



This document is strictly CONFIDENTIAL.

**A Benchmarking Approach to Evaluating Effectiveness Research: Benchmarking The Lefkoe Method**

Journal:	<i>Psychotherapy Research</i>
Manuscript ID:	Draft
Manuscript Type:	Method Paper
Keywords:	Outcome Research
Keywords (user):	effectiveness research, benchmarking

SCHOLARONE™  
Manuscripts

## Abstract

We introduce a benchmarking protocol for effectiveness research and pilot its use evaluating two self-help programs: The Lefkoe Method (TLM) and Anthony Robbins' *Unleash the Power Within*. Participants were: (1) 32 purchasers who completed Natural Confidence, a DVD program using TLM, (2) 10 undergraduates who completed Natural Confidence, (3) 5 undergraduates who completed Robbins' program, and (4) 11 undergraduate controls. Despite sample limitations, effect sizes observed for completers of Natural Confidence were comparable to those found in CBT trials. Results provide preliminary support for TLM as an effective intervention. We discuss obstacles encountered and make recommendations for future researchers.

Keywords: Lefkoe Method, self-help, benchmarking, effectiveness research

## A Benchmarking Approach to Evaluating Effectiveness Research:

## Benchmarking The Lefkoe Method

Kazdin and Blase (2011) argued the dominant model of treatment in clinical practice, in-person individual or small group<sup>1</sup> treatment, is unsatisfactory. Clinical researchers spend too much time perfecting, understanding and disseminating resource-intensive forms of treatment; while important, these activities cannot meet the ever-increasing burden of mental health treatment<sup>2</sup>. If clinical researchers are to contribute to a significant reduction in this burden, other models of treatment delivery must be considered, models that can be tailored to meet the needs of the broader mental health community; *self-help* models were just one option mentioned in the portfolio presented.

Expanding treatment models to include less intensive forms of delivery will significantly expand the agenda of clinical researchers. Careful consideration of alternative models<sup>3</sup> and different standards in evaluating outcome research are needed (Kazdin & Blase, 2011). Less efficacious treatments that reach more of those burdened by mental health problems may be preferable to more efficacious treatments that meet the needs of relatively few.

Two important details are missing from this account: (1) a systematic method for evaluating the existing alternatives available to the public and (2) a statement about who should be doing those evaluations. We claim that clinical researchers are well-suited to evaluate such protocols, propose a method for doing so, and pilot its use with a brief evaluation of two self-help interventions.

---

<sup>1</sup> couples and families are included in this category

<sup>2</sup> According to Kazdin and Blase, this would be the case even if the number of professionals in this field doubled.<sup>[1]</sup>

<sup>3</sup> what Kazdin and Blase referred to as a portfolio of methods of delivery; they laid out one such portfolio in the paper.<sup>[1]</sup>

### *Self-Help Interventions*

Collins, Westra, Dozois, and Burns (2004) identified systematic barriers to accessing face-to-face treatment; of the thirteen identified, self-help protocols arguably eliminate or substantially reduce five: Willingness to disclose problems, fear of stigma and embarrassment, lack of time for treatment, geographic influences, and desire to handle problems on one's own. Consumers often opt for self-help programs citing financial and convenience considerations. This documented preference is analogous to the tendency for individuals suffering from specific phobia to prefer virtual reality (*in silico*) treatments to live (*in vivo*) procedures. The analogy breaks down when considering the availability of efficacy evaluations of virtual procedures. Increasingly, clinicians can make empirically-informed recommendations to clients preferring virtual options (for a review, see Myerbrocker & Emmelkamp, 2010); presently, clinicians and paraprofessionals cannot do the same for those favoring self-help options.

### *Roles for Treatment Developers, Treatment Providers, and Clinical Researchers*

McFall's (1991) manifesto compelled treatment providers to always disclose the nature of treatment before its administration; paraphrasing McFall, we argue ethics obligate a professional (e.g., self-help product provider) inform potential consumers of four intervention characteristics: (1) what the intervention entails, (2) what the treatment claims are, (3) what research suggests the efficacy and effectiveness are in relation to claims made, and (4) what research suggests are potential side effects. Moving beyond McFall's list, we pose a fifth: (5) a representative list of empirically validated alternatives (with a brief synopsis of 1-4 for each). This addition places responsibility on researchers to: (1) evaluate available alternatives empirically, regardless of one's affiliation or fondness for them, and (2) disseminate the findings to practicing clinicians who can pass the results on to consumers. This addition places

1  
2  
3 responsibility on treatment providers to: (1) stay informed about viable alternatives and their  
4 empirical status, and (2) apprise their consumers about these alternatives. Practically speaking,  
5  
6 we recognize this is not always possible; under such circumstances, practitioners should consider  
7  
8  
9  
10 referral.

11  
12 Others have examined the efficacy of internet-based self-help protocols (e.g.,  
13  
14 Christensen, Griffiths, Mackinnon, & Brittliffe, 2006 and Abott, Klein, & Ciechomski, 2008).  
15  
16 Most were attempts to disseminate empirically validated treatments (CBT primarily) in  
17  
18 convenient and readily available forms. A concentrated effort to help the consumer considering  
19  
20 alternatives that may be less expensive and/or more readily available than empirically validated  
21  
22 interventions is lacking; if clinical researchers do not make the effort, it is unlikely anyone will.  
23  
24  
25

### 26 27 *A Proposed Model for Evaluating Effectiveness*

28  
29 Effectiveness research, owing largely to increases in practical constraints and a vanishing  
30  
31 ability to constrain potential confounds, differs from efficacy research considerably. To argue  
32  
33 the two forms of research should be conducted the same way borders on ridiculous.  
34  
35 Interventions developed outside of the laboratory are largely in need of empirical evaluation.  
36  
37 Clinical researchers are well suited to take on the task, but the question of how to do so requires  
38  
39 careful thought.  
40  
41  
42

43  
44 We propose effectiveness researchers do the following: (1) present measurable treatment  
45  
46 claims (predictions), (2) select measures with sound psychometric properties to directly measure  
47  
48 claims, (3) administer measures pre-post and follow-up, (4) include control groups (e.g.,  
49  
50 alternative treatments with similar claims, wait list-controls, etc.), and (5) benchmark findings by  
51  
52 comparing effect sizes obtained to those observed in clinical trials for empirically-supported  
53  
54  
55 treatments.  
56  
57  
58  
59  
60

1  
2  
3 *Following this Model*  
4

5 We present a pilot study following these steps. We aimed to evaluate the effectiveness of  
6  
7  
8 The Lefkoe Method claims to reduce the impact of daily stress, increase self-esteem, decrease  
9  
10 negative affect, increase positive affect, and decrease anxiety in social situations. Because this is  
11  
12 real-world work, the Ns are small and unequal, we could not sample randomly, and many  
13  
14 participants dropped out of the study. Nonetheless, effect sizes were impressive when  
15  
16 benchmarked against those seen in trials of CBT. We have, then, preliminary empirical support  
17  
18 for The Lefkoe Method, and a series of methodological suggestions for future researchers who  
19  
20 elect to follow this benchmarking model for effectiveness research.<sup>4</sup>  
21  
22  
23

24 *The Lefkoe Method*  
25

26  
27 Appealing to McFall's (1991) manifesto, we introduce this method by making relevant  
28  
29 disclosures<sup>5</sup>.  
30

31  
32 *Disclosure 1:* the nature of the intervention. The full Lefkoe Method (TLM) consists of  
33  
34 seven interventions; two were used within the Natural Confidence program.  
35

36  
37 The first is the Lefkoe Belief Process (LBP), designed to eliminate beliefs formed earlier  
38  
39 in life hypothesized to cause the presenting problem. A typical presenting problem is fear of  
40  
41 public speaking; some of the 8-10 beliefs Lefkoe argues *cause* this problem include: "Mistakes  
42  
43 and failure are bad", "What I have to say isn't important", "I'm not good enough", "I'm not  
44  
45 important", and "What makes me good enough and important is having people think well of  
46  
47  
48

---

49 <sup>4</sup> It should be noted that our recommendation to use benchmarking in evaluating findings is not unique (e.g.,  
50 Franklin, Abramowitz, Kozak, Levitt, & Foa, 2000 and Wade, Treat, & Stuart, 1998)

51  
52  
53 <sup>5</sup> We leave out the fifth (a list of alternatives) here purposefully, as in research settings this is taken care of by the  
54  
55 choice of suitable comparison groups and the use of benchmarking in the evaluation of findings. In the therapeutic  
56  
57 context this is something the therapist should disclose directly to the client.  
58  
59  
60

1  
2  
3 me”. The LBP consists of a series of questions and distinctions asked of a client (in this case) on  
4  
5 an interactive Internet program. The LBP facilitator’s first job, after helping the client to identify  
6  
7 the presenting problem clearly, is to help them identify the relevant beliefs thought to cause the  
8  
9 problem. Once identified, the client uses the LBP to eliminate the beliefs. The client states the  
10  
11 belief at the start of the Process, acknowledging the belief “feels true.” At the end of the  
12  
13 Process, the client again states the words of the belief and comments on them. The phrase “feels  
14  
15 empty, with no feeling; the statement does not feel true” provides a provisional sign that the  
16  
17 Process eliminated the belief.  
18  
19  
20  
21

22 The second is the Lefkoe Stimulus Process (LStimP), designed to extinguish responses to  
23  
24 previously conditioned stimuli. Using the same fear of public speaking problem mentioned  
25  
26 above, typical “conditionings” include: Fear associated with criticism and judgment, fear  
27  
28 associated with not meeting expectations, fear associated with people putting their attention on  
29  
30 me, and fear associated with rejection. Again, this Process consists of a few questions and  
31  
32 distinctions.<sup>6</sup>  
33  
34  
35

36 *Disclosure 2: treatment claims.* Lefkoe claims eliminating all beliefs and conditionings  
37  
38 thought to cause psychological problems eliminates those problems. Specifically, he refers to a  
39  
40 large body of anecdotal evidence (citing a minimum of 500 written testimonials<sup>7</sup>) that people  
41  
42 who used TLM reported significant reductions in anxiety and stress and improvement in self-  
43  
44 esteem.  
45  
46  
47

48 *Disclosures 3 and 4: empirical evidence and evidence of iatrogenic effects.* Here, there is  
49  
50 little to disclose. Lefkoe has sought independent empirical evaluation of TLM. The single  
51  
52  
53  
54

---

55 <sup>6</sup> More information about the Lefkoe Belief Process can be found in *The Therapist*, May 2, 2001, published by the  
56 California Association of Marriage and Family Therapists.

57 <sup>7</sup> Some of these testimonials can be found at <http://blog.recreateyourlife.com>.  
58  
59  
60

1  
2  
3 existing evaluation of TLM suggests treatment participants significantly reduced and in some  
4 cases completely eliminated self-reported fear of public speaking (on a 1-10 scale)  
5  
6  
7  
8 (Cunningham, Lefkoe & Sechrest, 2006). Specifically, after an average of 3.3 one-hour sessions,  
9  
10 mean reported fear decreased from 7 to 1.5 on a 10-point scale; a 1 on the scale referred to “no  
11  
12 fear at all.”

13  
14  
15 *Anthony Robbins' Unleash the Power within Program*

16  
17  
18 Appealing again to McFall's (1991) manifesto, we introduce this method by making  
19  
20 relevant disclosures.

21  
22 *Disclosure 1:* The version used consists of twelve audio sessions in MP3 format. Table  
23  
24  
25 1 provides the content outline of these sessions.

26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

---

Insert Table 1 here

---

34 *Disclosure 2:* The website promoting the material<sup>8</sup> details nine claims: the user will learn  
35  
36 to (1) break old patterns and destroy limiting beliefs, (2) become a decisive person who takes  
37  
38 massive, intelligent action, (3) condition the self mentally, emotionally, and physically to create  
39  
40 staggering results, (4) find and develop a strategy of success, (5) have more confidence and self-  
41  
42 esteem, (6) develop new communication skills, (7) generate more passion and enthusiasm, (8)  
43  
44 become aware of the higher purpose for your life, and (9) discover what drives and what holds  
45  
46 the self back. Unfortunately, only one of these (number 5) is directly measurable without  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

---

<sup>8</sup> See [www.tonyrobbins.com/events/unleash-the-power-within/](http://www.tonyrobbins.com/events/unleash-the-power-within/)



1  
2  
3 considerable extrapolation. If accurate, these claims predict improvements on measures of the  
4  
5 impact of daily stress, of positive and negative affect, and of anxiety.<sup>9</sup>  
6  
7

8 *Disclosures 3 and 4:* There is nothing to disclose. To our knowledge, no empirical effort  
9  
10 to examine these claims has been made.  
11

## 12 Materials and Methods

13  
14 This study was conducted in compliance with the University of Arizona Institutional  
15  
16 Review Board. Students were provided a verbal description of the project and viewed a written  
17  
18 consent form before volunteering to participate. Online purchasers saw this same approved  
19  
20 consent form, and were not permitted to access study questionnaires before providing electronic  
21  
22 consent.  
23  
24  
25

### 26 *Participants*

27  
28 Three convenience samples served: (1) a sample of purchasers who completed Natural  
29  
30 Confidence online, (2) a sample of students who completed Natural Confidence online, and (3) a  
31  
32 sample of students randomly assigned to complete either the Robbins comparison program or the  
33  
34 waitlist control condition.  
35  
36  
37

38  
39 Lefkoe gathered the voluntary group of purchasers, by offering the opportunity to  
40  
41 participate in exchange for receiving the next program the Lefkoe Institute offered free of  
42  
43 charge. Of those who volunteered, 102 completed the pre-measures. Of those, 32 completed  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54

---

55 <sup>9</sup> These predictions are also made by TLM claims, making the Tony Robbins program a suitable comparison/control  
56  
57 condition.  
58  
59  
60

1  
2  
3 usable portions<sup>10</sup> of the post-measures. The 32 purchasers (13 female, 19 male) ranged in age  
4  
5 from 22 to 75 years, with a mean of 45.4 and a median of 44 years.  
6  
7

8 The remaining two samples were collected at the University of Arizona. The decision to  
9  
10 gather a convenience sample of undergraduates was the result of financial, time, and other  
11  
12 practical constraints. Ideally, we would randomly assign participants to conditions, but the  
13  
14 comparison conditions were not ready when the initial opportunity for convenience sampling  
15  
16 presented itself.<sup>11</sup> The second sample, then, was of college students who would complete  
17  
18 Natural Confidence online. Students in a university course were offered the opportunity to  
19  
20 participate in partial fulfillment of course requirements. In a class of 34, 21 elected to complete  
21  
22 the pre-measures; only 10 accessed the online material and completed the post-measures<sup>12</sup>. The  
23  
24 10 students (8 female, 2 male) ranged in age from 19 to 35 years, with a mean of 22.9 and a  
25  
26 median of 21 years.  
27  
28  
29  
30

31 The third sample was of college students who would complete one of two conditions: the  
32  
33 Tony Robbins (comparison) program or the waitlist (time) control. Students in another  
34  
35 university course were offered the opportunity to participate in partial fulfillment of course  
36  
37 requirements. In a class of 26, 19 volunteered to participate. We randomly assigned the first 14  
38  
39  
40  
41

42  
43 <sup>10</sup> If a participant skipped only one outcome measure, they were simply not included in the analysis of that outcome  
44  
45 measure. If they left items blank on a given outcome measure they were similarly stricken from the analysis of that  
46  
47 outcome measure. This procedure alone accounts for differing sample sizes within a condition across measures.  
48

49 <sup>11</sup> Repeated attempts to secure materials from providers of Anthony Robbins' *Unleash the Power Within* audio self-  
50  
51 help program to evaluate its effectiveness ultimately proved unsuccessful. Eventually we decided to buy a small  
52  
53 number of audio programs to use as a comparison condition, but we made this decision only after exhausting other  
54  
55 options for securing the material.

56  
57 <sup>12</sup> The remaining 11 opted for alternative research participation or writing assignments to fulfill course requirements.  
58  
59  
60

1  
2  
3 students sampled to either the comparison (Robbins program) or control (sampling changes in  
4  
5 measures pre-post, in the absence of an intervention) condition. Of those, 7 completed the pre-  
6  
7 measures for each condition; 5 accessed the Robbins audio material<sup>13</sup> and completed the post-  
8  
9 measures, and 6 controls completed the post-measures. An additional 5 students completed the  
10  
11 control condition, bringing the control sample size closer to that of the first student sample. The  
12  
13 5 students (2 female, 3 male) completing the Robbins audio program and usable portions of the  
14  
15 pre-and post-measures ranged in age from 23 to 37 years, with a mean of 27.4 and a median of  
16  
17 25 years. The 11 students (6 female, 5 male) not completing an intervention but completing  
18  
19 usable portions of the pre-and post-measures ranged in age from 18 to 20 years, with a mean of  
20  
21 18.4 and a median of 18 years.  
22  
23  
24  
25  
26

### 27 *Intervention Materials*

#### 28 *Natural Confidence, online version*

29  
30  
31 The program consists of a single DVD, available for download from a secure website,  
32  
33 which uses LBP to help viewers eliminate 19 beliefs and the LStimP to extinguish four  
34  
35 conditioned responses. Lefkoe offers a lifetime, money-back guarantee that users of this DVD  
36  
37 will achieve the following results: significantly improve confidence, stop rejection fear, stop  
38  
39 social anxiety, stop concerns about the opinions of others, significantly reduce stress, stop the  
40  
41 “little voice” that is critical of whatever one does, stop procrastination, and stop perfectionism.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
Natural Confidence has been available for about two years and roughly 10% of purchasers have  
requested a refund.

#### 51 *Anthony Robbins' Unleash the Power Within, MP3 version*

---

54  
55  
56  
57  
58  
59  
60  
<sup>13</sup> The Robbins' intervention was expensive; thus, we only purchased seven of them; five of the seven individuals  
given access to the material completed the pre- and post- measures

1  
2  
3 This version consists of twelve audio MP3 files, with the content specified in Table 1; the  
4  
5 program was purchased from Nightingale-Conant (an independent distributor).  
6  
7

### 8 *Outcome Measures*

#### 9 10 *Daily Stress Inventory (Brantley, Waggoner, Jones, & Rappaport, 1987)*

11  
12 Used to assess level of impact of daily stressors; this 58-item measure assesses frequency  
13 and associated subjective distress relating to events that may have occurred within the past 24  
14 hours. We selected this measure because of evidence that it correlates with biochemical  
15 measures of daily stress (Brantley, McKnight, Jones, Dietz & Tulley, 1988). Changes observed  
16 on the DSI appear to predict health-related improvements. Traditionally, this inventory produces  
17 two measures: (1) the *Frequency* of events endorsed, and (2) the *Sum* of stress ratings of  
18 endorsed events); we used the latter in this study.  
19  
20  
21  
22  
23  
24  
25  
26  
27

#### 28 29 *Positive and Negative Affect Scale (Watson, Clark, & Tellegen, 1988)*

30  
31 Used to assess levels of positive and negative affect; matched-valence responses on these  
32 scales were summed to create two overall levels of reported affect: positive and negative.  
33  
34  
35

#### 36 37 *Rosenberg Self-Esteem Scale (Rosenberg, 1989)*

38  
39 Used to assess level of self-esteem; test-retest correlations typically range from 0.82 to  
40 0.88, with Cronbach's alpha values ranging from 0.77 to 0.88 (e.g., Rosenberg, 1989 and  
41 Blascovich & Tomaka, 1993).  
42  
43  
44

#### 45 46 *Spielberger State-Trait Anxiety Inventory (Spielberger, Gorsuch, & Lushene, 1970)*

47  
48 Used to assess temporary state (STAI-S) and more stable trait (STAI-T) anxiety; the  
49 mean reliability coefficient is 0.765 (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983).  
50  
51  
52

### 53 *Procedure*

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
Sample 1: Program purchasers could click a link on the Lefkoe Website inviting them to participate voluntarily in exchange for receiving a free copy of the Lefkoe Institute's next program. If clicked, the link generated an email providing instruction required to access a secure website where they could fill out pre- and post-measures. Upon entering the website, participants received a consent statement noting their responses would be de-identified and requesting to use their data for research purposes. The instructions specified they answer only questions they felt comfortable answering, they quit at any time without penalty, and they return to complete the post-measures when they had finished working with the treatment material. Unlike students in our next sample, purchasers could eliminate up to four beliefs using TLM before purchasing the program<sup>14</sup>.

27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
Samples 2 and 3: Student participants completed an informed consent procedure, received an informational packet detailing how to log into a secure website, and, once there, completed the pre-measures. For all but the time-control participants, the packet also detailed how to access the assigned treatment material. The Natural Confidence materials and the Robbins' materials were available via websites requiring only the provided password. The former site permitted modular video streaming of individual components of the DVD-based treatment; the latter provided for the download of the audio MP3 files listed previously. Participants in both treatment conditions signed forms agreeing not to provide the treatment material to anyone else. Participants were allotted two weeks to work with the treatment material, and were then asked to log back into the secure website to complete the post-measures.

---

51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
<sup>14</sup> The web site offers three initially. A couple of weeks after a person signs up to gain access to the three they are mailed a link to a fourth. That is, purchasers could eliminate up to four of the nineteen beliefs before participating in the study. It is possible that outcome measures would have revealed even poorer pre-values for this group had purchasers had no contact with treatment materials at the time they completed the pre-measures.

1  
2  
3 At that time, they were asked to report how far into the program they went before completing  
4  
5 post-measures; the data were completely de-identified.  
6  
7

### 8 *Statistical Analyses*

9  
10 *First Analysis:* Because we used convenience samples, we conducted an ANOVA for  
11  
12 each outcome measure to assess group-level differences in the pre-values of the constructs we  
13  
14 measured. We describe these first to contextualize the pre-post differences we found in  
15  
16 subsequent analyses.  
17  
18

19  
20 *Second Analysis:* Following a comparison of group-level differences in pre-values, we  
21  
22 conducted a series of independent t-tests<sup>15</sup> for each measure by treatment group. Given these  
23  
24 independent comparisons increase the likelihood of making a Type I error, we used a Bonferroni  
25  
26 correction to arrive at an alpha level of 0.01 for each significance test.<sup>16</sup> That is, for each  
27  
28 outcome measure we used an alpha level of 0.01 to evaluate the significance of a two-tailed  
29  
30 paired t-test assessing pre-post changes among participants in each condition. Only those  
31  
32 completing usable portions of the pre- and post-measures were included in these analyses.  
33  
34

35  
36 *Third Analysis:* Finally, we made an effort to benchmark our findings against available  
37  
38 effect sizes observed on the same outcome measures from CBT studies. We used two methods  
39  
40 for our search.  
41  
42

43  
44 The first search was a more systematic review of CBT studies targeting anxiety. We  
45  
46 identified these based on the results of an ongoing meta-analysis examining the results of  
47  
48 component-controlled CBT studies currently underway in our laboratory (Jacobs, Hill, Gable, &  
49

---

50  
51 <sup>15</sup> We evaluated the conditions independently because of small and unequal across-group sample sizes.

52  
53 <sup>16</sup> Bonferroni corrections require assessing statistical significance given an alpha level of  $1/n$  times what it would be  
54  
55 if only one hypothesis were tested. Since we tested five hypotheses within each sample (associated with the five  
56  
57 outcome measures employed), this rendered our alpha level 0.01.  
58  
59  
60

1  
2  
3 Brown, 2011). Although systematic, this method rendered a large number of studies using  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Brown, 2011). Although systematic, this method rendered a large number of studies using measures that differed from our own. The obvious suggestion is to consider carefully which studies will be used as benchmarks *before* selecting outcome measures. This method did render several studies using Spielberger's trait anxiety measure across which we could make benchmark comparisons.

The second search was less systematic and did not emphasize a prioritizing of strong methods; we searched using keywords relating to our other specific outcome measures (always including the keyword "CBT"). Again, we suggest future researchers select studies against which they will benchmark based on strong methods and appropriate sampling techniques; outcome measures can then be selected to match those used in the benchmarking studies<sup>17</sup>.

In calculating effect sizes, we used a method recommended by Dunlop, Cortina, Vaslow and Burke (1996). Specifically, in calculating Cohen's *d* for our paired t-tests, we elected to use the original means and standard deviations from our samples; this provides a more conservative measure of effect size, as pooled standard deviations that correct for level of correlation among measures result in effect size overestimation.<sup>18</sup>

## Results

*First Analysis:* A series of one-way ANOVAs revealed no significant pre-test differences across treatment group on the impact of daily stress [ $F(3, 53) = 1.027, p = 0.388$ ], levels of

---

<sup>17</sup> Adding additional measures is not problematic, providing measures used in the original study are included as well.

<sup>18</sup> The degree to which the resulting effect size is an overestimate depends upon the magnitude of the intercorrelation; intercorrelations of at least 0.8 produce effect sizes that are more than twice the magnitude of the effect size as calculated using original standard deviations. For arguments in favor of using pooled standard deviation calculations, see Rosenthal (1991) or Mullen & Rosenthal (1985).

1  
2  
3 negative affect [ $F(3, 54) = 2.096, p = 0.111$ ], or self-esteem [ $F(3, 54) = 2.160, p = 0.103$ ]. Pre-  
4  
5 test differences across treatment group were observed in age [ $F(3, 54) = 25.267, p = 0.000$ ],  
6  
7 level of positive affect [ $F(3, 54) = 6.921, p = 0.001$ ] and trait anxiety [ $F(3, 51) = 3.926, p =$   
8  
9 0.013]. Tukey post hoc analyses suggest that differences were constrained to those between  
10  
11 purchasers of the online Natural Confidence program and one or more of the student conditions.  
12  
13 We expected such differences given the different demographic; this provides further support for  
14  
15 our decision to gather a student group of completers of Natural Confidence for comparison.  
16  
17  
18  
19

20 Purchasers of Natural Confidence were significantly older than were students completing  
21  
22 Natural Confidence (Mean difference = 23.013,  $p = 0.000$ ), the Robbins Program (Mean  
23  
24 difference = 18.413,  $p = 0.000$ ), and the time control (Mean difference = 18.413,  $p = 0.004$ );  
25  
26 pair-wise comparisons of the three student conditions detected no significant mean differences in  
27  
28 age. Purchasers of Natural Confidence reported significantly less positive affect than students in  
29  
30 the time control (Mean difference = 10.205,  $p = 0.001$ ) and scored higher on measures of trait  
31  
32 anxiety than students completing the Robbins Program (Mean difference = 15.048,  $p = 0.033$ ).  
33  
34 Tukey post hoc analyses revealed no other significant pair-wise comparisons of differences in  
35  
36 positive affect or trait anxiety.  
37  
38  
39  
40  
41

42 Findings render comparisons across the student conditions clearly defensible, given the  
43  
44 absence of significant differences in pre-test values across those conditions. Given the  
45  
46 differences observed in reported level of positive affect and trait anxiety in purchasers of Natural  
47  
48 Confidence, we watched for significant pre-post differences observed in the purchasers group but  
49  
50 *not* in the student Natural Confidence group; such changes could arguably be an artifact of a  
51  
52 ceiling effect.  
53  
54  
55  
56  
57  
58  
59  
60



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

*Second Analysis:* A series of paired samples t-tests, at a Bonferroni-corrected alpha of 0.01, were performed to investigate pre-post differences within treatment groups. As shown in Tables 2 and 3, there were no significant pre-post differences on any measures for time control participants or for completers of the Robbins program.

We had the smallest number of completers (5 students) in the Robbins condition. A quick look at tested significance levels indicates that two pre-post differences might reach significance with a larger sample size: level of reported negative affect and level of trait anxiety. Power analyses were conducted to determine the sample sizes that would be necessary for differences of this size to be detected at a power of 80% (the standard typically sought by researchers). With respect to the negative affect measure, the present sample size (5) afforded a power of only 8.4% to detect a true difference of this magnitude; a sample size of 133 would be needed to increase the power to 80%. With respect to the trait anxiety measure, the present sample size (5) afforded a power of only 11.7% to detect a true difference of this magnitude; a sample size of 80 would be needed to increase the power to 80% (would actually be 80.2%). That is, there is a non-significant trend for Robbins program completers to reduce their levels of negative affect and trait anxiety following completion of the program.

---

Insert Tables 2 and 3 here

---

As Table 4 illustrates, among student completers of Natural Confidence there were no significant changes in negative affect or self-esteem; impact of daily stress was reduced, levels of reported positive affect increased, and trait anxiety decreased. As Table 5 illustrates, among online purchasers (completers) of Natural Confidence there were significant pre-post differences on all measures: impact of daily stress was reduced, levels of reported positive affect increased,

1  
2  
3 levels of reported negative affect decreased, self-esteem increased, and trait anxiety decreased.  
4  
5  
6 Importantly, observed pre-post differences in levels of reported positive affect and trait anxiety  
7  
8 were observed in *both* the purchaser and student completer group, suggesting these differences  
9  
10 cannot be attributed simply to pre-test differences across groups<sup>19</sup>.  
11

12  
13  
14 

---

Insert Tables 4 and 5 here

---

15  
16  
17  
18 *Third Analysis:* Tables 6 through 9 display the effect size calculations for pre-post  
19  
20 differences on outcome measures across conditions. In general, effect sizes ranging from 0.1 to  
21  
22 0.3 are considered small, those ranging from 0.3 to 0.5 are considered moderate, and those above  
23  
24 0.5 are considered large (Cohen, 1988). Table 10 displays the effect size calculations for studies  
25  
26 we found that used a CBT-related intervention and one or more of the outcome measures used  
27  
28 here<sup>20</sup>. By CBT-related, we mean interventions using at least a cognitive restructuring element;  
29  
30 we did not exclude studies including exposure components/elements.  
31  
32

33  
34  
35 

---

Insert Tables 6-10 here

---

36  
37  
38  
39 Recall significant pre-post differences occurred only in completers of the Natural  
40  
41 Confidence Program. Student completers showed significant improvements in the hypothesized  
42  
43 direction in daily stress (effect size = 0.27; considered small), positive affect (effect size = -1.03;  
44  
45 considered large), and trait anxiety (effect size = 1.08; considered large). Online purchasers who  
46  
47  
48

49  
50  
51 <sup>19</sup> if you recall, online purchasers were significantly lower in reported levels of affect than time control participants  
52  
53 and higher in trait anxiety than Robbins completers

54  
55 <sup>20</sup> All effect sizes were calculated using the original means and standard deviations from the pre- and post-treatment  
56  
57 samples.  
58  
59  
60

1  
2  
3 completed the Natural Confidence program showed significant improvements in the  
4  
5 hypothesized direction in daily stress (effect size = 1.08; considered large), self-esteem (effect  
6  
7 size = -0.50; considered moderate), negative affect (effect size = 1.31; considered large), positive  
8  
9 affect (effect size = -1.19; considered large), and trait anxiety (effect size = 1.25; considered  
10  
11 large).  
12  
13

14  
15 The best benchmark comparison available is of trait anxiety. The average effect size for  
16  
17 pre-post differences in trait anxiety in the CBT-related studies reviewed was 1.22 (considered  
18  
19 large). This average effect size is only slightly larger than student completers of Natural  
20  
21 Confidence (effect size = 1.08) and slightly smaller than purchasers (effect size = 1.25). Given  
22  
23 the problematic sampling techniques and abnormally low sample sizes, these effect sizes are  
24  
25 impressive and comparable to those seen in CBT-related interventions.  
26  
27

28  
29 We could also make benchmark comparisons of differences in self-esteem, positive  
30  
31 affect, and negative affect. These comparisons are preliminary, however, as only a single study  
32  
33 is cited for each. Nonetheless, it gives us a reasonable idea about the size of these effects in  
34  
35 comparison to those found in studies using CBT-related interventions. Purchasers demonstrated  
36  
37 moderate improvements in self-esteem (effect size = -0.50); this is only slightly smaller than that  
38  
39 shown in the referenced benchmark study (effect size = -0.77). In sad affect, CBT completers in  
40  
41 the benchmark study showed improvement (effect size = 0.60) smaller than that observed in  
42  
43 purchasers of Natural Confidence (effect size = 1.31). Improvements among CBT-completers in  
44  
45 positive affect (effect size = -0.35) were relatively small in comparison to those among student  
46  
47 completers of Natural Confidence (effect size = -1.03) and purchasers (effect size = -1.19).  
48  
49

50  
51 Overall, then, the effect sizes observed among completers of Natural Confidence are impressive  
52  
53 when benchmarked across available comparison studies using CBT-related interventions.  
54  
55  
56  
57  
58  
59  
60

## Discussion

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54

Self-help programs are increasingly available; consumers who opt for these over other available options have little empirical guidance. We propose clinical researchers make concentrated efforts to evaluate available self-help treatments empirically. A protocol for effectiveness research was introduced, borrowing heavily from McFall's (1991) Manifesto. In an attempt to follow this protocol, the effectiveness of The Lefkoe Method (in the form of an online Natural Confidence program) at reducing the impact of daily stress, increasing self-esteem, decreasing negative affect, increasing positive affect, and decreasing anxiety was compared to that of Anthony Robbins' *Unleash the Power Within*. Observed effects were evaluated in a benchmarking fashion, in relation to effect sizes observed in available CBT studies using the same outcome measures.

There were minimal pre-test differences across conditions, all of which specified differences between the online purchasers of Natural Confidence and another group(s); these were anticipated. No statistically significant pre-post changes occurred among students in the time-control or Robbins program conditions. Significant pre-post changes occurred on all measures among purchasers of Natural Confidence; student completers of the program reported reduced levels of the impact of daily stress, increased levels of positive affect, and decreased levels of trait anxiety<sup>21</sup>. The observed effect sizes appear comparable or slightly larger than those observed in comparison CBT studies; this is the case despite several shortcomings which would render this outcome less likely, including small sample sizes, non-standard sampling methods, and high attrition rates. Although this result is promising, we should remind readers

---

<sup>21</sup> In other words, no significant changes were observed in levels of negative affect or self-esteem among student completers.

1  
2  
3 that our search for CBT studies was not carefully constrained; instead, it represents comparisons  
4  
5 with readily available data.  
6  
7

### 8 Conclusions

9  
10 We propose researchers evaluate effectiveness using five steps: (1) present measurable  
11 treatment claims (predictions), (2) select measures with sound psychometric properties tied  
12 directly to treatment claims, (3) administer measures pre-post and follow-up, (4) include control  
13 groups (e.g., alternative treatments with similar treatment predictions, wait list-controls, etc.),  
14 and (5) benchmark findings by comparing effect sizes obtained to those observed in clinical trials  
15 for empirically-supported treatments.  
16  
17  
18  
19  
20  
21  
22  
23

24 On the basis of our experience piloting this protocol with the samples discussed herein,  
25 we offer two critical suggestions to future researchers: (1) select measures used by studies  
26 reported in meta-analyses, and (2) use *consumer samples*<sup>22</sup>, as college students lack comparable  
27 levels of motivation and have different pre-test characteristics.  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51

---

52  
53 <sup>22</sup> This word choice was purposeful. Choose participants carefully; our primary conclusion is that undergraduate  
54 samples are likely to prove more trouble than they are worth. By consumers we include treatment-seekers; our use  
55 of this term is similar to the clinical term therapy *customers*.  
56  
57  
58  
59  
60

## References

- Abbott, J.A.M., Klein, B., & Ciechomski, L. (2008). Best practices in online therapy. *Journal of Technology in Human Services, 26*, 360-375.
- Blascovich, J., & Tomaka, J. (1993). Measures of self-esteem. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (3rd ed; pp. 115-160). Ann Arbor: Institute for Social Research.
- Borkovec, T.D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology, 61*(4), 611-619.
- Borkovec, T.D., Newman, M.G., Pincus, A.L., & Lytle, R. (2002). A component analysis of cognitive-behavioral therapy for generalized anxiety disorder and the role of interpersonal problems. *Journal of Consulting and Clinical Psychology, 70*(2), 288-298.
- Brantley, P.J., McKnight, G.T., Jones, G.N., Dietz, L.S., & Tulley, R. (1988). Convergence between the daily stress inventory and endocrine measures of stress. *Journal of Consulting and Clinical Psychology, 56*(4), 549-551.
- Brantley, P.J., Waggoner, C.D., Jones, G.N., & Rappaport, N.B. (1987). The daily stress inventory: development, reliability, and validity. *Journal of Behavioral Medicine, 10*(1), 61-73.
- Brown, J.S.L., Elliott, S.A., Boardman, J., Andiappan, M., Landau, S., & Howay, E. (2008). Can the effects of a 1-day CBT workshop on self-confidence be maintained after 2 years? A Naturalistic study. *Depression and Anxiety, 25*, 632-640.
- Bryant, R.A., Moulds, M.L., Guthrie, R.M., Dang, S.T., Mastrodomenico, J., Nixon, R.D.V., Felmingham, K.L., & Hopwood, S. (2008). A randomized controlled trial of exposure

1  
2  
3 therapy and cognitive restructuring for posttraumatic stress disorder. *Journal of*  
4  
5  
6 *Consulting and Clinical Psychology*, 76(4), 695-703.

7  
8 Christensen, H., Griffiths, K. M., Mackinnon, A. J., & Brittliffe, K. (2006). Online randomized  
9  
10 controlled trial of brief and full cognitive behaviour therapy for depression.  
11  
12 *Psychological Medicine*, 36, 1737-1746.

13  
14  
15 Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence  
16  
17 Erlbaum Associates.

18  
19  
20 Collins, K.A., Westra, H.A., Dozois, D.J.A., & Burns, D.D. (2004). Gaps in accessing treatment  
21  
22 for anxiety and depression: Challenges for the delivery of care. *Clinical Psychology*  
23  
24 *Review*, 24(5), 583-616.

25  
26  
27 Cunningham, V., Lefkoe, M., & Sechrest, L. (2006). Eliminating fears: An intervention that  
28  
29 permanently eliminates the fear of public speaking. *Clinical Psychology &*  
30  
31 *Psychotherapy*, 13, 183-193.

32  
33  
34 Dunlop, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of  
35  
36 experiments with matched groups or repeated measures designs. *Psychological Methods*,  
37  
38 1, 170-177.

39  
40  
41 Franklin, M. E., Abramowitz, J. S., Kozak, M. J., Levitt, J. T., & Foa, E. B. (2000). Effectiveness  
42  
43 of exposure and ritual prevention for obsessive-compulsive disorder: randomized  
44  
45 compared with nonrandomized samples. *Journal of Consulting and Clinical Psychology*,  
46  
47 68, 594-602.

48  
49  
50 Jacobs, W.J. Hill, D., Gable, J., & Brown, S., (2011). A meta-analysis of component-controlled  
51  
52 CBT studies. Presented to the 44th Annual Meeting of the Association for Behavioral  
53  
54 and Cognitive Therapies, November, 2010, San Francisco, CA.

- 1  
2  
3 Kazdin, A.E., & Blase, S.L. (2011). Rebooting research and practice to reduce the burden of  
4  
5 mental illness. *Perspectives on Psychological Science*, 6(1), 21-37.  
6  
7  
8 McFall, R.M. (1991). Manifesto for a science of clinical psychology. *The Clinical Psychologist*,  
9  
10 44(6), 75-88.  
11  
12  
13 Meyerbrocker, K., & Emmelkamp, P.M.G. (2010). Virtual reality exposure therapy in anxiety  
14  
15 disorders: A systematic review of process-and-outcome studies. *Depression and Anxiety*,  
16  
17 27, 933-944.  
18  
19  
20 Mullen, H., & Rosenthal, R. (1985). *BASIC meta-analysis: Procedures and programs*. Hillsdale,  
21  
22 NJ: Earlbaum.  
23  
24  
25 Parrish, B.P., Cohen, L.H., Gunthert, K.C., Butler, A.C., Laurenceau, J., & Beck, J.S. (2009).  
26  
27 Effects of cognitive therapy for depression on daily stress-related variables. *Behaviour*  
28  
29 *Research and Therapy*, 47, 444-448.  
30  
31  
32 Paunovic, N., & Ost, L. (2001). Cognitive-behavioral therapy vs. exposure therapy in the  
33  
34 treatment of PTSD in refugees. *Behaviour Research and Therapy*, 39(10), 1183-1197.  
35  
36  
37 Rosenberg, M. (1986). *Conceiving the self*. Krieger: Malabar, FL.  
38  
39 Rosenberg, M. (1989). *Society and the adolescent self-image* (revised ed.). Middletown, CT:  
40  
41 Wesleyan University Press.  
42  
43  
44 Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.  
45  
46  
47 Salaberria, K., & Echeburua, E. (1998). Long-term outcome of cognitive therapy's contribution  
48  
49 to self-exposure in vivo to the treatment of generalized social phobia. *Behavior*  
50  
51 *Modification*, 22(3), 262-284.  
52  
53  
54 Spielberger, C. D., Gorsuch, R.L., & Lushene, R.E. (1970). *Manual for the state-trait anxiety*  
55  
56 *inventory*. Palo Alto, CA: Consulting Psychologists Press.  
57  
58  
59  
60



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Spielberger, C. D., Gorsuch, R.L., Lushene, R.E., Vagg, P.R., & Jacobs, G.A. (1983). *Manual for the state-trait anxiety inventory (Form Y)*. Palo Alto, CA: Consulting Psychologists Press.

Wade, W.A., Treat, G.L. & Stuart, T.A. (1998). Transporting an empirically supported treatment for panic disorder to a service clinic setting: a benchmarking strategy. *Journal of Consulting and Clinical Psychology*, 6(2), 231-239.

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*, 54, 1063-1070.

Table 1. Robbins' Unleash the Power Within, MP3 Session Topics

Session #:	Topic:
1	Seizing Your Extraordinary Life
2	Understanding Your Emotional Triad
3	Mastering Your Communication
4	Designing Your Ultimate Success Formula
5	Creating Lasting Change
6	Embracing Change and Creating Momentum
7	Meeting Your Six Human Needs
8	Building Your Powerful Team
9	Understanding and Developing Rapport
10	Mastering Strategies for Consistent Results
11	Strategizing Your Everyday Life
12	Unleashing Your Power Within

*Table 1.* This table displays the topics for each of the twelve MP3 files associated that constitute Robbins' *Unleash the Power Within* series.

Table 2. Pre-Post Paired T-tests for Students in the Time Control Condition

Paired Samples Test

		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Pair 1	Daily Stress Inventory Pre-Post	10.91	71.51	.51	10	.62
Pair 2	PANAS Positive Affect Pre-Post	3.64	7.75	1.56	10	.15
Pair 3	PANAS Negative Affect Pre-Post	-3.18	11.03	-.96	10	.36
Pair 4	Rosenberg Self-Esteem Pre-Post	.40	3.44	.37	9	.72
Pair 5	Spielberger Trait Anxiety Pre-Post	-1.36	9.76	-.46	10	.65

*Table 1.* This table displays the results for the paired samples t-tests conducted to determine whether there were any significant pre-post differences within the time control group on each of five measures. As this was a simple time control, no pre-post differences were predicted. Consistent with this hypothesis, no significant pre-post differences were observed in this group at a Bonferroni-corrected alpha of 0.01.

Table 3. Pre-Post Paired T-tests for Students in the Robbins Condition

## Paired Samples Test

		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Pair 1	Daily Stress Inventory Pre-Post	24.20	51.24	1.056	4	.35
Pair 2	PANAS Positive Affect Pre-Post	-4.20	12.029	-.78	4	.48
Pair 3	PANAS Negative Affect Pre-Post	2.80	3.56	1.76	4	.15
Pair 4	Rosenberg Self-Esteem Pre-Post	-.80	1.92	-.93	4	.41
Pair 5	Spielberger Trait Anxiety Pre-Post	4.80	4.76	2.25	4	.087

*Table 3.* This table displays the results for the paired samples t-tests conducted to determine whether there were any significant pre-post differences within the Robbins group on each of five measures. Because this was a treatment group, significant differences on each of the five measures were predicted. No significant pre-post differences were observed in this group at a Bonferroni-corrected alpha of 0.01 though a power analysis suggest that pre-post differences in negative affect and trait anxiety would reach significance given a larger sample size.

Table 4. Pre-Post Paired T-tests for Students in the Natural Confidence Condition

Paired Samples Test

		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Pair 1	Daily Stress Inventory Pre-Post	19.70	19.69	3.16	9	.011*
Pair 2	PANAS Positive Affect Pre-Post	-5.90	3.51	-5.32	9	.000**
Pair 3	PANAS Negative Affect Pre-Post	3.00	7.97	1.19	9	.264
Pair 4	Rosenberg Self-Esteem Pre-Post	-.20	2.20	-.29	9	.780
Pair 5	Spielberger Trait Anxiety Pre-Post	10.90	8.94	3.86	9	.004**

\*\* significant at a Bonferroni-corrected alpha of 0.01

\*significant at a non-corrected alpha of 0.05

Table 4. This table displays the results for the paired samples t-tests conducted to determine whether there were any significant pre-post differences within the student Natural Confidence group on each of five measures. Because this was a treatment group, significant differences on each of the five measures were predicted. As shown, significant pre-post differences in the hypothesized directions were observed in the impact of daily stress, reported levels of positive affect, and trait anxiety.

Table 5. Pre-Post Paired T-tests for Online Purchasers of the Natural Confidence Program

## Paired Samples Test

		Paired Differences		t	df	Sig. (2-tailed)
		Mean	Std. Deviation			
Pair 1	Daily Stress Inventory Pre-Post	51.032	44.94	6.32	30	.000**
Pair 2	PANAS Positive Affect Pre-Post	-8.50	6.96	-6.91	31	.000**
Pair 3	PANAS Negative Affect Pre-Post	10.41	9.69	6.074	31	.000**
Pair 4	Rosenberg Self-Esteem Pre-Post	-1.31	2.48	-2.99	31	.005**
Pair 5	Spielberger Trait Anxiety Pre-Post	11.79	14.092	4.51	28	.000**

\*\* significant at a Bonferroni-corrected alpha of 0.01.

*Table 5.* This chart displays the results for the paired samples t-tests conducted to determine whether there were any significant pre-post differences within the online purchasers of the Natural Confidence program on each of five measures. As this was a treatment group, significant differences on each of the five measures were hypothesized. As shown, significant pre-post differences in the hypothesized directions were observed on all five measures: the impact of daily stress, reported levels of positive affect, reported levels of negative affect, self-esteem, and trait anxiety at a Bonferroni-corrected alpha of 0.01.

Table 6. Effect Size Estimates for Students in the Time Control Condition

		Pre-test	Pre-test Std.	Post-test	Post-test Std.	Cohen's <i>d</i>	Cohen's <i>d</i>	Percentage
		Mean	Deviation	Mean	Deviation	(effect size)	95% Confidence Interval	Change (%)
Pair 1	Daily Stress Inventory Pre-Post	118.09	66.97	107.18	49.18	0.19	-39.38-29.26	-9.24
Pair 2	PANAS Positive Affect Pre-Post	38.45	5.34	34.82	6.40	0.65	-2.51-4.43	-9.44
Pair 3	PANAS Negative Affect Pre-Post	19.45	5.16	22.64	9.22	-0.45	-3.50-5.00	16.40
Pair 4	Rosenberg Self-Esteem Pre-Post	28.36	5.80	26.90	4.51	0.30	-3.30-3.09	-5.15
Pair 5	Spielberger Trait Anxiety Pre-Post	44.45	10.88	45.82	11.26	-0.13	-6.56-6.52	3.08

Table 6. This table displays the results effect sizes for observed pre-post differences within the time control group on each of five measures. Recall that none of these observed changes were significant at a Bonferroni-corrected alpha level of 0.01.

Table 7. Effect Size Estimates for Students in the Robbins Condition

		Pre-test	Pre-test Std.	Post-test	Post-test Std.	Cohen's <i>d</i>	Cohen's <i>d</i>	Percentage
		Mean	Deviation	Mean	Deviation	(effect size)	95% Confidence Interval	Change (%)
Pair 1	Daily Stress Inventory Pre-Post	60.80	53.16	36.60	13.72	0.70	-45.90-12.72	-39.80
Pair 2	PANAS Positive Affect Pre-Post	37.20	9.066	41.40	5.90	-0.61	-8.56-4.56	11.29
Pair 3	PANAS Negative Affect Pre-Post	21.20	8.58	18.40	7.70	0.38	-7.14-7.13	-13.21
Pair 4	Rosenberg Self-Esteem Pre-Post	26.00	2.00	26.80	0.84	-0.58	-2.34-0.15	3.08
Pair 5	Spielberger Trait Anxiety Pre-Post	35.40	11.19	30.60	8.91	0.53	-9.28-8.34	-13.56

*Table 7.* This chart displays the results effect sizes for observed pre-post differences within student completers of the Robbins' condition on each of five measures. Recall that none of these observed changes were significant at a Bonferroni-corrected alpha level of 0.01.



Table 8. Effect Size Estimates for Students in the Natural Confidence Condition

		Pre- test Mean	Pre-test Std. Deviation	Post- test Mean	Post-test Std. Deviation	Cohen's <i>d</i> (effect size)	Cohen's <i>d</i> 95% Confidence Interval	Percentage Change (%)
Pair 1	Daily Stress Inventory Pre- Post	84.90	83.74	65.2	66.75	0.27*	-51.63- 41.65	-23.20
Pair 2	PANAS Positive Affect Pre-Post	34.10	5.97	40	6.11	-1.03**	-4.73-2.76	17.30
Pair 3	PANAS Negative Affect Pre-Post	20.00	8.81	17	5.14	0.44	-5.02-3.62	-15
Pair 4	Rosenberg Self- Esteem Pre-Post	26.60	2.55	26.8	2.044	-0.09	-1.67-1.18	0.75
Pair 5	Spielberger Trait Anxiety Pre-Post	40.90	13.50	30	6.63	1.08**	-7.29-5.19	-26.65

\*\* significant at a Bonferroni-corrected alpha level of 0.01.

\*significant at a non-corrected alpha level of 0.05.

Table 8. This chart displays the results effect sizes for observed pre-post differences within student completers of the natural confidence condition on each of five measures.

Table 9. Effect Size Estimates for Online Purchasers of the Natural Confidence Program

		Pre- test Mean	Pre-test Std. Deviation	Post- test Mean	Post-test Std. Deviation	Cohen's <i>d</i> (effect size)	Cohen's <i>d</i> 95% Confidence Interval	Percentage Change (%)
Pair 1	Daily Stress Inventory Pre- Post	98.26	58.69	47.23	34.04	1.08**	-19.58-13.06	-51.93
Pair 2	PANAS Positive Affect Pre-Post	28.25	7.82	36.75	6.65	-1.19**	-3.90-1.11	30.09
Pair 3	PANAS Negative Affect Pre-Post	25.91	10.043	15.50	5.52	1.31**	-2.17-3.22	-40.18
Pair 4	Rosenberg Self- Esteem Pre-Post	25.25	2.96	26.56	2.37	-0.50**	-1.52-0.32	5.19
Pair 5	Spielberger Trait Anxiety Pre-Post	50.45	10.045	38.66	9.10	1.25**	-2.40-1.25	-23.37

\*\* significant at a Bonferroni-corrected alpha level of 0.01.

Table 9. This chart displays the results effect sizes for observed pre-post differences within purchasers who completed the Natural Confidence condition on each of five measures.

Benchmarking Effectiveness Research

Table 10. Benchmark Effect Size Estimates for CBT-Related Interventions

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Source	Measure	Approximate Sample Size (n) <sup>1</sup>	Pre-Test Mean	Pre-Test Standard Deviation	Post-Test Mean	Post-Test Standard Deviation	Cohen's <i>d</i>	Cohen's <i>d</i> 95% Confidence Interval	Percentage Change (%)
Borkovec, Newman, Pincus, Lyte [23]	Spielberger Trait	24	57.3	8.54	41.4	9.48	1.80	-1.62-5.59	-27.69
Borkovec, Costello [24]	Spielberger Trait	18	54.7	8.4	40.0	9.5	1.69	-2.19-6.08	-26.87
Paunovic, Ost [25]	Spielberger Trait	8	63.9	14.0	46.1	13.3	1.39	-8.31-10.61	-27.86
Bryant, Moulds, Guthrie, et al. [26]	Spielberger Trait	28	56.9	12.75	43.4	16.11	0.95	-3.78-6.91	-23.73
Salaberria, Echebura [27]	Spielberger Trait	23	34	9.2	25.1	10.7	0.91	-2.85-5.28	-26.18

<sup>1</sup> Occasionally the exact sample size within the cognitive intervention group was not directly reported. In those instances the sample size was estimated according to the information provided. For example if the total number of participants was given in addition to the rates of dropout according to condition, the total number was divided evenly into the number of groups to find initial sample size; the number of participants reportedly dropping out of the relevant condition was then subtracted from this number to get the approximate sample size.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Brown, Elliott, Boardman, et al. [28]	Spielberger Trait	42	60.6	7.80	55.4	11.04	0.55	-1.81-3.89	-8.61
Brown, Elliott, Boardman, et al. [28]	Rosenberg Self-Esteem	42	14.1	6.81	19.4	7.07	-0.77	-2.83-1.37	37.20
Parrish, Cohen, Gunthert, et al. [29]	PANAS Sad Affect (five items)	54	2.94	0.86	2.41	0.91	0.60	0.37-0.85	-18.03
Parrish, Cohen, Gunthert, et al. [29]	PANAS Positive Affect (five highest items)	54	2.02	0.75	2.29	0.81	-0.35	-0.55- -0.13	13.37

*Table 10.* This chart displays the results effect sizes for observed pre-post differences within samples found during our perusal of the CBT treatment literature.