

Emotional Freedom Techniques (EFT) Reduces Intense Fears: A Partial Replication and Extension of Wells, Polglase, Andrews, Carrington, & Baker (2003)

A. Harvey Baker & Linda S. Siegel

Abstract

Wells, Polglase, Andrews, Carrington, and Baker (2003) found that Emotional Freedom Techniques (EFT; an intervention involving manual stimulation of a specific set of acupuncture points accompanied by certain verbalizations) produced greater decrease in intense fear of small animals than did a comparison condition. The present partial replication and extension assessed whether such findings reflected (a) nonspecific factors common to many forms of psychotherapy, (b) some methodological artifact (such as regression to the mean, fatigue, or the passage of time), and/or (c) therapeutic ingredients specific to EFT. Participants were randomly assigned to EFT, a supportive interview, or no-treatment control. On a majority of

the dependent variables, participants in the EFT condition showed significant decrease in fear of small animals immediately after, and again 1.38 years after, one 45-min intervention, whereas the other two conditions did not. These findings lend support for EFT's efficacy in the treatment of intense fear, but further research is needed regarding the range of problems for which EFT may be efficacious, the treatment procedures required to maintain clinical gains, the relative power of EFT compared with other established therapies, and the mechanism(s) that produce EFT's effects.

Keywords: anxiety reduction; desensitization; phobias; Emotional Freedom Techniques; EFT; energy psychology

A. Harvey Baker, PhD, Department of Psychology, Queens College of the City University of New York; **Linda S. Siegel**, MEd, Lawrenceville, New Jersey. Corresponding concerning this article should be addressed to Linda S. Siegel, 14 Fairway Court, Lawrenceville, NJ 08648. Email: activist@shaw.ca.
Disclosures: The authors declare that they have no conflicts of interest.

medicine and spiritual practices; (d) publicized demonstrations claiming effectiveness of some of these interventions; and (e) very few supportive, peer-reviewed investigations that included appropriate control conditions.

Emotional Freedom Techniques (EFT) is a member of a group of therapeutic modalities sometimes referred to as Meridian-Based Therapies (Wells, Polglase, Andrews, Carrington, & Baker, 2003). These approaches have emerged in the past two decades and are associated with (a) clinical procedures that involve physical interventions (e.g., tapping locations on one's own body) that make use of meridian acupuncture points; (b) claims of extraordinary, rapid, and dramatic psychological and/or behavioral change (e.g., Callahan & Callahan, 1997); (c) the purported mediating process(es) alleged to have effected these changes are based on traditional Chinese

The present study focused on whether EFT is efficacious in the treatment of intense, irrational fears. Two prior studies have focused on this issue. Using phobic participants, a 30-min treatment intervention, multiple self-report measures, and a behavioral dependent measure, Wells et al. (2003) reported significantly greater decrease in fear of small animals for EFT (which involved certain verbal statements plus tapping on meridian end-points) than for a Diaphragmatic Breathing intervention (DB; which involved a parallel set of verbal statements and diaphragmatic breathing), both immediately after 30 min of treatment and also 6 to 9 months after that treatment session. Waite and Holder (2003), using undergraduate volunteers, a 2- to 3-min intervention interval, and a single self-report measure, found that EFT and two variations

that each substituted a newly devised procedure for one of the elements of EFT led to equal and significant decreases in fear, whereas a control condition showed no significant decrease. Because the two newly devised variations of EFT produced improvement equal to that of EFT, they speculated that the mechanism(s) that mediated this observed efficacy was not new or uniquely associated with EFT but simply involved processes long known to be associated with psychotherapy, such as distraction or systematic desensitization.

In a review of the mental health research on Meridian-Based Interventions, Feinstein (2007) pointed out that the two newly devised conditions used by Waite and Holder (2003) may have unwittingly stimulated acupuncture points and thus may have had a therapeutic impact. He argued that Waite and Holder's conclusion that stimulating acupuncture points is not an active ingredient in EFT treatments may thus be erroneous; in any case, it seems premature. In a companion article to the present study (Baker, Carrington, & Putlin, 2007), we argued that Waite and Holder (a) ignored methodological limitations of their study (the decrease in fear observed for their three intervention conditions may simply have reflected demand characteristics or response expectancy) and (b) failed to consider alternative hypotheses. Until either the particular features of their three efficacious conditions that led to decrease in fear are identified or the action mechanism of this decrease is delineated, or both, it is premature to dismiss EFT as involving no new or interesting process.

The present study (which added both a comparison placebo-control condition involving treatment ingredients common to many forms of psychotherapy, see below, and a no-treatment control condition) sought to shed light on Wells et al.'s (2003) findings. Specifically, the goal was to determine whether the efficacy of EFT demonstrated by Wells et al. would be maintained under more stringent experimental controls.

Background: Choice of EFT

Callahan (e.g., Callahan & Callahan, 1997) originally developed a Meridian-Based Intervention, which he called Thought Field Therapy (TFT), based on observations of a remarkably rapid phobia cure with a patient in his clinical psychology practice after serendipitously applying techniques he had been studying derived from

the fields of applied kinesiology and acupuncture. According to Callahan, one must first (a) diagnose the patient's problem using procedures such as manually testing the relative strength of muscles associated with acupuncture meridians and then (b) apply an appropriate algorithm to it, which involves tapping on a precise sequence of certain acupuncture meridian end points (end points that are associated with that particular diagnosis) on the patient's body while repeating certain verbal phrases. EFT was developed by one of Callahan's earliest students, Gary Craig (1995, 1999), who—based on his own and others' reported clinical experience—concluded that (a) use of a precise sequence of tapping locations was not necessary and (b) rather than tapping a different set of meridian end points for each different diagnostic condition, one can achieve the same results using one common generic set of meridian end points for all conditions. A recent study involving random assignment of participants to one of two conditions (Pignotti, 2005b) by a publicly disaffected former colleague of Callahan's (Pignotti, 2007) found equally high success rates using Callahan's precise sequences of tapping locations with one group and random sequences for the other group. Because Pignotti (2005b) found these results surprising and disconcerting, she delayed several years before publishing these findings, which were clearly not due to the researcher-therapist's expectations.

A primary difference between EFT and TFT is the latter's maintenance that a precise sequence of points to be tapped is necessary for treatment to succeed. Not only has there been no evidence supporting Callahan's claim that sequence is relevant, but Pignotti's (2005b) study also provided empirical evidence consistent with the view that sequence is irrelevant. Because TFT advocates have not met the burden of proof on this issue (see Pignotti, 2005a), we conclude that it is plausible to follow Wells et al. (2003) in choosing EFT as the focus of our study because EFT is a simpler and more parsimonious procedure.

Scope of the Present Study

The purpose of this study was to investigate whether EFT, as reported by Wells et al. (2003), is efficacious for reducing fear of small animals. Comparison of EFT with other treatments that have been shown to be efficacious with fearful individuals, such as Therapist-Directed Exposure Therapy

(e.g., Ost, 1996), will be necessary if the efficacy of EFT is established. Delineating the mechanisms underlying any therapeutic effect of EFT is also beyond the scope of this study. In our own formulation, the putative acupuncture meridian system is not the only viable explanatory hypothesis. Our working assumptions, in fact, are that (a) some aspect of EFT induces a relaxation response (see **Discussion** section) and that (b) the coupling of relaxation with attention focused on the feared object in EFT brings about a novel and rapid form of desensitization or counterconditioning.

The Present Study

The present study built on the work of Wells et al. (2003), in which EFT was contrasted to an experimentally devised comparison-placebo condition, systematic DB. Both conditions in that study had several features in common: (a) equal focus on the feared animal throughout each treatment procedure by repeatedly stating a reminder phrase (e.g., “my fear of rats”) or focusing on an image representing the fear; (b) use of procedures presumed to induce relaxation and reduce arousal (i.e., tapping or DB); (c) equal brief stops in the periods of tapping or DB; (d) equal frequency of assessing level of fear (subjective units of distress; SUD); and (e) equivalent address to each separate aspect or issue related to the problem by additional rounds of EFT or DB.

Because the design of the Wells et al. (2003) study specified equality on the several factors just listed, so that DB was as similar as possible to EFT, it is possible to draw certain conclusions from the finding that participants in the EFT condition showed significantly greater improvement than did those in DB: (1) EFT involves frequent assessment of SUD (fear) level. It can be argued that this creates a demand characteristic that results in participants reporting less fear after treatment to please the experimenter. This possibility cannot account for the greater improvement with EFT in Wells et al., as there was equal frequency of assessment of SUD level for EFT and for DB. (2) EFT involves frequent focusing on the feared object. This can be reasonably construed as a form of exposure, which is known to reduce fear. However, differences in exposure cannot account for participants' greater improvement for EFT compared with DB in the Wells et al. study because there was *equal* exposure for each condition. (3)

Because the two conditions were designed to be as similar as possible—except for the specific physical treatment intervention (tapping plus verbalizations for EFT vs. diaphragmatic breathing plus verbalizations for DB) aimed at increasing relaxation/reducing arousal—EFT's greater improvement cannot be explained as reflecting greater total duration of treatment, greater frequency of treatment intervals, or greater attention directed by the experimenter to the participant.

One hypothesis that the Wells et al. (2003) study did not address was the “common factors” hypothesis, to which Waite and Holder (2003) alluded in speculating about their findings. “Common factors” refers to features that are presumed to be present in most forms of psychotherapy (e.g., contact with a supportive and caring person, focused attention on the presenting problem, effects of expectancy and hope, demand characteristics, suggestion; Behar & Borkovec, 2003). After reviewing the strengths and limitations of a common (nonspecific) factors or placebo comparison design, Behar and Borkovec (2003; pp. 215–217) commented that

If one does decide to use a common factors control group, perhaps the best type of choice involves the provision of a nondirective or supportive listening condition . . . [this condition] thus represents a significant core of what is often meant by the phrase *common factors*. (p. 217)

We therefore included a supportive interview (attention-placebo) comparison condition that met two criteria: It appeared to constitute a plausible treatment intervention for reducing fear and participants did not a priori expect less help from it than they did for EFT. Specifically, we used a supportive interview with active listening components (Gordon, 1974), an approach derived from the work of Carl Rogers. Rogerian counseling had been used on panic and agoraphobia symptoms with a positive outcome in a study involving 50 or more sessions (Teusch, Bohme, & Gaspar, 1997), indicating that a supportive interview approach could be regarded as a plausible treatment to reduce fears.

In contrast to Wells et al. (2003), who employed two conditions involving exposure to the feared animal, our use of a supportive interview compared EFT (a treatment that involved exposure) with an interview intervention procedure (a

treatment that contained a number of nonspecific therapeutic components but offered little in terms of specific active ingredients for overcoming fear).

Another way in which the present study differed from Wells et al. (2003) was in its inclusion of a no-treatment control condition in which participants were asked to sit in the laboratory for the same duration as participants in the EFT condition, either studying or reading magazines provided for them, none of which dealt with fears. Wells et al. had found that there was some significant improvement from pre- to postintervention for each of their two conditions (EFT and DB). This variance *shared* by EFT and DB could have been due to such causes as regression to the mean, fatigue, the passage of time, or practice effects associated with repeated measurement. The Wells et al. (2003) study was not designed to clarify these additional issues. The present study sought to fill this gap by inclusion of a no-treatment condition. Ernst and Resch (1995) have, in fact, argued for including such a no-treatment control condition in order to accurately distinguish between true placebo effects and perceived ones. Our working hypothesis was that no effects would be observed for the no-treatment condition. Results from a pilot no-treatment condition were consistent with this hypothesis.

Method

Participants

Volunteers with an intense fear of rats, spiders, or water bugs were recruited through newspaper articles, flyers, and a television news story. Participants had to (a) be at least 18 years old; (b) report that they were very afraid of rats, spiders, and/or water bugs; and (c) be willing to come to our laboratory (some participants traveled 4 hr roundtrip). In addition, as in Wells et al. (2003), (d) any participant able to stand (during the pretest) at the closest point to the feared animal on a Behavioral Approach Task (BAT) and still report a SUD level (defined below) of less than 5 was also excluded from the study. Participants were not paid for participation in the initial phase of the study but were paid for returning for follow-up.

Volunteers were screened with a highly structured telephone interview patterned closely after the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., *DSM-IV*; American Psychiatric Association, 1994) criteria for specific phobia, resulting in 39 people who reported an intense fear

of rats, spiders, and/or water bugs and were willing to participate. Eight, who met the first three criteria above, were subsequently ruled out based primarily on criterion d, above, leaving 31 participants (2 male, 29 female; with a mean age of 34.9 years, ranging from 18 to 68 years) in the study. This predominance of women is consistent with earlier research into fears and phobias (Ost, 1987). Of these 31 participants, seven met all six *DSM-IV* criteria for a diagnosis of specific phobia in adults, nine met five of the six *DSM-IV* criteria for a diagnosis of specific phobia in adults (of these nine, two did not regard their fear as excessive or unreasonable and seven did not feel that their fear interfered with their lives). The remaining 15 participants reported great fear but did not meet *DSM-IV* criteria for a diagnosis of specific phobia. All participants signed a consent form approved by the Queens College internal review board.

Design

Participants were randomly assigned to EFT ($n = 11$), supportive interview ($n = 10$), or no-treatment control ($n = 10$) conditions. Two experimenters worked with each participant. One experimenter, "E_d," collected all data before the intervention (during the 20-min pretest) and after the intervention (during the 15-min posttest). Following the admonition of Barrios (1978) regarding the need for masked data collection when studying fear of small animals, we took stringent steps to keep E_d blind to treatment condition. Thus, no one was assigned to a condition until all pretest data were collected and E_d was absent. The second experimenter, "E_i," administered the intervention (which took 45 min) and was kept blind regarding all data. Then E_d returned and performed the posttest assessments.

Measures

Fear Questionnaire (FQ). A modified version of the Marks and Matthews (1979) Brief Standard Self-Rating for Phobic Patients was used. This included three of their four measures: Main Target Phobia, Global Phobia, and Anxiety-Depression. They reported that these measures had good test-retest reliability (ranging from .79 to .93) and that they were sensitive to clinical improvement following treatment of phobic patients.

Fear of Specific Animal Questionnaire (FOSAQ). This is an 11-item questionnaire (Baker, Quitchon, & Putilin, 2007) with a 0 to 8 Likert-type

response scale. There were three parallel forms, each assessing fear of a specified animal (here, rat, spider, or water bug, according to which of these participant most feared). Baker, Quiatchon, and Putilin (2007) presented reliability and validity findings with the FOSAQ that indicate that it is legitimate to combine data from such parallel forms, thus allowing a researcher to include in a single study people afraid of a variety of small animals and insects. They found that this instrument displayed high internal consistency ($r = .94$) and test-retest reliability (correlations ranged from .74 to .81) as well as significant validity correlations, including the correlation between FOSAQ scores and SUD level when participants imagined the feared animal ($r = .60, p < .001$) and the correlation between FOSAQ scores and how close one was willing to get to the feared animal during the BAT ($r = -.44, p < .005$).

SUD Scale when imaging the feared animal (SUD Imagined). Participants rated how they felt on an 11-point SUD Scale (Wolpe, 1958) ranging from 0 (*no fear or distress at all*) to 10 (*intense fear or distress*) while imagining their feared animal. They were to respond in terms of how they felt “at this moment” and not in terms of how they imagined they would feel in a real situation. According to Thyer, Papsdorf, Davis, and Vallecorsa (1984), such physiological indicators of anxiety as pulse rate ($r = .39, p < .05$) and peripheral vasoconstriction ($r = -.84, p < .01$) show concurrent validity with SUD scores.

Pulse rate. Each participant’s pulse rate was assessed using a Cateye Model PL 6000 Heartbeat Counter, with one part attached to an ear lobe and another part to the upper arm. Participants’ pulse rate was measured before taking the FOSAQ and FQ questionnaires (Pulse Rate Before BAT) and at each step in the BAT (Pulse Rate During BAT), during both the pretest and the posttest.

BAT. The BAT is an in vivo performance test (administered during the pretest and posttest) that assesses the participant’s intensity of fear when faced with the feared animal. (The treatments, described below, were not in vivo.) At each decision point, participants were asked whether they were willing to move closer to the feared animal. There were nine decision points on this task: (1) while seated in the interview room, the participant was informed that there was a [whichever animal the participant most feared] that was visible but in a totally secure container in the next room, about 20 ft (6.10 m) away. Participant was then asked “Are

you willing to allow me to open the door to this adjacent room?” If participant said “Yes,” the door was opened and the person stood up and moved near the open doorway, with the experimenter standing next to her or him. (2) Participant was asked whether she or he was willing to step into the doorway of the room. (3) Experimenter called attention to the fact that there was a series of six lines (each 27.5 in. [69.9 cm] apart on the floor) and asked whether participant was willing to take a step onto the first line. The next five points involved asking participant “Are you willing to take another step to the next line?” Responding in the affirmative at the 8th point would bring the participant adjacent to the container with the feared animal in it. The ninth point involved asking whether she or he was willing to put one hand on the external side of the container. (This did not involve touching the feared animal.) At each of these nine decision points, the task ended if participant replied “No.” No pressure of any kind was put on the participant to say “Yes” at any point in this task. Participant’s SUD level and pulse rate were recorded at each decision point. E_d kept reminding the participant to look directly at the feared animal throughout the task. In a separate sample, test-retest reliability was assessed before and after a 1-hr interval and showed $r(33) = .95, p < .001$.

SUD During BAT. At each measurement point on the BAT, the SUD rating was recorded along with the pulse rate. An SUD average was calculated for each participant by averaging SUD scores for each point taken in common between pretest and posttest for that participant’s BAT (i.e., if participant took five steps on the pretest and seven steps on the posttest, only the first five SUD scores from pretest and posttest would be averaged). The higher the SUD level at each point in the BAT, the more distress the participant felt upon approaching the feared animal.

Autonomic Response Questionnaire. This was a modified version of the Autonomic Perception Questionnaire (Kern & MacDonald, 1980). It was administered right after the BAT where E_d asked participants whether they experienced any of the typical symptoms of anxiety (e.g., lump in throat, dry mouth) during the BAT. Participants were also asked to indicate any other physical symptoms they experienced while approaching the feared animal. The score was the total number of different autonomic-type symptoms the participant specified.

The FQ, SUD Imagined, the two pulse rate measures, the BAT, and SUD During BAT were used in the Wells et al. (2003) study. The 11-item FOSAQ was developed for use in the present study in part to have a more extensive measure of the degree to which a participant feared a given animal. The Autonomic Response Questionnaire was employed to obtain a measure of physiological responses in the presence of the feared animal as subjectively perceived by participant.

Procedure

Upon arrival at the laboratory, each participant was greeted by the E_d , who led her or him into the small interview room within the laboratory. All participants were told never to inform E_d about which condition they eventually participated in. Then audiotape recording began (used primarily to refer to later if there were any uncertainties) and pretest measures were administered. The door separating the interview room from the room with the feared animal was kept shut except during the BAT.

After the pretest E_d left and E_i entered the interview room, opened a sealed envelope, and informed the participant about the treatment condition to which she or he was randomly assigned. E_i described the treatment and asked participants to rate the degree to which they expected that this condition would help alleviate their fear. This was recorded on an 8-point scale, that ranged from *will not help at all* to *will help me to completely get over my fear*. E_i then delivered the appropriate intervention.

After the 45-min intervention E_i left the room to get E_d . Both experimenters were directed not to exchange information about the condition to which the participant had been assigned. Posttest measures were administered by E_d . The participant was then reminded that she or he had agreed during the telephone interview to be contacted regarding participation in a possible follow-up study. No instructions were given regarding using EFT at home in the interval between the original intervention and the possible follow-up, nor was the participant encouraged to use EFT on her or his own.

Treatment

The no-treatment control, supportive interview, or EFT treatment was given without delay after the pretest measures. All treatment condi-

tions lasted 45 min, and immediately after the treatment, the posttest was administered.

EFT condition. This was based on the EFT “Basic Recipe” outlined by its developer (Craig, 1995, 1999). It consisted of having the participant imitate E_i in tapping various acupuncture meridian endpoints on his or her own body while maintaining a strong focus on the fear by repeating a brief reminder phrase; for example, “fear of rats” could be the reminder phrase for “I am afraid of rats.” This reminder phrase was stated once at every acupuncture meridian endpoint that was being tapped.

A single round of EFT (Craig, 1995) involves the following:

Part 1. The intensity of whatever single negative affect was to be worked on during the EFT intervention was assessed on a 0- to 10-point SUD scale. (SUD data obtained intermittently during the intervention were recorded but not used for analysis.)

Part 2. E_i stated a standard setup phrase (e. g., “Even though I have this fear of rats, I deeply and completely accept myself”); the participant then repeated this phrase three times and, while doing so, rubbed two pressure points located on the upper left and right portions of the chest, one with each hand. These spots were rubbed firmly with the index and middle fingers of each hand while the participant repeated the phrase.

Part 3. Participant then gently tapped five to seven times on five different acupuncture meridian endpoints on the face and two on the upper torso.

Part 4. Participant similarly tapped four different acupoints near the front ends of the fingertips of one hand and a fifth point on the side of the hand. A short reminder phrase (e.g., “fear of rats”) was recited once at each location, to focus on the issue, during Parts 3 and 4.

Part 5. Then, while tapping very rapidly three to five times per second with three fingertips of one hand on the back of the other hand at a location halfway between and 0.5- in. back from the knuckles between the ring finger and the little finger, the participant carried out a series of simple tasks, such as close eyes; open eyes; look down to the floor while holding the head steady with eyes directed to the right, to the left, rotated clockwise, rotated counterclockwise; hum a few notes from a song (“Happy Birthday”); count from one to five out loud; hum again.

Parts 6 and 7. These involved repeating Parts 3 and 4, respectively.

Part 8. This was a repeat of Part 1.

A 45-min session of EFT could therefore involve 15 to 30 or more such rounds.

If participant made progress on the first three rounds (that is, if the SUD level decreased), a shortened EFT procedure was then followed, as prescribed by Craig (1995, 1999), involving only Parts 1, 2, two administrations of Part 3, and Part 1 above repeated. On the other hand, if the participant's SUD level did not decrease, a variety of special interventions were used (e.g., the participant was asked to speak in a very loud, dramatic voice while doing Part 2; see Craig, 1995, 1999).

Treatment focused on the general issue of fear of whichever animal the participant feared until the SUD level was reduced to 2 or less. Then the focus shifted to specific aspects or separate issues related to the feared animal (e.g., a rat's tail, the sound of a rat running, being bitten by a rat). These specific aspects were worked on within the 45-min time limit.

Every aspect that applied to a participant's fear was separately subjected to the tapping procedure outlined earlier. E_i was trained to assist participants in identifying all significant aspects of the feared animal.

Supportive interview. The supportive interview involved the use of empathy, compassion, and active listening. The procedure was designed to make someone who was being interviewed feel positive about the experience because he or she really felt "fully listened to." E_i always followed the participants' lead in discussing their fear of rats, spiders, or water bugs. Some participants exhausted what they wanted to say about their fear before the 45 min had passed. The experimenter then directed the interview to one or more of a series of topics that people generally enjoy talking about, for example, "A recent vacation which you took." This was done so that all interviews lasted the full 45 min.

No-treatment control. The participant read a magazine (which involved no content related to fears) or studied in the same interview room between the pre- and posttest.

Statistical Procedures

Analyses. We performed eight separate 3×2 analyses of variance (ANOVAS) to assess the effects of the three levels of the between-groups treatment variable (EFT vs. supportive interview

vs. no-treatment control) and the two levels of the repeated measures variable of time (pretest vs. posttest), with one ANOVA on each of the eight dependent variables. If EFT showed significantly greater improvement from pretest to posttest than did one or both of the other two between-groups conditions, this would result in a significant Treatment \times Time interaction. Whenever one was observed, we pinpointed exactly which of the three conditions differed significantly from the others using a follow-up 2×2 Treatment \times Time ANOVA as specified below. All statistical tests were conducted using the SPSS Base 10.0 Statistical Package (2002).

Graphic presentation. For graphic presentation only, the data were reduced by subtracting the posttest mean from the corresponding pretest mean for each dependent variable, thus allowing for direct depiction of degree of improvement from pretest to posttest and facilitating comparison with similar difference-score graphs in Wells et al. (2003). These three difference scores for each dependent variable were not used in statistical analyses. Rather, for each dependent variable, the six scores (3 conditions $\times 2$ times) were analyzed using 3×2 ANOVAs.

Results

We first consider initial comparability of conditions, then assess evidence of EFT's efficacy, and finally consider the role of participants' expectations.

Comparability of Conditions at the Time of the Pretest

Age. The 1×3 ANOVA with age as the dependent variable and conditions as the independent variable was not significant, $F(2, 30) = 0.42$, *ns*. Ages were EFT, 34.1 years ($SD = 16.9$); supportive interview, 38.1 years ($SD = 13.5$); and no treatment, 32.5 years ($SD = 10.4$).

Dependent variables. A 1×3 ANOVA was conducted on the pretest data for each of the eight dependent measures with condition as the independent variable. If random assignment of participants was successful, none of these should be significant; indeed, for seven of these, results were not significant, $F_s(2, 25-28)^1$ ranged from 0.29 to 1.98. For one variable, Pulse Rate Before BAT, there was a significant outcome, $F(2, 24) = 4.90$, $p < .02$. Presumably, this reflected a

chance outcome, given the random assignment of participants. To ensure that this pretest difference between conditions did not distort the results from the 3×2 ANOVA for this measure, the ANOVA was supplemented by a 1×3 analysis of covariance (ANCOVA) with the pretest score of Pulse Rate Before BAT as the covariate.

Effects of Treatment Conditions

Parallel results from the 3×2 ANOVAS on five dependent measures. Parallel results occurred on the FQ, the SUD Imagined, the FOSAQ, SUD During BAT, and the BAT, respectively, $F_s(2, 25-28)^{ii} = 6.78, 28.19, 24.96, 10.00,$ and $6.20; p_s < .004, .001, .001, .001,$ and $.006,$ respectively. These significant interactions indicate that the outcomes for the condition with the largest amounts of improvement (EFT in all instances) differed significantly from the outcomes for the condition with the smallest amounts of improvement (always no treatment). To assess whether the condition with an intermediate amount of change (always the supportive interview) differed significantly from the outcome of either of the other two conditions, we conducted two follow-up 2×2 Condition \times Treatment ANOVAS for each of the

five dependent measures. The five 2×2 ANOVAs comparing EFT and the supportive interview conditions showed significant interactions, $F_s(1, 17-19)^{ii} = 6.84, 24.20, 23.94, 13.52,$ and $5.65; p_s < .02, .001, .001, .002,$ and $.03,$ respectively: EFT participants showed significantly greater improvement than did the supportive interview participants (for whom little if any change was observed). Results from the five 2×2 ANOVAs comparing supportive interview and no treatment failed to show any significant interactions, $F_s(1, 16-18)^{ii} = 0.00, 2.66, 1.79, 0.01,$ and $0.62,$ respectively; all p_s are nonsignificant—and both of these conditions showed little if any improvement.

Figure 1 further elucidates these findings. Improvement on fear measures would manifest itself as a decrease on the posttest compared with the pretest results, while improvement on the BAT would increase. This is, indeed, what was found. EFT participants showed striking decrease in fear at the posttest for each of the four fear measures, whereas the other two condition participants showed little if any change (see Figs. 1a–1d). On the BAT, EFT participants showed increase at the posttest, whereas the other two condition participants showed slight decrease (see Fig. 1e). In sum, EFT participants showed substantial improvement

Table 1. Effects of Variation in Type of Intervention for the First Five Dependent Variables: Pretest Versus Posttest

Variable	Intervention	Pretest	Posttest
		<i>M (SD)</i>	<i>M (SD)</i>
Fear Questionnaire	EFT	25.27 (13.59)	12.36 (8.08)
	Interview	28.50 (12.38)	28.00 (13.39)
	No treatment	19.40 (4.11)	18.90 (4.43)
SUD Imagined	EFT	8.08 (1.03)	3.40 (2.27)
	Interview	8.90 (1.85)	8.00 (2.45)
	No treatment	7.30 (2.94)	7.40 (2.76)
FOSAQ	EFT	6.15 (0.85)	2.88 (1.67)
	Interview	6.12 (1.48)	5.90 (1.60)
	No treatment	6.46 (0.96)	6.45 (1.04)
SUD During BAT	EFT	6.23 (1.74)	1.91 (1.67)
	Interview	7.01 (1.90)	6.34 (1.90)
	No treatment	6.11 (2.45)	5.50 (3.11)
BAT	EFT	5.45 (3.36)	7.73 (1.42)
	Interview	6.50 (2.84)	6.40 (2.76)
	No treatment	5.90 (3.28)	5.40 (3.57)

Note. EFT = Emotional Freedom Techniques; SUD = subjective units of distress; FOSAQ = Fear of Specific Animal Questionnaire; BAT = Behavioral Approach Task.

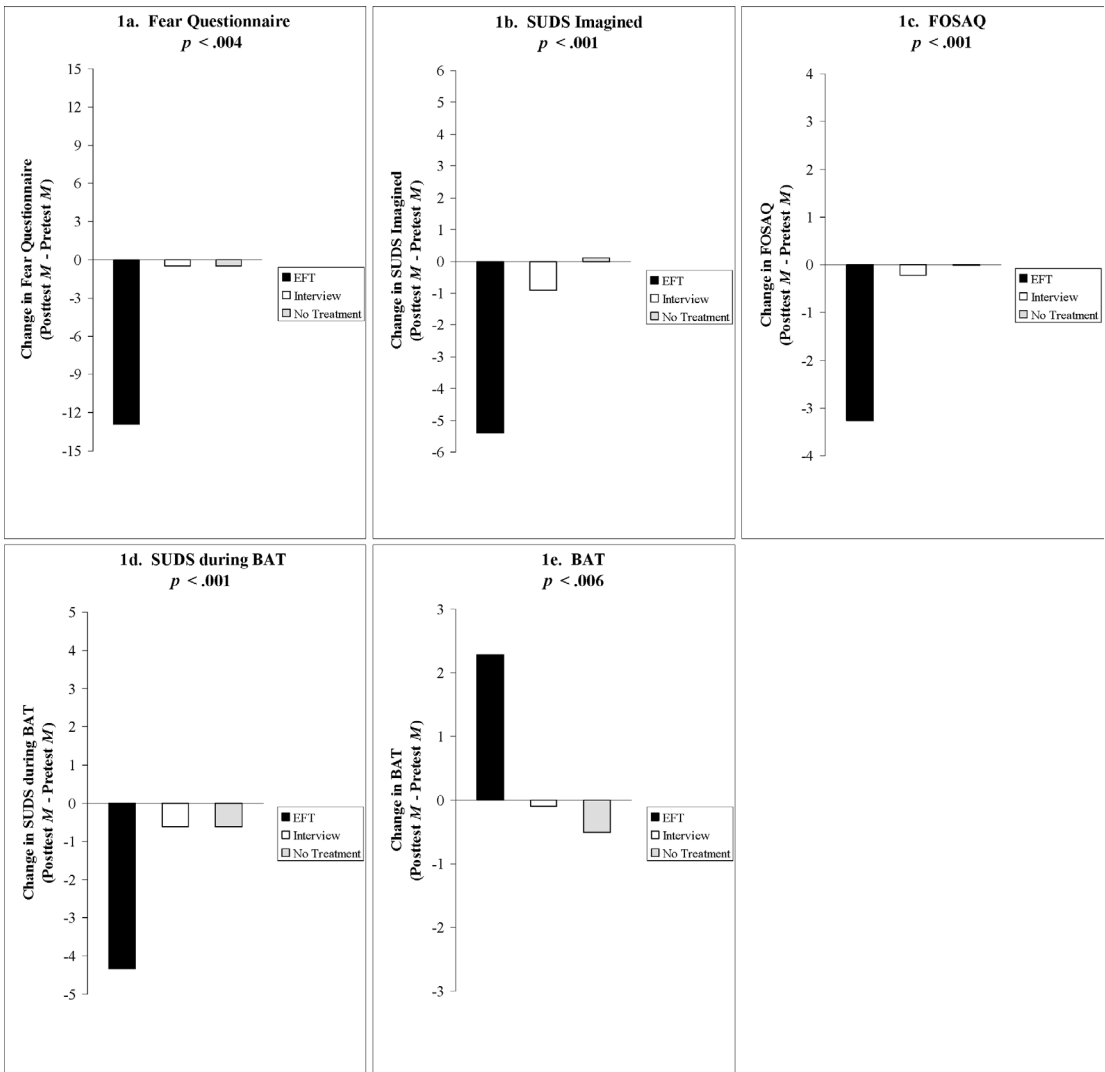


Figure 1. Immediate effects of EFT compared with supportive interview and no-treatment conditions. EFT = Emotional Freedom Techniques; SUDS = subjective units of distress; FOSAQ = Fear of Specific Animal Questionnaire; BAT = Behavioral Approach Task.

on each of these five measures, whereas the other two condition participants showed very little change and did not differ significantly from each other.

Autonomic reactivity. Although EFT showed larger decrease on this score from pretest to posttest than did the other conditions, this change was not significant, $F(2, 24) = 4.17, ns$.

Pulse Rate Before BAT. The 3×2 interaction was not significant, $F(2, 20) = 2.79, ns$. Because this was the one dependent variable on which the three conditions differed at the time of the pretest, an ANCOVA was also undertaken, with the pretest

scores for Pulse Rate Before BAT as the covariate. This 1×3 ANCOVA on the posttest scores for this dependent measure was also not significant, $F(2, 17) = 2.21, ns$, indicating no effect of condition upon these posttest scores.

Pulse Rate During BAT. The results for the 3×2 interaction for this variable failed to show significance, $F(2, 25) = 0.57, ns$, indicating no effect of conditions.

Effect size and explained variance. Table 2 presents data regarding how large the observed effects are. Of primary interest here are the 3×2 Condition \times Time interactions. The third column

Table 2. Size of Findings for Immediate Effects of EFT.

Type of variable Specific task	Condition × Time		EFT: pretest vs. posttest		Improvement on EFT % (pretest to posttest)
	(η^2)	Explained variance %	Cohen's <i>d</i>	Explained variance %	
Behavioral					
BAT	.31	31	0.78	13	42
Emotional					
SUD Imagined	.68	68	2.33	58	61
SUD during BAT	.42	42	2.46	60	69
Fear of Specific Animal	.67	67	1.78	44	53
Fear Questionnaire	.33	33	1.12	24	51
Physiological					
Autonomic Reactivity	.14	14	1.62	40	69
Pulse Rate Before BAT	.22	22	0.08	0	2
Pulse Rate During BAT	.04	4	1.38	32	13
Mean	.35	35	1.44	34	45

Note: EFT = Emotional Freedom Techniques; SUD = subjective units of distress; BAT = Behavioral Approach Task.

in Table 2 presents the relevant values of η^2 , the usual measure of effect size associated with ANOVAS. The mean value across the eight dependent values of η^2 is .35.

The fourth column in Table 2 presents the corresponding values of the percentage of explained variance, indicating a mean value of 35%.

The remaining three columns of Table 2 are of secondary interest. Some researchers who want to explore the effects of EFT compared with conditions different from those used in the present study may find it valuable to know the effect size and explained variance within the EFT condition itself (i.e., EFT pretest vs. posttest). Across the eight dependent variables, the mean value of Cohen's (1988) *d* was 1.44, corresponding to 34% explained variance.

Practitioners, however, are primarily interested in "clinical significance"—that is, whether the effect involved is large enough to be noticed in everyday life. We thought it helpful therefore to report, in the extreme right column, the results within the EFT condition of the percentage improvement from pretest to posttest. As may be noted, the mean value here is 45%.

Participant's Expectations

Expectation may affect a wide range of psychological phenomena (e.g., Kirsch, 1985) as well

as psychotherapy outcome (e.g., Weinberger & Eig, 1999). Therefore, any study assessing the effects of a treatment technique must be alert to the possible role of participant's expectations. This took on additional weight here because the majority of participants were recruited in response to a television news report. Although the duration between the time of the news report and the participation of these people was at least 2 months (usually more), it was still important to assess whether they expected EFT to help them with their fears more than the comparison placebo condition. If this were the case, it could in part or whole account for any findings.

Right after participants were told to which condition they had been assigned and given a detailed description of that condition, they were asked to rate the degree to which they expected this condition to help. These ratings were on a scale that ranged from 0 (*it will not help at all*), to 2 (*it will help slightly*), 4 (*moderately*), 6 (*greatly*), or 8 (*It will help me to completely get over my fear*).

Results of a 1×3 ANOVA indicated that the expectation scores for EFT ($M = 3.64$, $SD = 2.54$) differed significantly from those of the no-treatment condition ($M = 0.0$, $SD = 0.0$), $F(2, 26) = 8.65$, $p < .001$. Results of a 1×2 follow-up ANOVA indicated that the expectation scores for the supportive interview ($M = 2.20$, $SD = 1.75$) differed significantly from those of no treatment,

$F(1, 16) = 12.47, p < .003$. Results of an additional 1×2 follow-up ANOVA showed no significant difference between EFT and the supportive interview conditions, $F(1, 19) = 2.23, ns$.

In addition, between-conditions differences in results did not parallel between-conditions differences in participants' expectation: No significant difference in outcome was observed between the supportive interview and the no-treatment control conditions, but significant difference in terms of level of expectation was observed between these two conditions. In contrast, a significant difference in outcomes was observed between supportive interview and EFT conditions, but no significant difference in terms of level of expectation was observed between these two conditions.

Because the issue of expectation is a very important one, it seemed advisable to adopt a cautious approach here: Although EFT and supportive interview did not differ significantly in terms of participants' expectation, the observed difference was suggestive (i.e., it reached the $p < .15$ level). For each of the five dependent measures with significant results on the ANOVAS (i.e., the FQ, the SUD Imagined, the FOSAQ, SUD During BAT, and the BAT), we therefore undertook a 3×2 ANCOVA, with all three conditions as the independent group variable and with participants' expectations as the covariate. Here the results were $F_s(2, 22-25)^{ii} = 3.30, 12.65, 13.65, 5.23, \text{ and } 4.09$, with $ps = .05, < .001, .001, .02, \text{ and } .03$, respectively, and with the corresponding values of $\eta^2 = .21, .51, .55, .30, \text{ and } .25$, respectively (mean η^2 here is .36). We also compared the outcomes for EFT and the supportive interview on the same five dependent measures in 2×2 ANCOVAs with the following results: $F_s(1, 16-18)^{ii} = 3.93, 18.94, 20.01, 9.72, \text{ and } 5.04$, with $ps = .06, < .001, .001, .006, \text{ and } .04$, respectively, and with the corresponding values of $\eta^2 = .18, .53, .56, .35, \text{ and } .22$, respectively (mean η^2 here is .37). These results indicate that even when we institute statistical control of a participant's a priori expectations (by using ANCOVA), there are still statistically significant positive findings for EFT, which offers support of EFT's efficacy.

The particular one-item measure of expectation available here has been used, to our best knowledge, in only one prior study: Wells et al. (2003). If this one-item measure does not validly assess participants' expectations, then the ANCOVAs will not adequately control for the role of a

priori expectations. Some time after collecting the data for the present study, we learned of a fairly widely used measure of expectation devised by Borkovec and Nau (1972). Since this latter measure includes a focus on the degree of help that participants expect from the treatment they are to receive, we expect there is a substantial correlation between the Borkovec and Nau measure and the one-item measure employed here. We compared both measures in three subsequent (unpublished) samples, each involving a variant of EFT. The correlations between the two measures were $r_s(25, 20, 62) = .84, .86, \text{ and } .61$. The average correlation across these three samples was 0.74., indicating that the one-item measure is assessing something very similar to Borkovec and Nau's measure.

Follow-Up Study: Method

Duration of Time Until Follow-Up

The mean duration between initial testing and follow-up was 1.38 years ($SD = 0.69$).

Participants

Of the 31 participants in the original study, one had moved out of state and two could not be located. Of the 28 potential participants for the follow-up study, 24 participated. In light of the difficulties getting participants to return for the follow-up in the Wells et al. (2003) study, we obtained a small grant that allowed us to pay each participant in this follow-up study \$125, which contributed substantially to our high rate of follow-up participation.

Procedures

The same procedures, questionnaires, and tasks used in the pretest were readministered at follow-up.

Analyses

First, we assessed the amount of change from pretest to follow-up for the EFT condition compared with that for the two comparison conditions by using a repeated measures Condition \times Time of Testing (pretest vs. follow-up) ANOVA. Second, to determine the degree to which the significant effects observed for EFT at the posttest were maintained, we considered (a) whether the effects were maintained (follow-up vs. pretest difference); (b)

Table 3. Size of Findings for Follow-Up Effects of EFT.

Type of variable	Specific task	Condition × Time		EFT: pretest vs. posttest		Improvement on EFT % (pretest to follow-up)
		h2	Explained variance %	Cohen's d	Explained variance %	
Behavioral	BAT	.09	9	0.38	3	21
Emotional	SUD Imagined	.29	29	0.68	10	18
	SUD During BAT	.14	14	0.50	6	30
	Fear of Specific Animal	.15	15	0.55	7	14
	Fear Questionnaire	.23	23	0.66	10	29
	Mean	.18	18	0.55	7	22

Note: EFT = Emotional Freedom Techniques; SUD = subjective units of distress; BAT = Behavioral Approach Task.

to what extent effects dissipated (follow-up vs. posttest differences); and (c) the size of this difference from pretest to follow-up for EFT compared with the other conditions using Cohen's *d* (see Table 3).

At the time of the original intervention, the outcomes for the supportive interview condition and the no-treatment control condition were absolutely indistinguishable. To gain statistical power in undertaking the follow-up analyses, we decided a priori to pool these two conditions in comparing their outcomes to that of the EFT condition. As for each analysis there was a clearly specified directional hypothesis involved in these follow-up analyses, one-tailed statistical tests were employed.

Results of the Follow-Up Study

Follow-Up for FQ

EFT participants scored relatively high on the original pretest ($M = 26.3$, $SD = 13.9$) and substantially lower on the posttest ($M = 13.2$, $SD = 8.00$). At the follow-up, a substantial part of this decrease had been lost ($M = 18.7$, $SD = 6.72$); for posttest versus follow-up, $t(9) = 5.10$, $p < .001$. However, the mean score at follow-up had not returned to the original value and was significantly lower than the original pretest value, $t(9) = 2.08$, $p < .04$. With respect to the ANOVA comparing EFT with the two (pooled) comparison conditions, there was a significant Treatment × Time interaction, $F(1, 22) = 6.4$, $p < .01$ (see Fig. 2a).

The effect size (η^2) for this interaction showed 22.5% explained variance. These last three findings, considered together, provide clear evidence

that the effect for EFT persisted for this variable at the time of follow-up.

Follow-Up for SUD Imagined

On this variable, EFT participants scored relatively high on the original pretest ($M = 8.67$, $SD = 1.00$) and substantially lower at the posttest ($M = 3.30$, $SD = 2.21$). At the follow-up, a substantial part of this decrease had been lost ($M = 7.30$, $SD = 2.16$); for posttest versus follow-up, $t(9) = 8.49$, $p < .001$. However, the scores had not returned to their original value and were significantly lower than the original pretest value, $t(8)^{ii} = 2.03$, $p < .04$. With respect to the ANOVA comparing EFT with the two (pooled) comparison conditions, there was a significant Treatment × Time interaction, $F(1, 21)^{ii} = 8.69$, $p < .004$ (see Fig. 2b). The effect size (η^2) for this interaction showed 29.3% explained variance. These last three findings, considered as a set, provide evidence that the effect for EFT persisted for this variable at the time of the follow-up.

Follow-Up for the FOSAQ

EFT participants scored relatively high on the original pretest ($M = 6.17$, $SD = 0.83$) and substantially lower on the posttest ($M = 2.97$, $SD = 1.74$). At the follow-up, a substantial part of this decrease had been lost ($M = 5.19$, $SD = 1.40$); for posttest versus follow-up, $t(8)^{ii} = 5.82$, $p < .001$. However, the mean scores at the time of the follow-up had not returned to the original value and were still lower than the pretest value, $t(9) = 1.73$, $p < .06$. With respect to the ANOVA comparing EFT with the two (pooled) comparison conditions, there was a significant Treatment × Time interaction, $F(1,$

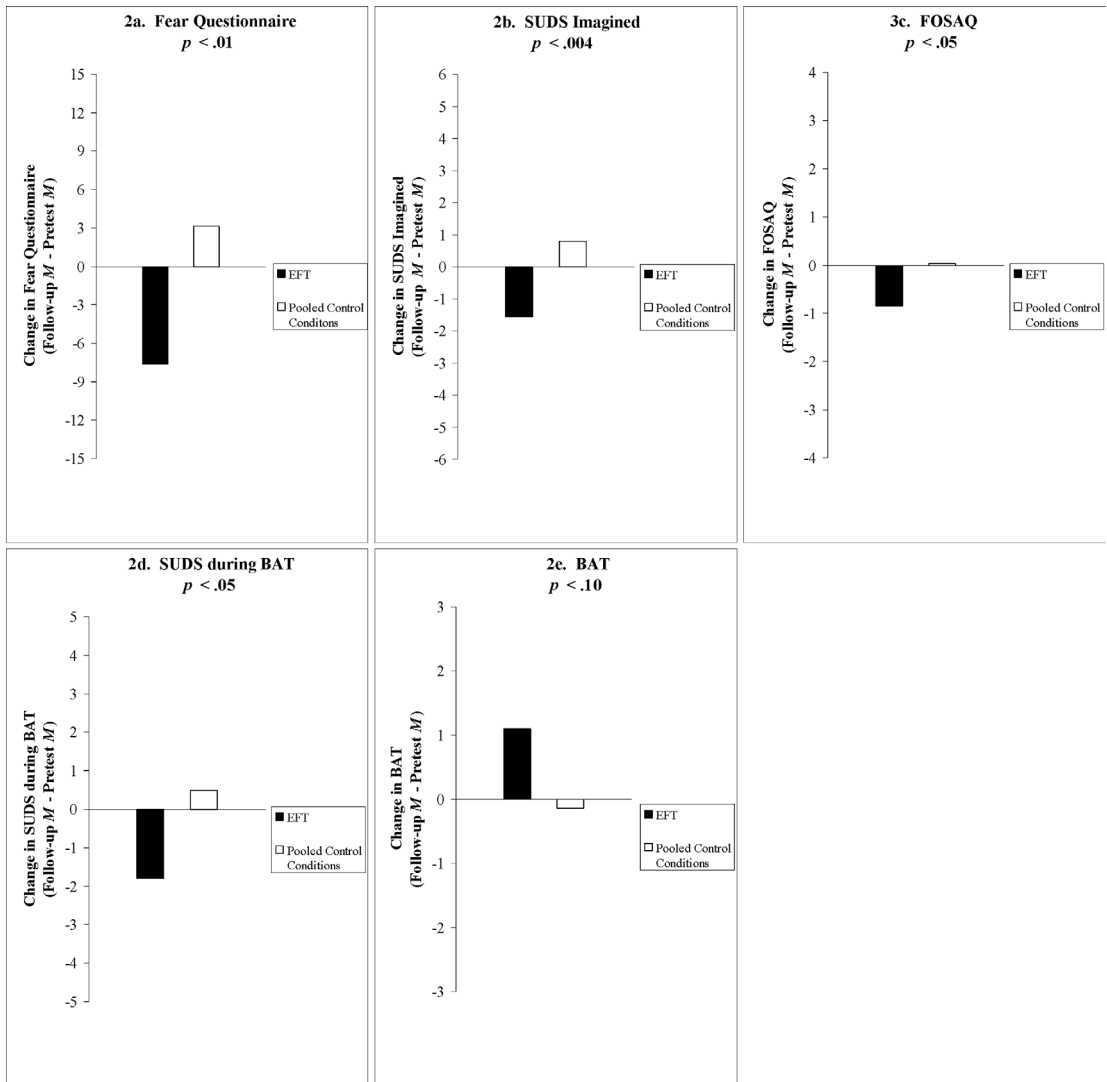


Figure 2. Follow-up effects of EFT compared with pooled control conditions (supportive interview and no treatment). EFT = Emotional Freedom Techniques; SUDS = subjective units of distress; FOSAQ = Fear of Specific Animal Questionnaire; BAT = Behavioral Approach Task.

21)ⁱⁱ = 3.55, $p < .05$ (see Fig. 2c). The effect size (η^2) for this interaction showed 14.5% explained variance. These last two findings, coupled with the marginal outcome for the pretest versus follow-up t test, provide a basis for concluding that there is evidence here that the effect for EFT persisted for this variable at the time of the follow-up.

Follow-Up for SUD During BAT

EFT participants scored relatively high ($M = 6.06$, $SD = 1.97$) at the time of the original pretest

and substantially lower at the posttest ($M = 2.51$, $SD = 2.21$). At the follow-up, a substantial part of this decrease had been lost ($M = 4.26$, $SD = 2.15$); for posttest versus follow-up, $t(9) = 2.51$, $p < .02$. However the scores at the time of the follow-up had not returned to their original value and were still lower than the original pretest value, $t(9) = 1.52$, $p < .09$. With respect to the ANOVA comparing EFT with the two (pooled) comparison conditions, there was a significant Treatment \times Time interaction, $F(1, 22) = 3.59$, $p < .05$ (see Fig. 2d). The effect size (η^2) for this interaction showed

14% explained variance. These last two findings, coupled with the marginal outcome for the pretest versus follow-up *t* test, provide a reasonable basis for concluding that there is evidence that the effect for EFT persisted for this variable at the time of the follow-up.

Follow-Up for the BAT

EFT participants scored relatively low on the original pretest (that is they did *not* closely approach the feared animal, $M = 5.30$, $SD = 3.50$) and showed substantial improvement ($M = 7.60$, $SD = 1.43$) at the posttest. At the follow-up ($M = 6.40$, $SD = 3.06$), there was some dissipation of this improvement, but it did not reach a level of statistical significance: For posttest versus follow-up, $t(9) = 1.23$, *ns*. Although the scores at the time of the follow-up had not returned to their original value, they no longer differed significantly from the values of the pretest, $t(9) = 1.19$, $p < .15$, *ns*. With respect to the ANOVA comparing EFT with the two (pooled) comparison conditions, the Treatment \times Time interaction approached but did not reach the conventional level of significance, $F(1, 22) = 2.10$, $p < .10$ (see Fig. 2e). The effect size (η^2) for this interaction showed 8.7% explained variance. Although the patterning of the data here is exactly parallel to the preceding four outcomes for the other dependent measures, this is at best suggestive, given the marginal significance.

A Comparison Across These Five Dependent Measures

The Treatment \times Time interactions were significant in four instances and approached ($p < .10$) the conventional level in the fifth. The mean amount of explained variance associated with these interactions was 18%. Taken together, this provides considerable evidence that the effects of EFT were persistent, as seen in Table 3.

Discussion

This study investigated EFT in the treatment of specific fear of certain small animals. More specifically, it explored whether these effects were due to common factors associated with various forms of psychotherapy, and it also investigated whether certain methodological artifacts were the source of the effects observed in an earlier study by Wells et al. (2003). On both the behavioral and emotional

levels of assessment, but not on the physiological level, there was evidence that a single session of EFT produced immediate fear reduction. The sizes of the observed effects were robust: (a) the mean amount of explained variance for the analyses involving all three treatment conditions (across all dependent measures, including those that failed to show significance) was 35% (see Table 2) and (b) the mean size of Cohen's *d*—now comparing only EFT's pretest and posttest means—was 1.44, a large effect. Although these initial effects of EFT did show considerable and occasionally significant shrinkage over time, they did not disappear. It is puzzling why pulse rate, both here and in the Wells et al. study, did not show greater improvement for EFT, although this result is consistent with Ost (1991) and Turpin (1989), who indicated that behavioral interventions yield changes on physiological measures less regularly than they do on behavioral and self-report measures.

Positive Contributions of This Study

Critics of EFT (e.g., Waite & Holder, 2003) raised the possibility that EFT's observed efficacy is simply due to common (i.e., nonspecific) factors long-known to be associated with various forms of psychotherapy (cf. Frank, 1971). We assessed this possibility by using a supportive interview as the comparison-placebo condition to provide a reasonable control for such common factors. As Behar and Borkovec (2003) concluded:

Although many types of nonspecific control groups that have been used in past research have suffered from . . . [various] . . . problems, one common factors condition is less severely handicapped by these difficulties . . . a nondirective or supportive listening condition. (p. 217)

The supportive interview condition produced no observable improvement, whereas EFT did; this offers no support for the hypothesis that the operation of factors common to most forms of psychotherapy produced the effects here observed for EFT.

The present findings constitute a partial replication of those of Wells et al. (2003), using as they did both multiple dependent measures and a duration of intervention comparable to that used in clinical practice. The interval between the original intervention and the follow-up—1.4 years—is

nearly double the interval of the Wells et al. study. The failure to find any hint of improvement for the no-treatment condition provides evidence that the effects observed for EFT are not due to such causes as regression to the mean, fatigue, the passage of time, or practice effects associated with repeated measurement. (The finding of no improvement for the supportive interview provides a second, independent basis for drawing the same conclusions.) This represents an advance over the more limited conclusions that could be drawn from the Wells et al. study.

The present study assessed participant's expectation of help immediately after the intervention condition to which she or he had been randomly assigned had been described. Expectation of help was significantly higher ($p < .003$) for the supportive interview compared with the no-treatment control condition; however, the outcomes for these two conditions were indistinguishable. There was no significant difference in expectations of help for EFT compared with the supportive interview, yet the outcomes for these two conditions differed significantly. Because the difference between the latter two conditions showed a probability of less than .15, ANCOVAs controlling for the effects of participants' expectations (something not undertaken in the Wells et al., 2003, study) were undertaken and still showed significant effects for EFT. Thus, the present study found no empirical evidence that participants' a priori expectations produced the observed efficacy of EFT. This is an advance over what could be concluded from the earlier study by Wells et al.

However, the present study's design could not rule out all possible contribution of expectations—for example, interaction *during* EFT with a therapist who had a positive attitude toward EFT may still have affected a participant's expectations.

Process(es) Which May Mediate the Observed Effects of EFT

The immediate effects of EFT in the treatment of fears (the present study) and specific phobias (Wells et al., 2003) indicate that EFT showed efficacy in reducing participants' fear of insects and small animals. Although this study was not designed to determine the mechanism of action of EFT, we here consider some processes that may mediate these effects. We later consider certain

other processes that might have artifactually produced these results (see **Limitations**).

Energy Meridian System. EFT is usually considered to be a Meridian-Based Intervention because it is based on the tapping of meridian end points used in traditional Chinese medicine. It may be that EFT's effects reflect stimulation of the body's so-called energy system as hypothesized by Craig (1995, 1999). Because the present study was *not* designed to assess the energy system hypothesis, it neither supports nor refutes it.

Imaginal exposure. The EFT procedure includes a frequent repetition of a Reminder Phrase that refers directly to the feared object as well as regular requests to imagine being in the feared situation during SUD assessments. Future research needs to determine the degree to which this imaginal exposure in and of itself contributes to EFT's observed efficacy. In the Wells et al. (2003) study, EFT and the comparison condition of systematic diaphragmatic breathing entailed a similar amount of imagining. Given that EFT and DB (both of which used imaginal exposure) in the Wells et al. study produced significant reductions in fear, whereas the supportive interview in the present study (which involved no use of imaginal exposure) did not, the hypothesis that imaginal exposure is a critical ingredient in the efficacy of EFT warrants further investigation. The fact that EFT showed significantly greater improvement than DB suggests that the physical intervention used with EFT was more effective than diaphragmatic breathing as a way of reducing the arousal (and hence increasing the relaxation) evoked by imaginal exposure.

Stimulating acupuncture points during imaginal exposure results in a novel form of desensitization. This formulation follows Wells et al. (2003, pp. 944–945): Some (as yet unidentified) aspect(s) of EFT may lead to exceptionally rapid relaxation. Both clinical patients and participants in EFT studies (the current study; Wells et al.) have spontaneously reported that they felt “very” and “quickly” relaxed after tapping on various body locations. Some evidence using fMRI suggests that stimulating certain acupuncture points sends messages to the amygdala and other brain centers that reduce arousal in fear centers (Hui et al., 2000). Other laboratory studies have suggested that stimulating other acupuncture points increases serotonin (Ruden, 2007). When such physiological interventions are coupled with

frequent imaginal focusing on the feared object—a focusing required during each round of the EFT procedure—we hypothesize that a novel form of desensitization, or counterconditioning, is induced, leading to fear reduction. This formulation has received empirical support from recent findings (Baker, Putlin, & Carrington, 2009) showing that relaxation increased significantly from pre-EFT to post-EFT and that the degree of increase in relaxation correlated significantly with degree of decrease in fear from pre-EFT to post-EFT.

Distraction. A question frequently directed at EFT asks whether the tapping process works by distracting the patient/participant from focusing on the fear. In a similar vein, Waite and Holder (2003) commented that “It is possible that systematic desensitization and distraction are mediators of EFT’s apparent effectiveness” (p. 24). It is not entirely clear how distraction might specifically operate in this context. A person undergoing EFT is repeatedly directed to focus on his or her fear (by saying out loud such a phrase as “fear of rats”). For each of the three intervention conditions involving tapping used by Waite and Holder (each of which produced decrease in fear), there were 25 such instances reminding the participant to focus on the fear; but there was only one such reminder before and during making a paper toy, their control condition (which produced no significant decrease in fear). In terms of any commonsense operational definition of distraction, EFT focuses attention directly onto the feared object rather than away from it, precisely the opposite of distraction.

Limitations of This Study, Possible Sources of Error, and Directions for Future Research

Generalizability of results. First, although EFT is offered as a form of treatment suitable for dealing with a wide array of psychological problems, we here have studied a very specific problem—high fear of certain small animals—the same type of issue addressed by Wells et al. (2003). Conclusions thus need to be limited to this type of problem. Second, although the participants were highly fearful and quite motivated to overcome this fear, only seven met and an additional nine came close to meeting the *DSM-IV* criteria for specific phobia (see **Participants**, above). It is thus not known whether these results would generalize to a strictly clinical population.

No specific mechanism of EFT’s action has been experimentally identified. There was no attempt here (or in Wells et al., 2003) to explore experimentally the specific mechanism of action that underlies EFT’s efficacy. Future research should undertake dismantling or other studies that directly addressed the issue of the specific mechanism of EFT’s action. Another possible approach would be to compare EFT with some other form of psychotherapy known to be efficacious (such as cognitive-behavior therapy or exposure therapy). If EFT produced equal and/or equally rapid results compared with one or more well-established forms of therapy in a well-controlled study, this would provide another basis for arguing that the efficacy of EFT is not simply due to common (nonspecific) processes associated with many forms of psychotherapy.

Shrinkage in EFT’s effects from initial session to follow-up session. From a clinical perspective, it is important that the percentage improvement for EFT shrank from 55% immediately following EFT to 22% at the time of the follow-up on the five measures reported in Table 3. The average patient would notice a 55% improvement, but it is doubtful whether the majority of patients would notice a 22% improvement. Thus, the immediate effects of EFT were quite large and clinically meaningful, but the effects at the time of the follow-up may not be clinically meaningful. This suggests that a single administration of EFT may well not be sufficient to produce a clinically meaningful long-term outcome for the average patient. We would recommend exploring whether the use of more (e.g., three) EFT sessions, coupled with homework assignments in-between, a procedure more in accord with the present clinical applications of EFT (Craig, 1999), might result in much less shrinkage. Further research can clarify this possibility.

The possible role of therapist allegiance. Luborsky et al. (1999) reported that the outcomes of psychotherapy research studies correlate highly with the allegiance of the researcher–therapist to the particular form of psychotherapy under study. This suggests that when a participant interacts with a therapist–researcher who has a positive (or negative) allegiance to the particular treatment under study, that participant may in some way be influenced to produce data supportive (or nonsupportive) of the efficacy of that particular form of therapy.

It seems both parsimonious and logical to expect that the effect of therapist allegiance, if any,

should manifest itself most clearly in terms of the efficacy of the condition to which allegiance is or is not directed, in this case, EFT. Wells et al. (2003) and our present study each found a significant decrease in fear from pre- to postintervention for the EFT conditions. Because each team of researchers had a positive allegiance to EFT, in neither study was it possible to rule out the possible role of such a therapist-allegiance effect. It would be of great interest to find evidence of EFT's efficacy in a study by someone with no prior allegiance to EFT. Now, in point of fact, Waite and Holder (2003) were "relatively neutral about whether EFT would work or not at the outset of the study" (M. Holder, personal communication, June 21, 2004). *When attention is restricted to their EFT condition*, Waite and Holder reported a decrease ($p = .003$) in fear level from baseline to posttreatment 1 (after a single round of treatment, lasting a few minutes only). In this restricted sense, their finding for their EFT condition is parallel to those of Wells et al. and ours, which also found a significant decrease in fear for the EFT conditions from pretest to posttest. Thus Waite and Holder's result for EFT only constitutes the first laboratory controlled study with evidence that EFT produces a significant effect in a study where it is quite unlikely that the possible role of therapist allegiance to EFT has contributed to the specific outcome for the EFT condition.

Future EFT studies should involve researchers who either (a) have no allegiance to EFT or (b) believe that EFT and the comparison treatment condition(s) are equally efficacious.

Demand characteristics. Some will argue that the present results reflect the operation of demand characteristics—that is, cues within the experimental procedures that directed participants to produce positive results. In medical research, the use of a design that is both double-blind and placebo controlled is regarded as the gold standard because it rules out various sources of bias. Double-blind studies are rare in the field of psychotherapy because it is not usually possible to have either the participant or the experimenter masked as to which treatment is being administered. The present study was single-blind: The data collectors were totally blind as to the particular intervention condition to which the participant had been assigned and the persons administering the interventions were totally blind to all data. One standard way of controlling for demand characteristics when a study

involves two well-known forms of therapy is to recruit therapists to conduct each form of intervention who have a positive allegiance to that form of intervention. This was not possible here. Although the present results are clearly positive and in keeping with the hypothesis, interpretation must thus be tentative.

Conclusions Regarding EFT

EFT shows immediate and long-term efficacy. These effects are replicable and not due to regression to the mean, practice effects, fatigue, or the passing of time. There is no evidence that they reflect differential a priori expectations of help on the part of participants. The present findings provide evidence that these effects of EFT are not produced by nonspecific common factors long known to be associated with most forms of psychotherapy. Further research is needed to explore the range of conditions for which EFT is efficacious, the treatment conditions required to maintain clinical gains, the relative power of EFT compared with other established therapies, and the mechanism(s) that produce the observed outcomes of EFT treatments.

Acknowledgments

We thank Phillip Ramsey for consulting with us regarding statistical issues; Timothy Rodier for assistance in executing the many statistical analyses; Jon P. Austin for help with the initial editing and with graphs and tables; Patricia Carrington, Nancy Hemmes, and Dimitri Putilin for their advice at many junctures in this research; Adela Oliver for help at various stages in this research; and David Feinstein for his excellent and searching critique of a draft of this article.

This study was supported in part by grants from the PSC-CUNY Research Awards Program for the years 2002–2003, 2003–2004, and 2004–2005 and by a grant from the Association for Comprehensive Energy Psychology.

We feel it important to note that all data were collected for the original intervention study (but not for the follow-up phase of the present research) before one of us (A. Harvey Baker) had any involvement whatsoever in statistical analyses and write up of the Wells et al. (2003) study. In that sense, the present intervention study (but not the follow-up) represents a replication by a then-independent laboratory.

References

- American Psychiatric Association. (1994). *Diagnosis and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Baker, A. H., Carrington, P., & Putilin, D. (2009). Theoretical and methodological problems in research on Emotional Freedom Techniques (EFT) and other meridian based therapies. *Psychology Journal*, 6(2), 34-46.
- Baker, A. H., Quiatchon, M. H., & Putilin, D. (2007). *Fear of Specific Animal Questionnaire: Reliability and validity*. Manuscript submitted for publication.
- Barrios, B. A. (1978). Note on demand characteristics in analogue research on small animal phobias. *Psychological Reports*, 42, 1264-1266.
- Behar, E. S., & Borkovec, T. D. (2003). Psychotherapy outcome research. In *Handbook of Psychology* (Vol. 2, 213-240). *Research Methods in Psychology*. New York: Wiley, 2003
- Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue and therapy rationales. *Journal of Behavior Therapy and Experimental Psychiatry*, 3, 257-260. doi:10.1016/0005-7916(72)90045-6
- Callahan, R. J., & Callahan, J. (1997). *Thought Field Therapy and trauma: Treatment and theory*. Indian Wells, CA: Author
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Craig, G. (1995). *Emotional Freedom Techniques: The manual*. Sea Ranch, CA: Author.
- Craig, G. (1999). *Emotional Freedom Techniques: The manual*. (Revised ed.). Sea Ranch, CA: Author.
- Ernst, E., & Resch, K. L. (1995). Concept of true and perceived placebo effects. *British Journal of Medicine*, 311, 551-553.
- Feinstein, D. (2008). Energy psychology: A review of the preliminary evidence. *Psychotherapy: Theory, Research, Practice, Training*, 45(2), 199-213.
- Frank, J. D. (1971). Therapeutic factors in psychotherapy. *American Journal of Psychotherapy*, 25, 350-361.
- Gordon, T. (1974). *Parent effectiveness training*. New York, NY: Van Rees.
- Hui, K. K. S., Liu, J., Makris, N., Gollub, R. W., Chen, A. J. W., Moore, C. I., . . . Kwong, K. K. (2000). Acupuncture modulates the limbic system and subcortical gray structure of the human brain: Evidence from fMRI studies in normal subjects. *Human Brain Mapping*, 9, 13-25. doi:10.1002/(SICI)1097-0193(2000)9:1<13::AID-HBM2>3.0.CO;2-F
- Kern, J. M., & MacDonald, M. L. (1980). Assessing assertion: An investigation of construct validity and reliability. *Journal of Consulting and Clinical Psychology*, 48, 532-534. doi:10.1037/0022-006X.48.4.532
- Kirsch, I. (1985). Response expectancy as a determinant of experience and behavior. *American Psychologist*, 40, 1189-1202. doi:10.1037/0003-066X.40.11.1189
- Marks, I. M., & Matthews, A. M. (1979). Brief standard self-rating for phobic patients. *Research and Therapy*, 17, 263-267. doi:10.1016/0005-7967(79)90041-X
- Ost, L.-G. (1987). Age of onset in different phobias. *Journal of Abnormal Psychology*, 96, 223-229. doi:10.1037/0021-843X.96.3.223
- Ost, L.-G. (1991). One-session therapist-directed exposure vs. self-exposure in the treatment of spider phobia. *Behavior Therapy*, 22, 407-422. doi:10.1016/S0005-7894(05)80374-0
- Ost, L.-G. (1996). One-session group treatment of spider phobia. *Behavior Research and Therapy*, 34, 707-715. doi:10.1016/0005-7967(96)00022-8
- Pignotti, M. (2005a). Callahan fails to meet the burden of proof for Thought Field Therapy claims: Rejoinder to Callahan. *Journal of Clinical Psychology*, 61, 251-255. doi:10.1002/jclp.20053
- Pignotti, M. (2005b). Thought Field Therapy voice technology vs. random meridian point sequences: A single-blind controlled experiment. *Scientific Review of Mental Health Practice*, 4, 38-47.
- Pignotti, M. (2007). Thought Field Therapy: A former insider's experience. *Research on Social Work Practice*, 17, 392-407. doi:10.1177/1049731506292530
- Ruden, R. A. (2007). A model for disrupting an encoded traumatic memory. *Traumatology*, 13, 71-75. doi:10.1177/1534765607299909
- Teusch, L., Bohme, H., & Gastpar, M. (1997). The benefit of an insight-oriented and experiential approach on panic and agoraphobia symptoms. *Psychotherapy and Psychosomatics*, 66, p. 293-301. doi: 10.1159/000289151
- Thyer, B. A., Papsdorf, J. D., Davis, R., & Vallecorsa, S. (1984). Autonomic correlates of the Subjective Anxiety Scale. *Journal of Behavior Therapy and Experimental Psychiatry*, 15, 3-7. doi:10.1016/0005-7916(84)90115-0
- Turpin, G. (1989). *Handbook of clinical psychophysiology*. New York, NY: Wiley.
- Waite, L. W., & Holder, M. D. (2003). Assessment of the emotional freedom technique: An alternative treatment for fear. *Scientific Review of Mental Health Practice*, 2, 20-26.
- Weinberger, J., & Eig, A. (1999). Expectancies: The ignored common factor in psychotherapy. In I. Kirsch (Ed.), *How expectancies shape experience* (pp. 357-382). Washington, DC: American Psychological Association. doi:10.1037/10332-015
- Wells, S., Polglase, K., Andrews, H., Carrington, P., & Baker, A. H. (2003). Evaluation of a meridian-based intervention, Emotional Freedom Techniques (EFT), for reducing specific phobias of small animals. *Journal of Clinical Psychology*, 59, 943-966. doi:10.1037/10332-015
- Wolpe, J. (1958). *Psychotherapy by reciprocal inhibition*. Stanford, CA: Stanford University Press.

Notes

¹ Because the data for some participants were missing, the degrees of freedom varied from 25 to 28 here.

² The data for one participant was missing here.