19.36 (1.41)	19.24 (2.07)
Ethnicity, *n* (% of sample)				
Asian/Asian-American	17 (48.6)	12 (36.4)	19 (44.2)	48 (43.2)
White/Caucasian	11 (31.4)	16 (48.5)	14 (32.6)	41 (36.9)
Other	4 (11.4)	1 (3.0)	4 (9.3)	9 (8.1)
Black/African-American	1 (2.9)	1 (3.0)	3 (7.0)	5 (4.5)
Latino/Hispanic	2 (5.7)	2 (6.1)	1 (2.3)	5 (4.5)
Native American	0 (0)	0 (0)	1 (2.3)	1 (0.9)
Ever received psychotherapy, *n* (% yes)	12 (34.3)	7 (21.2)	14 (32.6)	33 (29.7)
Past therapy type, *n* (% of sample)				
Cognitive-behavioral	5 (14.3)	5 (15.2)	7 (16.3)	17 (15.3)
Other individual	4 (11.4)	4 (12.1)	6 (14.0)	14 (12.6)
Family or couple	4 (11.4)	0 (0)	3 (7.0)	7 (6.3)
Group therapy	1 (2.9)	1 (3.0)	2 (4.7)	4 (3.6)
Length of therapy, *n* (% of sample)				
<1 month	7 (20.0)	2 (6.1)	2 (4.7)	11 (9.9)
1–6 months	1 (2.9)	3 (9.1)	3 (7.0)	7 (6.3)
6 months – 1 years	1 (2.9)	0 (0)	1 (2.3)	2 (1.8)
1–3 years	2 (5.7)	2 (6.1)	5 (11.6)	9 (8.1)
3–6 years	0 (0)	0 (0)	1 (2.3)	1 (.9)
>6 years	0 (0)	0 (0)	1 (2.3)	1 (.0)
Helpfulness of past therapy, *M* (*SD*)[a]	3.36 (1.80)	3.86 (2.27)	5.00 (2.00)	4.16 (2.07)

[a]Helpfulness of past therapy was measured on a 1 (not at all helpful) to 7 (very helpful) scale. The number of participants who responded to this item were as follows: control, n = 11; diary + feedback, n = 7; stress management, n = 13; full sample, n = 31.

Table 2.
Descriptive Statistics for Baseline Stress Measure by Cohort

Variable (scale range) and cohort	*M*	*SD*	*n*
Perceived Stress Scale (0-40)			
Spring & Summer 2014[a]	21.50	1.69	8
Autumn 2014	25.94	4.68	17
Winter 2015	24.79	3.41	48
Spring 2015	24.97	3.80	38

[a]The first two quarters of recruitment were combined because of the small cell size in the summer 2014 group (n = 1) and the similarity in recruitment strategy.

Table 3.
Numbers of Observed and Missing Cases and Rate of Missingness for Outcome Measures by Condition and Time Point

	Pre			Mid			Post		
Variable and condition	*n* obs.	*n* miss.	%	*n* obs.	*n* miss.	%	*n* obs.	*n* miss.	%
Perceived Stress Scale									
Control	35	0	0	30	5	14.3	26	9	25.7
Diary + Feedback	33	0	0	27	6	18.2	24	9	27.3
Stress Management	43	0	0	24	19	44.2	22	21	48.8
PANAS Negative Affect									
Control	35	0	0	30	5	14.3	25	10	28.6
Diary + Feedback	32	1	3.0	28	5	15.2	22	9	27.3
Stress Management	41	2	4.7	25	18	41.9	24	19	44.2
PANAS Positive Affect									
Control	33	2	5.7	31	4	11.4	26	9	25.7
Diary + Feedback	30	3	9.1	28	5	15.2	24	11	33.3
Stress Management	42	1	2.3	25	18	41.9	23	20	46.5

Table 4.
Descriptive Statistics for Outcome Measures of Stress and Distress by Condition

Variable (scale range) and condition	Pre			Mid			Post		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Perceived Stress Scale (0-40)									
Control	24.94	2.67	33	22.93	5.21	28	22.42	5.65	24
Diary + Feedback	24.72	4.05	32	24.19	5.09	27	22.35	6.39	23
Stress Management	25.42	4.37	38	22.29	4.58	21	19.95	5.24	19
PANAS Negative Affect (10-50)									
Control	25.97	5.46	33	27.07	5.13	28	25.96	4.33	23
Diary + Feedback	25.16	6.21	31	24.79	7.13	28	24.00	7.56	23
Stress Management	24.14	5.36	36	23.45	5.40	22	22.50	5.16	20
PANAS Positive Affect (10-50)									
Control	29.65	6.00	31	29.14	5.41	29	29.00	4.96	24
Diary + Feedback	30.62	5.25	29	28.14	5.05	28	28.29	6.02	21
Stress Management	27.03	5.60	37	26.45	5.70	22	26.16	5.66	19

Table 5.
Intervention Compliance by Condition for the Intent-to-Treat (ITT) and Per-Protocol (PP) Samples.

	ITT Sample			PP Sample		
Variable	Diary + Feedback (n = 33)	Stress Management (n = 43)	Total Sample (n = 76)	Diary + Feedback (n = 30)	Stress Management (n = 25)	Total Sample (n = 55)
Number of diary entries completed, *M* (*SD*)	16.06 (9.61)	13.95 (11.23)	14.87 (10.54)	17.53 (8.78)	20.72 (9.38)	18.98 (9.11)
Stress management sessions attended, *n*, (% of sample)						
0 sessions	-	14 (32.6)	-	-	-	-
1 session	-	2 (4.7)	-	-	-	-
2 sessions	-	2 (4.7)	-	-	-	-
5 sessions	-	15 (34.9)	-	-	15 (60.0)	-
6 sessions	-	10 (23.3)	-	-	10 (40.0)	-

Table 6.
Descriptive Statistics and Hypothesis Testing Results for Outcome Measures of Acceptability by Condition

	Diary + Feedback			Stress Management			Total Sample			*t* or *U* test			
Variable (scale range)	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*t* or *U*	*df*	*p*	*d*
Modified AQ[a] (1 – 5)													
Overall satisfaction	3.16	.36	20	3.38	.25	16	3.26	.33	36	2.13	34	.04*	.73
New self-awareness	3.11	.59	22	3.44	.36	17	3.25	.52	39	130.5	-	.11	-
Positive accurate mirroring	2.80	.58	21	2.80	.27	19	2.90	.47	40	181.5	-	.63	-
Negative feelings	2.46	.75	22	2.46	.64	17	2.38	.70	39	-1.22	44	.23	-
Treatment Credibility[b] (1 – 10)													
Overall	-	-	-	6.76	1.37	19	-	-	-	-	-	-	-
Logic of the program	-	-	-	6.95	1.81	19	-	-	-	-	-	-	-
Usefulness of the program	-	-	-	6.47	1.93	19	-	-	-	-	-	-	-
Confidence in the program	-	-	-	6.89	1.76	19	-	-	-	-	-	-	-
Would recommend to a friend	-	-	-	6.74	1.97	19	-	-	-	-	-	-	-

[a]Higher scores indicate greater agreement with the question content, ranging from 1 (strongly disagree) to 5 (strongly agree). Finn and colleagues published AQ norms from their Center for Therapeutic Assessment (N = 300) and reported the following: overall satisfaction, M = 3.6, SD = 1.1; new self-awareness, M = 3.8, SD = 1.1; positive mirroring, M = 3.5, SD = 1.5; negative feelings, M = 1.7, SD = 1.8. The authors note that their clients tend to be quite satisfied with their assessments and it is rare for samples to score substantially higher.
[b]Higher scores indicate greater satisfaction with and belief in the treatment program.

Table 7.
Descriptive Statistics for the Use and Helpfulness of the Feedback.

Variable	Diary + Feedback (n = 24)	Stress Management (n =20)	Total (N = 44)
Times feedback was viewed n (%)			
Never	0 (0)	1 (5.0)	1 (2.3)
Once	2 (8.3)	6 (30.0)	8 (18.2)
At least twice	13 (54.2)	5 (25.0)	18 (40.9)
Every time	9 (37.5)	8 (40.0)	17 (38.6)
Feedback usefulness[a], *M* (*SD*)			
1. Normative feedback[b]	4.59 (1.33)	4.50 (1.51)	4.55 (1.38)
2. Stressful situations[c]	4.40 (1.78)	4.50 (1.34)	4.44 (1.58)
3. Emotions and stress[d]	4.42 (1.35)	4.93 (1.33)	4.65 (1.35)
Did not remember feedback, n (%)			
1. Normative feedback	4 (19.0)	5 (29.4)	9 (23.7)
2. Stressful situations	1 (4.8)	3 (17.6)	4 (10.5)
3. Emotions and stress	2 (9.5)	2 (11.8)	4 (10.5)

[a]The usefulness of feedback was measured on a 1 (not at all useful) to 7 (very useful) scale.
[b]Normative feedback refers to the graph comparing the individual's stress level to the norm for the undergraduate population.
[c]The stressful situations feedback refers to a list of situation types that were most strongly associated with stress for that individual.
[d]The emotions and stress feedback refers to a list of emotions that were most strongly associated with stress for that individual.
[e]The coping strategies feedback refers to a list of the individual's most and least used and most and least helpful coping strategies.

Table 8. Mixed 3x3 ANOVA results across 10 imputations for the outcome of perceived stress (PSS).

Imputation Number and Source	*df*	*F*	*p*	partial η^2
1. Time	1.92	8.74	<0.001	0.08
Time x condition	3.84	3.06	0.02	0.06
Error	188.21			
2. Time	1.71	16.35	<0.001	0.14
Time x condition	3.41	3.78	0.01	0.07
Error	167.21			
3. Time	1.53	7.25	0.003	0.07
Time x condition	3.07	3.23	0.02	0.06
Error	150.40			
4. Time	1.72	8.60	<0.001	0.08
Time x condition	3.45	3.47	0.01	0.07
Error	168.94			
5. Time	1.89	11.30	<0.001	0.10
Time x condition	3.79	2.31	0.06	0.04
Error	185.57			
6. Time	1.90	8.89	<0.001	0.08
Time x condition	3.80	2.67	0.04	0.05
Error	186.43			
7. Time	1.83	16.62	<0.001	0.14
Time x condition	3.67	3.01	0.02	0.06
Error	179.60			
8. Time	1.31	40.96	<0.001	0.29
Time x condition	2.61	6.05	0.00	0.11
Error	128.05			
9. Time	1.72	25.87	<0.001	0.21
Time x condition	3.43	3.16	0.02	0.06
Error	168.29			
10. Time	1.95	10.74	<0.001	0.10
Time x condition	3.90	2.09	0.09	0.04

Error	190.87

Note: Three outliers were removed for this analysis. The reported statistics use the more conservative Greenhouse-Geisser correction, which adjusts the *df* and *p* statistics to account for violations in the assumption of sphericity.

Table 9.
Between-Group Effect Sizes at Mid-Study and Post-Study for Perceived Stress (PSS), Observed Data

Group Comparison	T2 between-group *d* [95% CI]	T3 between-group *d* [95% CI]
Control vs. Feedback	-.24 [-.77, .29]	.01 [-.56, .58]
Feedback vs. SMT	.39 [-.19, .96]	.41 [-.21, 1.01]
Control vs. SMT	.13 [-.44, .69]	.45 [-.17, 1.05]

Note. T2 = Mid-Study; T3 = Post-Study. Effect sizes were computed using the dataset in which the first cohort (n = 8) was removed.

Table 10.
Follow-Up Pre-Post t-Tests for Perceived Stress (PSS) by Condition, Observed Data

Condition	*M* difference	*SE* difference	95% CI of difference	*t*	*df*	*p*	*d*
Control	2.67	1.06	.46, 4.87	2.51	23	.02	.59
Diary + Feedback	2.43	1.26	-.18, 5.05	1.93	22	.07	.42
SMT	5.21	1.40	2.27, 8.15	3.72	18	.002	.87

Note. Analyses were computed using the dataset in which the first cohort (n = 8) was removed.

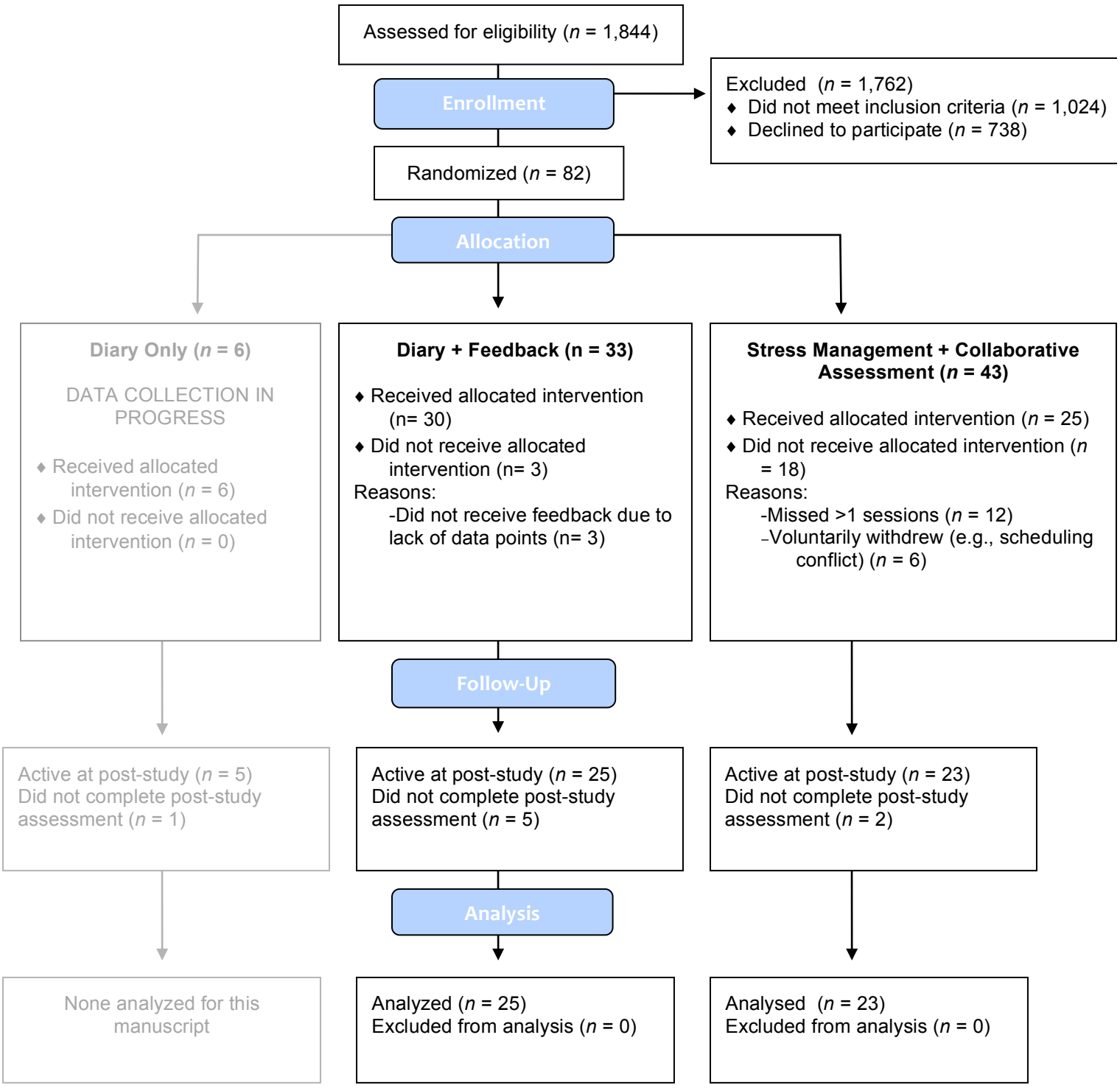

Figure 1a. CONSORT diagram for three intervention conditions.

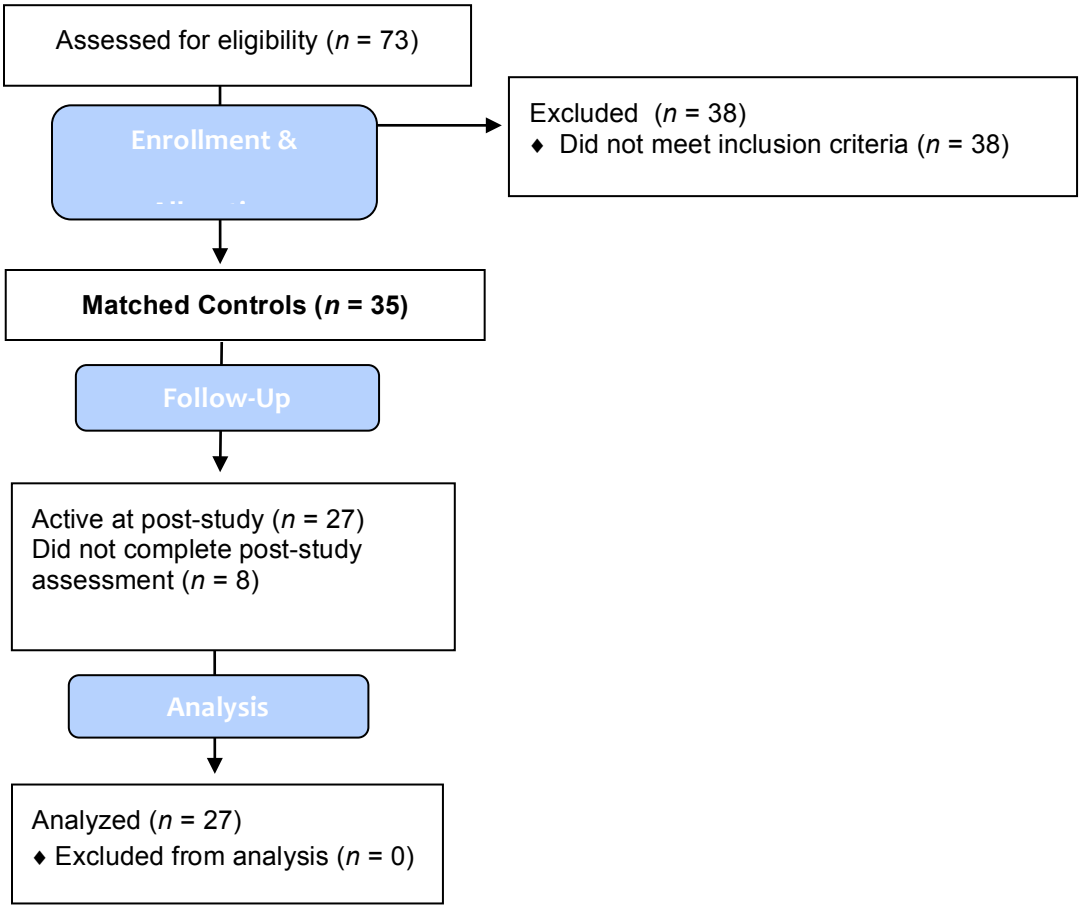

Figure 1b. CONSORT diagram for control condition.

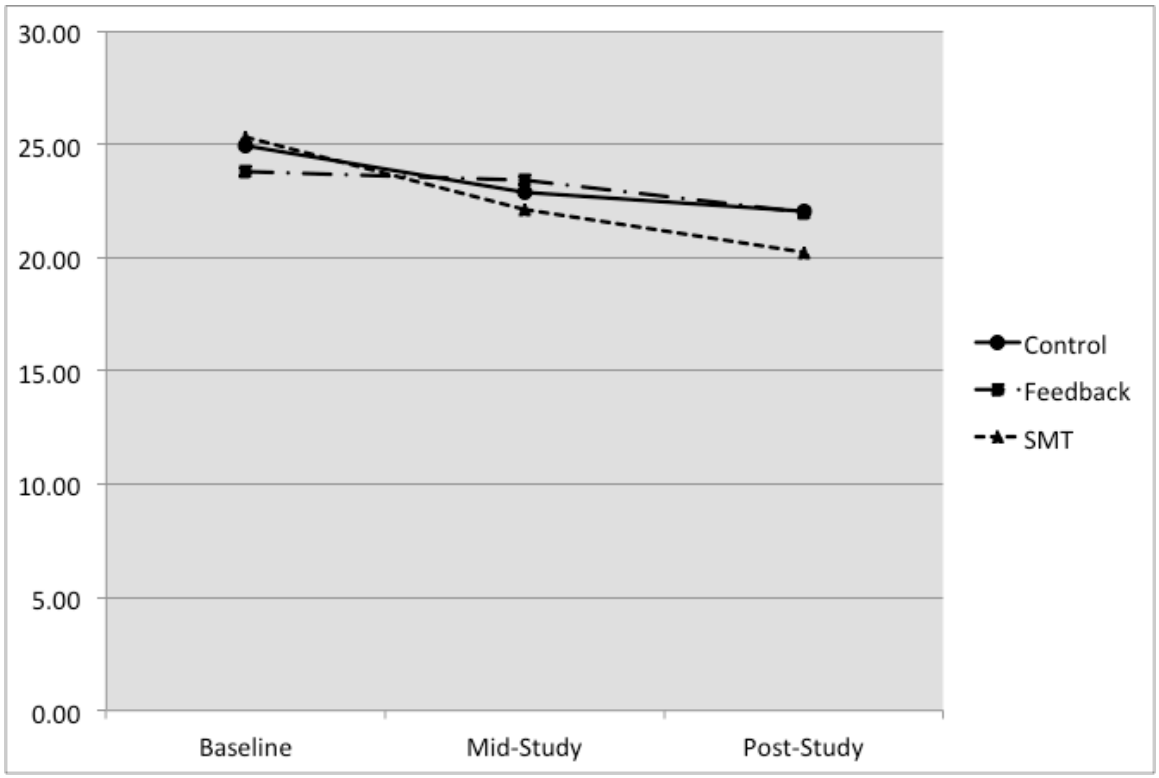

Figure 2a. Observed PSS scores at baseline, mid-study, and post-study for all three conditions.

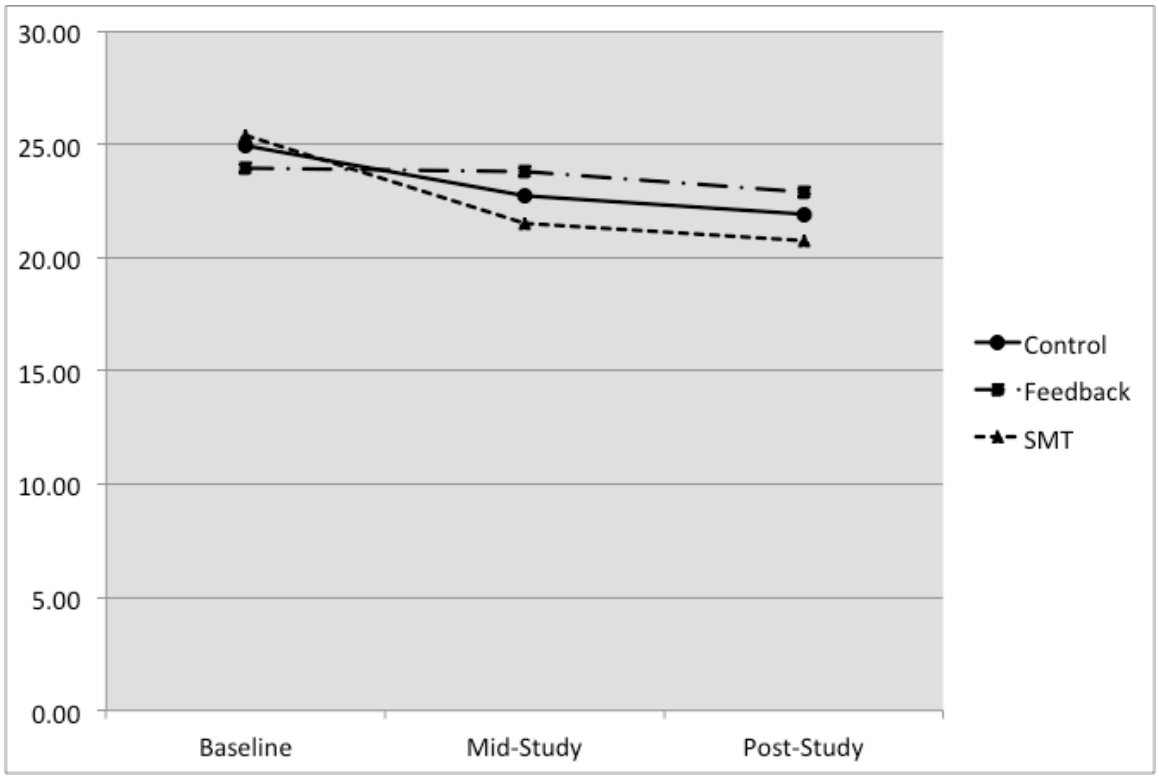

Figure 2b. Estimated marginal means, averaged across 10 imputations, of the PSS at baseline, mid-study, and post-study, separated by condition.

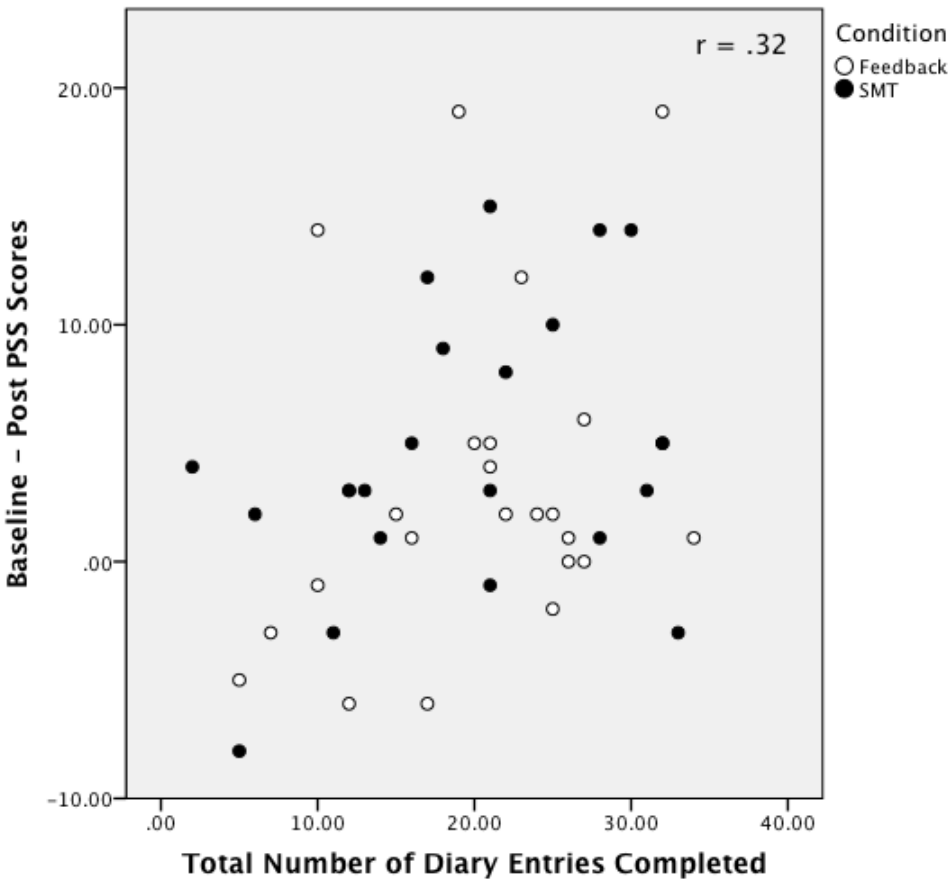

Figure 3. Change in PSS score (baseline minus post-study) plotted against the total number of diary entries completed, separated by condition (r = .32).

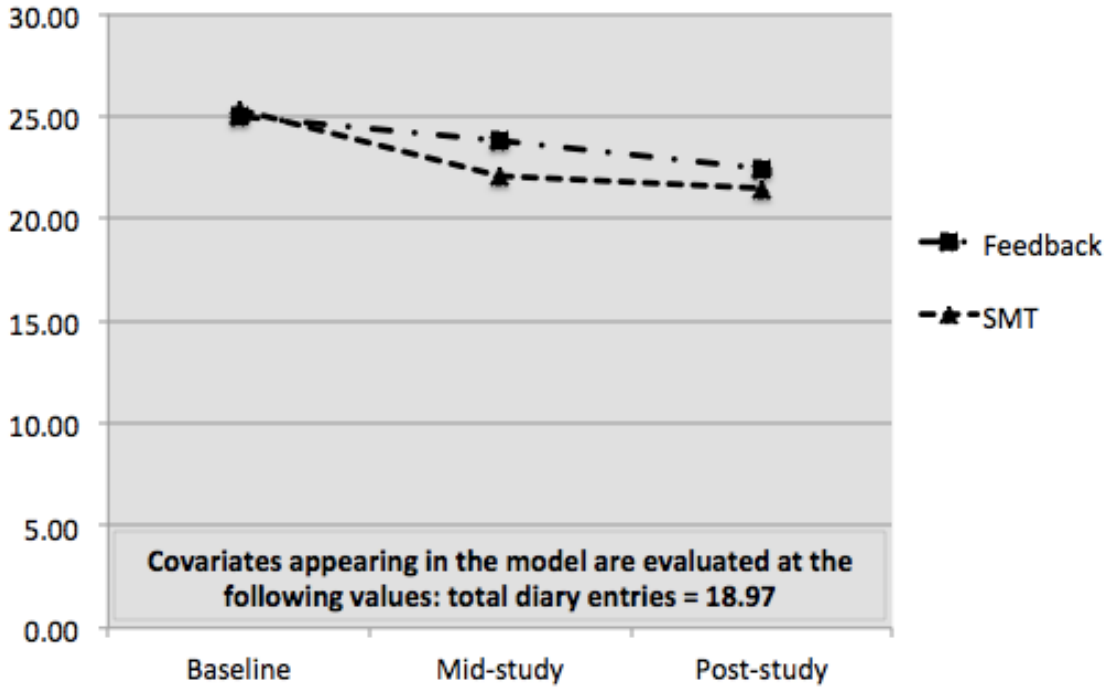

Figure 4a. Observed PSS scores at baseline, mid-study, and post-study for the two intervention conditions after controlling for diary adherence.

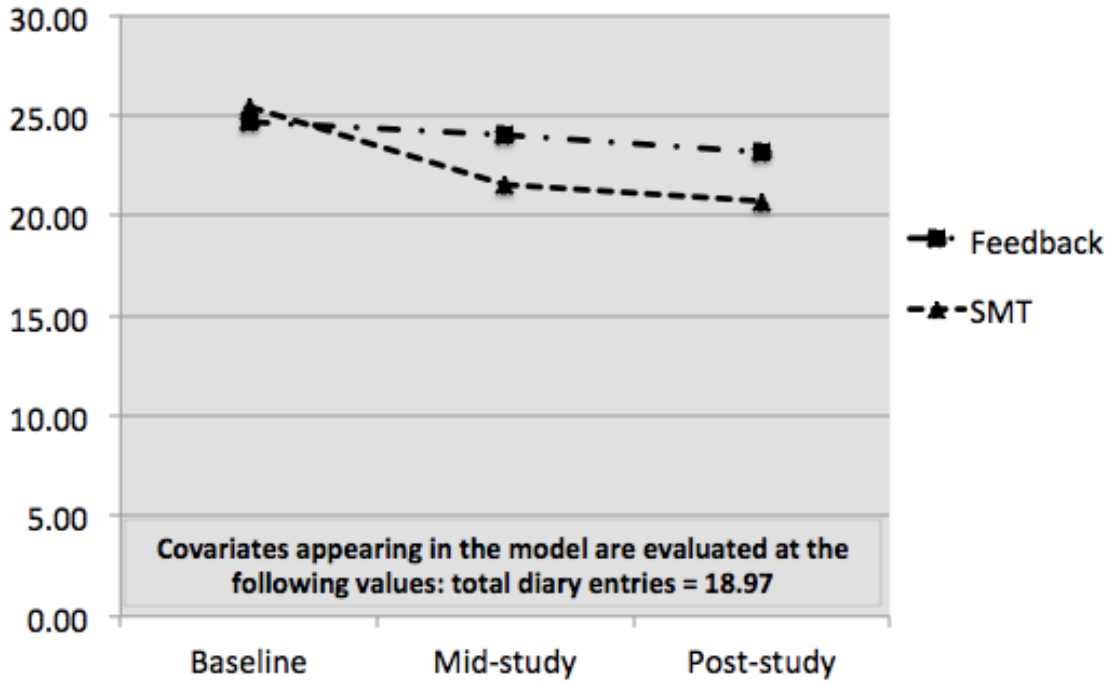

Figure 4b. Estimated marginal means, averaged across 10 imputations, of the PSS at baseline, mid-study, and post-study, separated by condition and controlling for diary adherence.

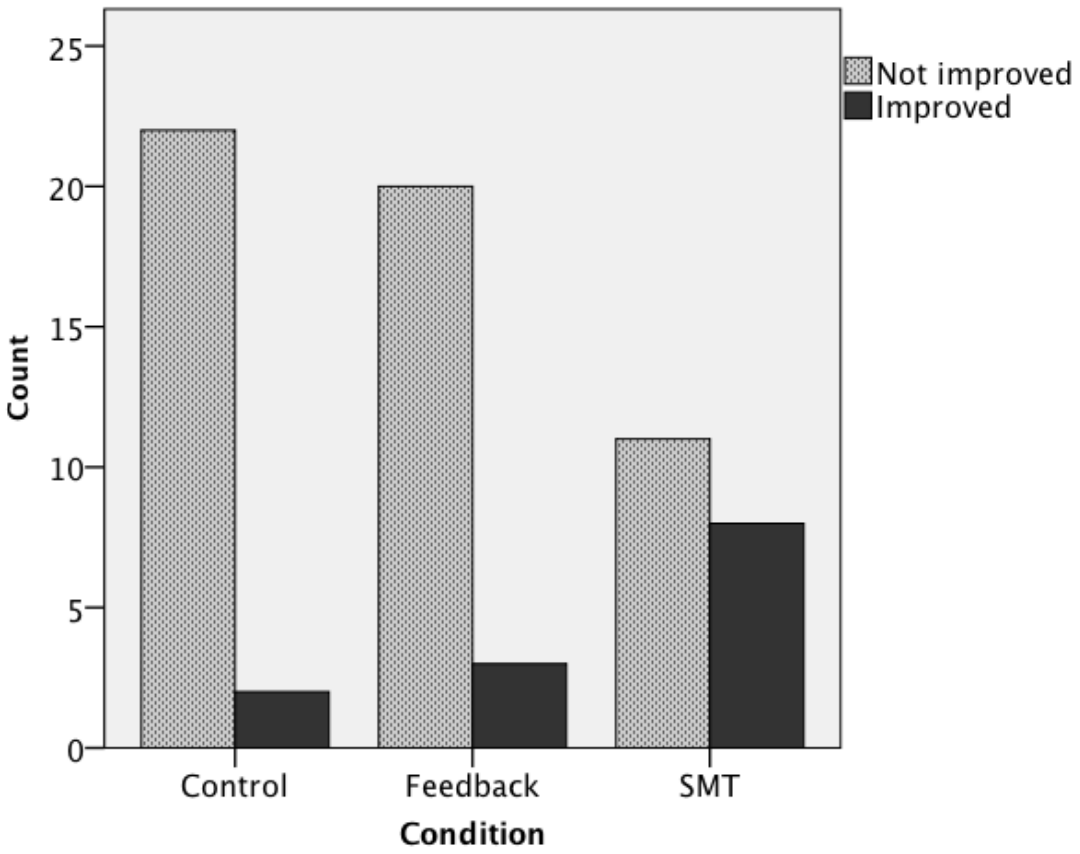

Figure 5. Number of participants who achieved clinically meaningful change, defined as a > 2 *SD* decrease in perceived stress from baseline to post study, by condition.

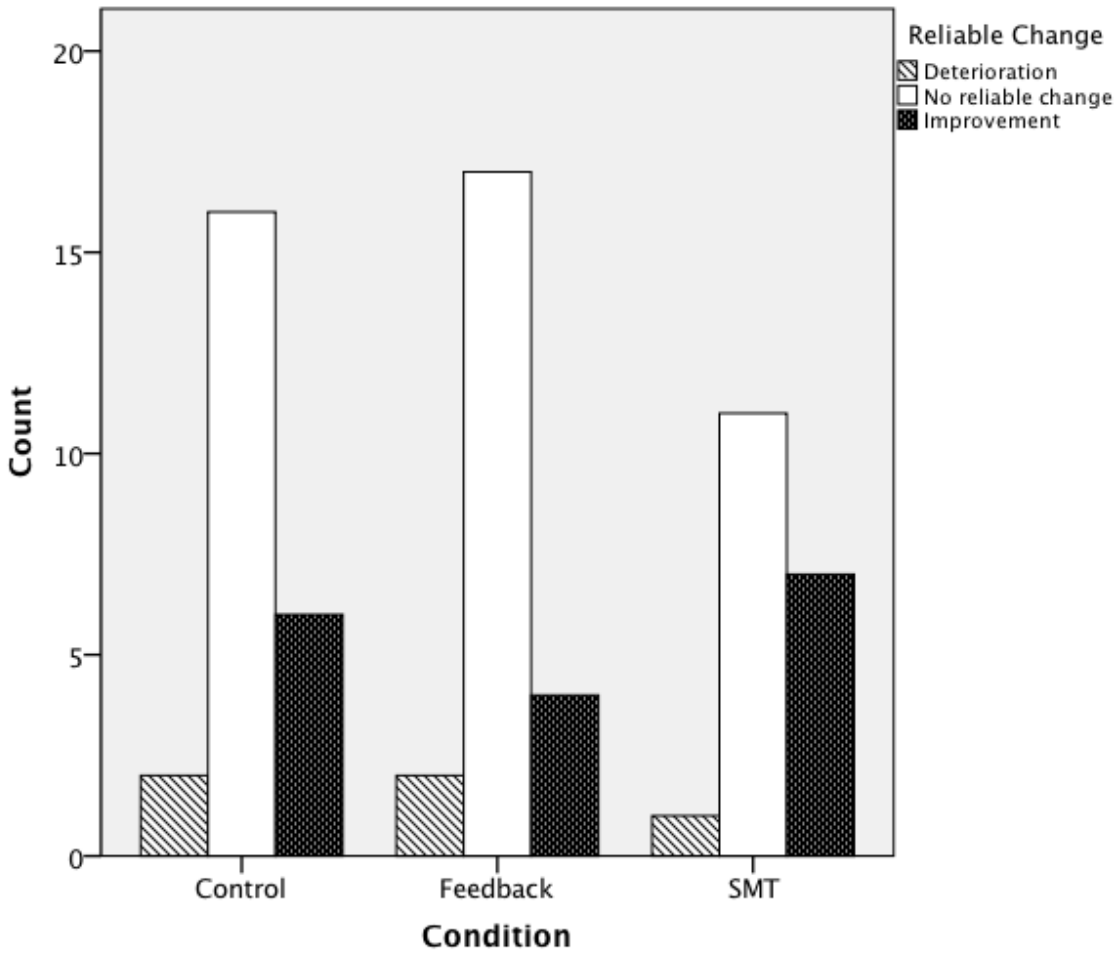

Figure 6. Number of participants who showed reliable improvement, reliable deterioration, or no reliable change, by condition. Reliable improvement was defined as a decrease in perceived stress greater than 1.96 times the standard error of the pre-post difference.

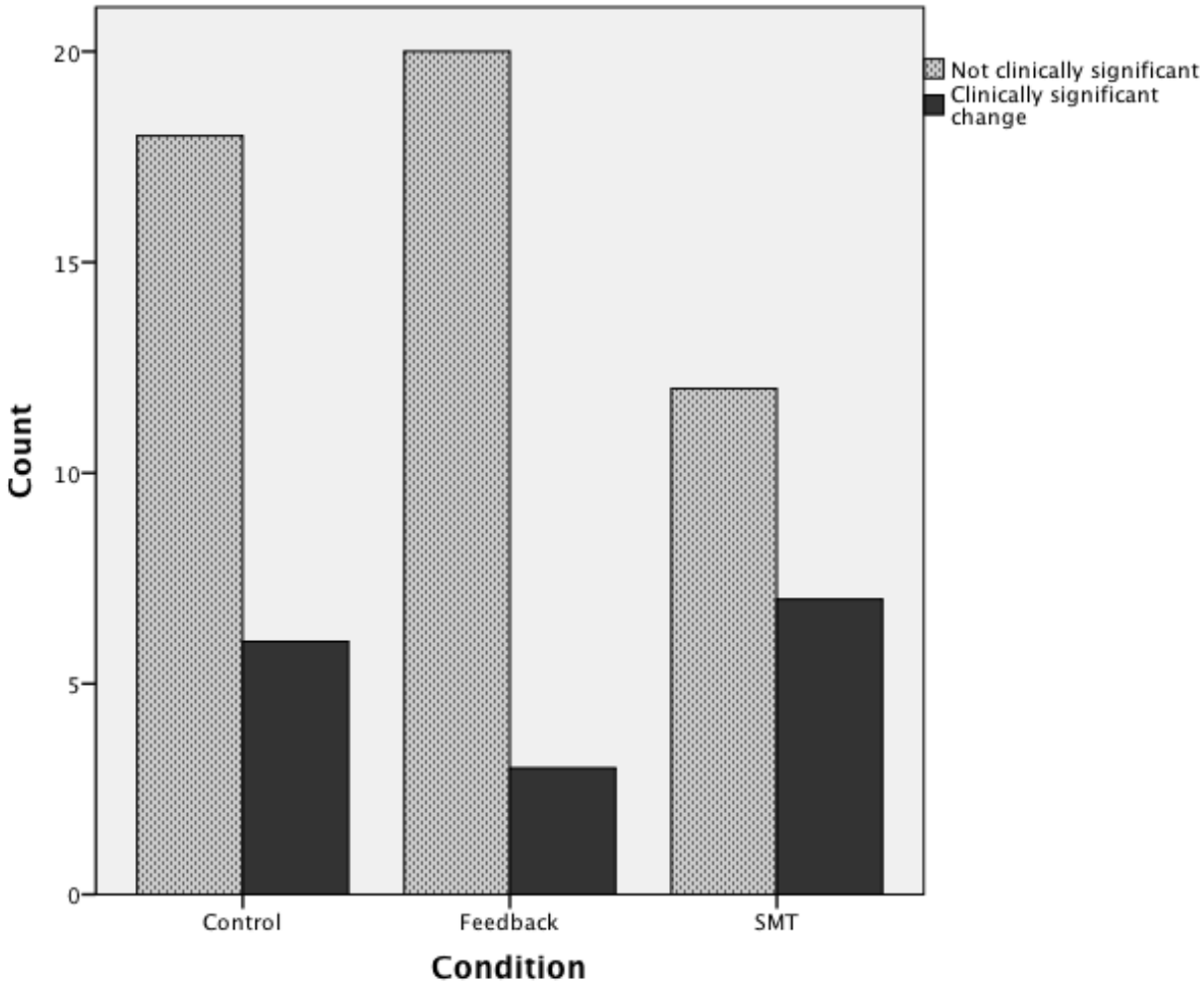

Figure 7. Number of participants who showed clinically significant change by condition. Clinically significant change was defined as both meaningful change (PSS < 17.25) and reliable improvement (decrease in stress > 5.01 points on the PSS).

Symbol	Definition	Value
M_1	Baseline mean for study sample	24.79
M_0	Mean of well-functioning, non-clinical population[a]	14.11
s_1	Standard deviation of study sample at baseline	3.77
s_0	Standard deviation of non-clinical population	3.79

[a]The mean of the non-clinical population was estimated using the larger pool of undergraduates screened for the study (N = 1,831). To adequately distinguish between "non-clinical" and "clinical" populations, the sample was split based on the eligibility criteria for this study (PSS <20 = non-clinical) and a mean and standard deviation were calculated for the non-clinical sub-sample.

Clinical Cutoff	Equation	Calculation	Cutoff Value
a	$M_1 - 2s_1$	$24.79 - 2(3.77)$	17.25*
b	$M_0 + 2s_1$	$14.11 + 2(3.77)$	21.66
c	$\frac{s_0 M_1 + s_1 M_0}{s_0 + s_1}$	$\frac{3.79\,(24.79) + 3.77\,(14.11)}{3.79 + 3.77}$	19.47

*Jacobson & Truax note, that when choosing which cutoff to use, *b* and *c* are usually preferable when norms are available. However, because the PSS is not a measure of psychopathology and the "clinical" and "non-clinical" distributions overlap highly, I opted to go with the most stringent value to avoid over-classifying individuals as having achieved meaningful change.

Figure 8a. Equations for clinical cutoffs of meaningful change as described by Jacobson & Truax (1991) and calculations for the present sample.

Symbol	Definition	Value
s_E	Standard error of the measure	see below
s_1	Standard deviation of study sample at baseline	3.77
r	Test-retest reliability of the measure	.77
s_{diff}	Standard error of the difference between baseline and post	see below
X_1	Baseline score for subject	-
X_2	Post study score for subject	-
RCI_i	Reliable Change Index for an individual subject	-
RCI_s	Reliable Change Index for the entire sample (cutoff)	see below

Symbol	Equation	Calculation	Value
s_E	$s_1\sqrt{1-r}$	$3.77\sqrt{1-.77}$	1.81
s_{diff}	$\sqrt{2(s_E{}^2)}$	$\sqrt{2(1.81^2)}$	2.56
RCI_i	$\frac{X_2 - X_1}{s_{diff}}$	-	-
RCI_s	$1.96(s_{diff})$	1.96 (2.56)	5.01

Figure 8b. Equations for cutoffs of reliable change as described by Jacobson & Truax (1991) and calculations for the present sample.

VITA

Jessica A. Chen was born in Mountain View, California. She graduated from Columbia University with a Bachelor of Arts in Psychology. She completed her Master of Science degree at the University of Washington; her master's thesis, titled "*The Shape of Symptom Change During Psychotherapy and Pharmacotherapy for PTSD,"* addressed how psychotherapy and medication differentially affect symptoms of post-traumatic stress disorder (PTSD). She completed her Doctor of Philosophy degree under the mentorship of Ronald E. Smith and Yuichi Shoda.