Published in final edited form as:

Behav Ther. 2006 June ; 37(2): 120–130.

Designing a Control for a Behavioral Group Therapy

Debra L. Safer [Assistant Professor]

Department of Psychiatry and Behavioral Sciences, Stanford University, dlsafer@stanford.edu

Emily M. Hugo [Doctoral Candidate]

Pacific Graduate School of Psychology-Stanford Psy.D. Consortium

Abstract

To evaluate whether the specific techniques of a newly developed behavioral therapy exceed the effects of its common factors (e.g. therapeutic alliance), an ideal comparison control must omit the unique ingredients of the active treatment while possessing the common factors in equal measure. Reviews of the controlled-trial psychotherapy research literature show that such systematic matching is not a common practice. Addressing this gap between theory and practice, we delineate seven steps for designing a credible control for a recently developed behavioral therapy adapted for Binge Eating Disorder—Dialectical Behavior Therapy (DBT). Our detailed description of matching the common factors to the extent possible and specifically measuring these offers a useful research design template for investigators of controlled behavioral treatment trials.

Designing a Control for a Behavioral Group Therapy

The specific therapeutic effect of a group psychotherapy treatment is most rigorously evaluated when the treatment under investigation is as closely matched as possible to a randomly assigned comparison treatment group (see for example Horvath, 1988; Lohr, DeMaoi, & McGlynn, 2003; Stevens, Hynan, & Allen, 2000). To control for the effects of the common factors (e.g. therapeutic alliance, therapeutic rationale, etc.) on outcome, a comparison group must possess these general therapeutic factors while omitting the unique active ingredients of the treatment under investigation (Horvath, 1988; Stevens et al., 2000). As Stevens et al. (2000) assert: “The task of demonstrating that specific techniques in complete treatments exceed common factors requires outcome studies that compare complete treatments with common factor controls.”

Much planning and effort is required for this task. Reviews of the controlled-trial psychotherapy research literature show that systematic matching is not a customary practice (Baskin, Tierney, Minami, & Wampold, 2003; Critelli & Neumann, 1984; Horvath, 1988; Stevens et al., 2000). Rather than precisely controlling for the common factors across psychotherapies, researchers typically use waiting lists, educationally focused groups, dissimilar comparative treatments, and treatment as usual for this purpose. Such inequivalences in design have been shown to impact outcome. For example, defining structural equivalence by including the same number of sessions and treatment modality (e.g. group therapy, pharmacotherapy), Baskin et al. (2003) found smaller effect size differences when active treatments were compared to structurally *equivalent* controls and larger effect size differences when controls were structurally *inequivalent*. Baskin and colleagues’ focus on requiring equivalent modalities in their definition may explain why earlier studies (Bowers & Clum, 1988; Stevens et al., 2000) were unable to detect meaningful group differences.

In initiating a study to evaluate the effectiveness of a group treatment using Dialectic Behavior Therapy (DBT) for binge eating disorder (BED), a key early task is to formulate and

systematically design a credible comparison therapy¹ that maintains internal validity without compromising external validity. By providing a detailed description of the step-by-step clinical choice-points faced by our treatment team to meet this challenge, we hope to frame the theoretical and practical issues facing the research field in demonstrating specific treatment efficacy beyond what would be caused by the common factors of psychotherapy. After a description of the literature on BED treatment outcome, the planning and resultant development of this control group for the DBT treatment will be described.

Background

Despite the existence of prominent theoretical and procedural differences among established BED treatments, there is a striking lack of evidence for their differential effectiveness (Wilfley, Wilson, & Agras, 2003). For example, the two most studied treatments for this disorder, Cognitive Behavioral Therapy (CBT) and Interpersonal Therapy (IPT) for BED, have very different conceptual bases yet have shown equivalent immediate and longer term binge abstinence rates in two head-to-head comparison trials (Wilfley et al., 1993; Wilfley et al., 2002).

Although CBT and IPT are efficacious, a substantial number of patients remain symptomatic at post-treatment or follow-up (Wilfley et al., 1993; Wilfley et al., 2002). Consequently, new treatments such as DBT for BED were developed. Drawing on an extensive literature linking negative affect and binge eating (Abraham & Beumont, 1982; Polivy & Herman, 1993), DBT for BED conceptualizes binge eating as a behavioral attempt to alter painful emotional states. Originally developed by Linehan (Linehan, 1993) to treat emotional dysregulation in borderline personality disorder, DBT was adapted by Telch and colleagues to target binge eating and has shown initial promise both in an uncontrolled and controlled comparison trial against a wait-list (Telch, Agras, & Linehan, 2000; Telch, Agras, & Linehan, 2001). Because the lack of differential outcomes among existing BED treatments may be accounted for by the effects of common factors, the current study was designed to isolate the specific active ingredients of DBT for BED from factors shared across psychotherapies.

The seven steps we used to design a credible comparison treatment for DBT for BED are: (1) Identify the common factors of psychotherapy (2) Identify the hypothesized specific elements of the experimental treatment (3) Identify the specific elements of currently existing treatment approaches for the disorder to avoid the use of known active techniques in the comparison therapy (4) Develop a credible comparison therapy (5) Match the two treatments regarding the common factors (6) Ensure that the two groups differ in terms of the specific active elements of the therapy being tested and (7) Measure the common factors as specifically as possible.

These seven steps are discussed in detail below.

Procedure

Step 1: Identify the common factors of psychotherapy—Our first step was to review the literature and identify the common factors of psychotherapy so as to allow systematic matching across our comparison and experimental treatments. Because definitions of the common factors vary, we turned to three thorough reviews (Altshuler, 1989; Grenavage & Norcross, 1990; Weinberger, 1993) as well as the influential model proposed by Frank & Frank (1991). The common factors distilled from these sources are defined in Table 1 and are referred to throughout this article.

¹The terms common factors, nonspecific effects, and placebo effects are often used interchangeably yet are not synonymous. See, for example, Weinberger (1993), Stevens et al. (2000), Lohr et al. (2003) for a thorough discussion of the various definitions of these terms. In this article we have chosen to use the term “credible comparison therapy” wherever possible to refer to a treatment attempting to control for common factors that are not specific to one school of psychotherapy.

Step 2: Identify the hypothesized specific elements of the experimental treatment (e.g. DBT)—Integral to DBT’s focus on the affect regulation model is the systematic linkage between dysregulated emotional states and binge eating. In this model, a repertoire of specific adaptive emotion regulation skills (Mindfulness, Emotion Regulation, Distress Tolerance) are taught in a didactic format to replace dysfunctional behaviors. One such “skill” involves a behavioral rehearsal of a verbal commitment to binge abstinence made during Session 1. Other specific elements of DBT for BED include weekly behavioral homework assignments such as filling out a diary card (to track the relationship between binge eating, emotional states and skills practice) and a behavioral analysis of binge episodes.

A comparison therapy then must avoid these hypothesized specific elements of DBT.² In other words, there should be no content overlap between the comparison treatment and the presentation of the affect regulation model nor any systematic linkage between emotional dysregulation and binge eating. Additionally, there should be no didactic instruction of DBT “skills” or behavioral homework reinforcing any of the specific elements outlined above.

Step 3: Identify the specific elements of currently existing treatment approaches for the disorder (e.g. BED) to avoid the use of known active techniques in the comparison therapy—A credible control treatment must avoid theoretical and procedural elements of active treatments. In addition to CBT and IPT, currently existing treatment approaches to BED include behavioral weight loss treatment (BWLTL) (See for example Agras et al., 1994; Marcus, Wing, & Fairburn, 1995), pharmacological treatment (See review by Carter et al., 2003), and psychodynamically-based treatment (Ciano, Rocco, Angarano, Biasin, & Balestrieri, 2002). To avoid content and procedural overlap with CBT (Fairburn, 1995; Marcus, 1997), our comparison therapy must not include: a) a focus on cognitions that identify irrational negative thoughts associated with binge eating episodes; b) a focus on the link between restrained eating and binge eating; c) a guide for patients to make specific changes to normalize their eating habits (e.g. eating 3 meals and 2-3 snacks per day); and d) the prescription of behavioral (e.g. food logs) or cognitive (e.g. thought records) assignments.

Similarly, to avoid content overlap with IPT (Klerman & Weissman, 1993; Wilfley et al., 1993; Wilfley et al., 2002), our comparison therapy must avoid systematic linkages between recent life events or relationships (e.g. interpersonal disputes, grief, interpersonal deficits, role transitions) and binge eating as well as the active systematic pursuit of specific interpersonal change strategies, including role playing to rehearse suggested changes in interpersonal behavior. To avoid content overlap with BWLTL (Agras et al., 1994; Marcus et al., 1995), our comparison therapy must not focus on a behavioral approach to weight loss and must avoid provision of specific weight loss strategies, goals, etc. This should not be misunderstood to preclude participants from discussing attempts to stop binge eating and to express hope that eventual weight loss might ensue with binge abstinence. Yet because DBT, like most psychotherapeutic treatments for BED, focuses on the goal of cessation of binge eating rather than weight-loss, a credible comparison must similarly avoid instruction in weight loss strategies.

There have been few randomized trials evaluating psychodynamic group treatments for BED. One of the available randomized studies, by Ciano and colleagues (2002), is not only quite small (five participants per condition) but it compared 28 weeks of psychoanalytic group therapy to 10 weeks of psychoeducational group treatment—an arguably unequally matched comparison condition. However, because psychodynamically-based treatments focus directly on the expression and “working-through” of emotional content, a comparison treatment based

²We are mindful of critiques, such as by Weisz, Hawley, Pilkonis, Woody, and Follette (2000) pointing out that the current research literature is unable to definitively identify the mechanisms by which even the best-researched treatments produce their effects.

on these tenets would lack internal validity in terms of content overlap with DBT's emotion regulation focus. Therapists in our comparison condition, therefore, must refrain from offering transference interpretations or other direct psychodynamically informed interpretations (e.g. interpretation of dreams). However, therapists are not prohibited from facilitating patient initiated discussions of their feelings about binge eating and its impact on their lives. Indeed, as identified in Step 1, a credible comparison treatment must include the common factor of emotional expression and the ventilation of affect. However, therapists must refrain from organizing the material in ways that would overlap with other emotion focused treatments such as DBT or psychodynamic treatments.

Finally, while concrete overlap with pharmacological treatments is obviously avoided by the fact that no pharmacological agents were dispensed with either treatment, our comparison therapy must avoid a focused presentation of a biological or medical model of eating disorders, such as a "chemical imbalance" (e.g. low serotonin) model.

Step 4: Develop a credible comparison therapy—Our goal was to develop a comparison therapy whose rationale and procedures would be credible enough to generate the common factors in full measure while lacking the specific elements of DBT for BED (identified in Step 2) and the specific elements of other BED approaches (identified in Step 3) (Horvath, 1988, p.217). A constant tension was meeting the above goals and producing a viable psychotherapeutic intervention (maintaining external validity without developing an entirely new specialty treatment for BED).

Our first consideration involved using a supportive-expressive control. Careful consideration of existing manuals for supportive expressive treatments, such as those used in treating breast cancer (Classen et al., 1993) revealed the unacceptability of such an approach due to its overt emotion evoking emphasis. We then considered basing our comparison treatment on the "nonprescriptive" treatment manual that served as a control against CBT group therapy for obese women with Type II diabetes (Kenardy, Mensch, Bowen, Green, & Walton, 2002). As part of that treatment manual, therapists are instructed to inform clients that "triggers of binge eating episodes are often related to a negative feeling which has not been directly acknowledged" —a form of the affect regulation model. The manual trains therapists to use an emotion-focused technique – a five step procedure entitled "focused unfolding."

The Kenardy et al. choice of a control usefully illustrates the real-world challenges facing researchers designing control treatments. We saw for ourselves that no specific control treatment could serve as an adequate comparison for all experimental treatments but must be designed carefully in relation to the experimental treatment under investigation (see, for example, Lohr et al., 2003). Kenardy et al.'s nonprescriptive treatment was specifically designed to serve as a comparison treatment to CBT. Its designation as a "nonprescriptive" therapy must be accepted with caution, however. While a focus on increasing awareness and acceptance of negative emotions may not be considered a prescriptive or active technique according to CBT theory, it clearly has the potential of being seen as active according to an alternative theoretical conceptualization. As Stevens et al. note: "One researcher's specific treatment may well be another's nonspecific, common factor" (Stevens, 2000 p.284). The fact that CBT showed no significant differences from the control at post-treatment (though did show significantly lower relapse rates at 3 month follow-up) (Kenardy et al., 2002) may be due to the lack of a specific therapeutic effect of CBT or to the incorporation of an active treatment rationale into the control condition.

In our search to find a control of sufficient power to credibly represent the common factors while avoiding content overlap with the affect regulation model of DBT or any known active treatment, we then turned to a manual written by John Markowitz and Michael Sacks (2002).

Originally intended to serve as a supportive psychotherapy control condition for a chronic depression study, the manual instructs therapists to follow a Rogerian approach (Rogers, 1951). Therapists are to encourage patients to find answers within themselves instead of offering patients specific techniques or skills. Self-esteem and self-efficacy are bolstered by highlighting patients' strengths over exploring their weaknesses. We believed the Markowitz & Sacks manual's therapeutic rationale could be credibly adapted for binge eating (i.e. bolstering self-esteem to enhance the ability to stop binge eating) while avoiding DBT's specific elements. In addition, our comparison therapy's specific ingredients (i.e. bolstering self-esteem) would seem to be indistinguishable from those evoked by the common factors of the therapeutic alliance and development of therapeutic optimism.

The final development of our comparison treatment included a pilot phase to allow for "road-testing" and potential revision. Unlike DBT's treatment manual, which we knew to be acceptable to therapists and patients from prior research, our newly designed comparison treatment manual had never been tested.

This pilot phase utilized two co-therapists (serving as the therapists for both the DBT and comparison conditions) and 14 participants (7 per condition) recruited from advertisements. Of particular interest was treatment acceptability. What did patients spontaneously express about the treatment? How did the therapists describe their own experiences? Except for an initial rating of patient expectations at the beginning of treatment, such "data" were not gathered via quantitative measures but were culled from therapist supervision sessions and careful review of audiotapes and transcripts of the sessions. We wanted to make sure, in our efforts to avoid designing an entirely new treatment with active elements that potentially exceeded the common factors, that our treatment's therapeutic rationale was sufficiently believable. We took seriously concerns expressed by researchers such as Westen, Novotny, and Thompson-Brenner (2004) about researchers' use of "intent-to-fail" conditions (ie control treatments not designed to maximize their efficacy) and findings, such as those reported by Luborsky and colleagues (Luborsky et al., 1999), that the therapy allegiance of a study's researcher is strongly associated with the outcome of a comparative study.

At the initiation of our pilot phase, patients were encouraged to focus on their strengths in other areas of their lives as a way of bolstering their low self-esteem, thereby enhancing their ability to stop binge eating. In ways that we had not anticipated, the presentation of our treatment model was missing an essential ingredient. This became clear in Session 5, in which a noticeable shift in optimism about the credibility of the treatment rationale was observed.

Therapist: [in response to a patient's description of a non-binge related success from the past] It really shows you're capable of doing some very difficult things. So I think the question is, and we talked about this last time too I think, how to apply the things you know in certain areas ...

Patient: I know there're other things I've conquered, there're other things I've done that I can give myself credit for. I realize all that, but when I think about not eating when I want all those things, it doesn't help me. I can't apply it to that. **I'm really reaching the point right now where I'm getting really discouraged because coming here was like "Ok, maybe this will help," and we're just far enough along where I'm going "But it isn't!"** I don't think I care enough anymore

Second patient: **Yeah, it isn't helping**

Third patient: **Yeah I know**

Fourth patient: **It should be helping me**

Fifth patient: **Exactly**

Patient: Then I get more frustrated and eat. I think I should be getting something here so that I don't keep doing this and I guess I haven't found it yet, whatever it is. **And I don't know how to build from other successes and apply it to this because I've sort of even tried that.** Because like years ago I drank too much and then I decided "No, that was not a good thing" and I was able to stop. I used to smoke, I don't smoke anymore. **I did all those things, but it almost doesn't seem to have any relationship to what I'm trying to overcome now.**

The demoralization the patients expressed in the pilot study, which was echoed by the therapists in the subsequent weekly team meeting, concerned us, especially as there was no counterpart demoralization being voiced in the DBT group. It could be argued that patients' optimism about treatment must only be equivalent at the beginning of treatment and that one's comparison treatment, in lacking specific active elements, is required to be relatively inefficacious. From this standpoint, an intervention to increase the level of any of the common factors, such as therapeutic optimism, would potentially be "loading the dice" in favor of our comparison treatment.

We would argue that such a strong expression of demoralization so early in treatment (Session 5 of 20) seriously threatens our requirement that our comparison treatment be a credible comparison to DBT. Adequate credibility, in our opinion, would mean that a comparison treatment would be believable enough to allow patients to express faith in it not only at the beginning of treatment but also throughout. While this would not require that our comparison therapy's ability to generate therapeutic optimism must equal that of DBT at session 5 (or session 20, for that matter), adequate believability should exist such that the majority of patients would not drop out of our "credible" comparison treatment.

Hence, we felt it necessary to reexamine our comparison therapy's treatment rationale carefully. During the initial screening process, many patients spontaneously identified themselves as having low self-esteem and spontaneously referred to this playing a major role in their binge eating. This appeared to indicate that the rationale itself was sufficiently plausible. The problem seemed not to lie in the rationale itself but in a delivery that lacked an ingredient we had not anticipated was necessary: the presence of sufficient focus, both during the initial delivery and during the following discussions referring to the rationale. In other words, it seemed as if without maintaining focus to adequately generate the common factor of a therapeutic rationale, the treatment was not able to subsequently generate the common factors of optimism and positive expectancy for change. Something specific needed to take place in our comparison therapy in order to generate a sense of hope and consequent movement. Our therapists needed to provide this focus without using a DBT or other specific known therapeutic rationale.

We were reminded that the DBT for BED treatment begins with a flow chart explicitly detailing the relationship between emotional dysregulation and binge eating. A similarly explicit flow-chart of the comparison treatment's therapeutic rationale was thus developed to provide missing structure for those patients who especially seemed to note its absence. Also, to heighten the comparability to our experimental treatment and to provide additional sources of therapeutic focus, patients were to update the group on their successes and failures in stopping binge eating during the prior week and to bring up difficulties about which they would like the group's support. Therapists were instructed to facilitate any patient-initiated discussions of binge eating and related topics (e.g. therapeutic goals) and to highlight the relationship between patients' self-esteem and binge eating when appropriate.

We found patients responded quite well to these changes and the demoralization did not recur. Interestingly, therapist attention to providing the discussion with focus was less and less

necessary over time with patients bringing up all manner of topics. Upon provision of an adequate focus, the sense of structure appeared self-generating.

Step 5: Match the two treatments so that both contain the same common factors (see Table 2)

Choice point: Therapeutic Alliance: Beneficial therapist qualities: The decision as to whether to keep the therapists constant in a treatment study comparing two (or more) treatments is one with no ideal answer. In making the decision to keep the therapist and co-therapist constant between groups, we were able to minimize variability due to therapist differences. Yet the decision to keep the same therapists in both treatments necessarily opened us to the potential criticism that the therapists might be biased towards one treatment arm over the other. This limitation is referred to in our discussion section. One practical constraint common to smaller scale studies such as ours is that it is often not possible to recruit and pay a large enough pool of therapists to allow for a random assignment between groups.

Choice point: Rituals to be observed/a healing setting: DBT utilizes the strategy of monitoring behavior and patients fill out a diary card each week. To match the specific ritual of filling out a diary card as well as to account for the known activity of monitoring alone (Kazdin, 1974), we included a diary card in our comparison therapy. In this diary, patients were asked to monitor their binge eating and self-esteem.

Step 6: Ensure that the two groups differ in terms of the specific active elements of the therapy being tested (See Table 3)

Choice point: Rituals to be observed/a healing setting: The active ingredients of the DBT treatment were avoided in our comparison treatment through making sure that the comparison treatment diary cards did not track the linkage between emotions and binge eating. Unlike in the DBT group, where review of the diary card is a mandatory part of the treatment, patients were free to use and refer to their diary cards during the group.

Choice point: The opportunity to express emotions or ventilate problems.: Since the facilitation of affect is one of the common factors in psychotherapy, we wanted our comparison group to generate this factor as fully as possible while avoiding specific affect regulation elements of the DBT model (see Step 2). Therapists were instructed to facilitate a supportive atmosphere in which patients' feelings about their binge eating and related aspects of their lives were expressed openly and validated. Unlike DBT therapists, however, who introduce and explore linkages between emotions and behavior, therapists in the comparison treatment were instructed not to introduce such linkages and to use reflection and restatement when patients did so.

Choice point: Acquisition and practice of new behaviors: Because DBT is a behavioral therapy with specific homework assignments, we wanted to make sure our comparison therapy matched DBT in the acquisition and practice of new behaviors but without provision of the specific active elements of DBT. Therapists were encouraged to provide a safe and supportive environment within which patients would be encouraged to take chances and experiment with new choices. Through the focus on self-esteem, therapists reassured patients that, for example, some failures along the way are common.

Step 7: Measure common factors as specifically as possible—Having identified the common factors that appeared to be most essential, the next step involved an attempt to measure them as specifically as possible to determine whether the common factors were generated in each treatment as well as to what degree. This was accomplished by having the participants and therapists complete rating scales (Table 4). Gathering such information in no

way implies that potential differences between the groups' common factors would be viewed as nuisance variables to be statistically controlled. Indeed, differences between the two groups' therapeutic alliances may be important in their own right. It was because our experimental and comparison treatment may differ in their ability to generate the common factors that we aimed to measure, to the extent possible, such differences, before drawing any conclusions. For example, it is possible that higher levels of the therapeutic alliance in one group may be in part a consequence of a more effective treatment. By measuring putative mediators of treatment outcome during the course of treatment, such as the therapeutic alliance, we would hope to capture whether decreases in binge eating precede higher levels of the therapeutic alliance or vice versa.

Preliminary Findings

Though our study is currently still in-progress and includes a 1 year follow-up, we present data on some of the common factors for those patients (n=60) treated to date. Treatment consisted of 20 sessions of group therapy, with patients randomly assigned to DBT for BED or our comparison treatment. Treatment was carried out in cycles, each including 18–22 patients (with 9–11 patients assigned to each treatment group). To date, the same PhD level psychologist serves as the lead co-therapist for each condition. Two doctoral level co-therapists have participated, with one co-leading during cycle 1 and the other during cycles 2 and 3. Hence during each cycle, the same therapists served as the therapists for both conditions. No differences between the groups on any of the baseline measures analyzed thus far (e.g. age, body mass index (BMI), binge frequency) were detected.

Using paired student t-tests, we compared patients on a number of the common factors (Refer to Table 5). Except for differences in session 12 therapist optimism about patients, where greater mean ratings in DBT were significant compared to the comparison treatment ($p = 0.015$), no other comparisons met statistical significance.

Discussion

Methodological issues concerning the concept of a common factors control group have been raised in a number of thoughtful critiques (Baskin et al., 2003; Chambless & Hollon, 1998; Critelli & Neumann, 1984; Horvath, 1988; Lohr et al., 2003; Parloff, 1986; Stevens et al., 2000). For example, while there appears to be agreement that common factors exist across psychotherapies (Weinberger, 1993), there is a lack of agreement over what they are, what they should be termed (e.g. “nonspecific” factors, “common factors,” “placebo effects”), how they should be operationally defined (Horvath, 1988), which factors are most important, how they are interrelated, and whether they can be reliably measured (Messer & Warren, 1995).

In actual use, control conditions vary considerably across treatments. Indeed, there is little empirical evidence for the general assumption that by definition common factor controls are composed of factors “common” across therapies (Baskin et al., 2003; Horvath, 1988). As Horvath cautions: “Assuming variables to be common or nonspecific does not make them so” (Horvath, 1988, p. 221). Furthermore, a large proportion of therapies and control conditions do not share common elements (Horvath, 1988).

Despite the above concerns, the concept of a common factor control remains in active use (See for example Baskin et al., 2003; Lohr et al., 2003). In order to provide an empirical basis for choosing an “active” treatment, an important question to address is whether a specific treatment's unique ingredients exceed the effects of its common factors. The BED research literature's lack of differential treatment effectiveness provides a case in point of the importance of designing a credible common factors control treatment. Before engaging in the expense of a sufficiently powered head to head comparison against another active treatment

(e.g. CBT or IPT), researchers of a newly developed treatment with demonstrated efficacy against wait-list such as DBT for BED should first compare that treatment against an “active placebo.” Indeed, as noted by Critelli and Neuman, “The placebo control group strategy provides the most direct and unambiguous answer to the question of whether a treatment shows a level of effectiveness beyond that of its placebo effects (Critelli & Neumann, 1984).

The aim of this paper has been to outline our steps for designing a credible comparison treatment for DBT for BED. We also present preliminary data (see Table 5) that indicate that our concerted efforts resulted in a treatment in which 13 of the 14 measures of the common factors did not statistically differ from our experimental treatment.

Several limitations of our attempt to design an implement a credible comparison therapy must be acknowledged. In keeping with longstanding criticisms of the entire enterprise of comparing experimental treatments with “placebo” treatments (See for example Parloff, 1986), our study was not double blinded (though our assessors were). In other words, our therapists knew which was the DBT and which the self-esteem group. Close supervision and ongoing adherence ratings allowed us to feel assured that we could manage this risk.³ Furthermore, much attention was placed on addressing the groups on equal terms (e.g. “Tuesday” versus “Thursday” group).

The statistically greater therapist ratings of therapeutic optimism about treatment outcome may be an example of another limitation of the study, namely that the active component of the experimental treatment may interact in unique ways with the common factors to produce additional differences which we would not know how to control (Horvath, 1988). An illustration of this difficulty is that even when we as researchers are aware of potential specific interactions between our experimental treatment and the common factors, the decision as how best to control for these is not straightforward. For example, DBT—with its focus on teaching adaptive emotion regulation skills—has a didactic format. It is possible that this specific ingredient of DBT had a particular effect on the common factor of therapeutic optimism that was not present in our comparison treatment. Perhaps including a didactic component of self-esteem enhancing “skills” would have made our comparison treatment a better control for DBT. We present this dilemma as one that we are unable to answer at present. For instance, just as it is likely that the didactic focus strengthened the common factor of therapeutic optimism, it may have weakened the common factor of the alliance and support between the participants and their fellow-group members. Furthermore, it is not clear we would have been able to successfully create a set of self-esteem skills that had no content overlap with existing treatments, as many such techniques involve a focus on altering cognitive distortions or encouraging behavioral change (e.g. mastery building).

Despite the limitations of our design, we believe that our detailed description of designing a credible comparison control and use of specific assessment measures of the common factors make a valuable contribution. The psychotherapy controlled trial research literature contains few detailed descriptions of the selection process involved in comparing newly developed treatments against controls. “Methods” sections of most such studies lack elaboration and often omit important information that can be obtained only through careful study (if available) of a control treatment manual. Unfortunately, as shown by Baskin et al.’s recent review of control

³As a point of caution, we would like to note that investigators conducting comparison studies must be on the lookout for the potential mismatch between the treatments from factors extrinsic to the design. For example, we made sure that the amount of supervision and attention to adherence was equally distributed between our experimental and comparison treatments. Luborsky et al. (1999) identified a number of ways in which the investigator’s allegiance may affect the outcome of a comparison study such as ours. Especially pertinent is the possibility that the researcher’s positive expectations for one treatment may boost the morale for therapists when working in this condition which may improve their performance with patients. Luborsky et al. (1999) suggest that therapists be supervised by those who represent that treatment mode. As this treatment was especially designed for this study, there are no such supervisors. However, a key member of the supervision team who had no allegiance to DBT carefully watched for the presence of any partiality by the investigator. This approach provides a firm foundation for clinical equipoise.

treatments (2003), a surprising number of control treatments failed to meet the researchers' rather basic definitions of equivalence (e.g. equal numbers and lengths of sessions).

Our study delineates our steps for designing a comparative therapy to answer whether DBT for BED is efficacious because of its specific intervention strategies and techniques or because of factors common to multiple treatment approaches. Clearly, the question of specific efficacy is one of theoretical and practical importance (Jones and Pulos, 1993). While our comparison treatment was designed specifically to control for the active elements of DBT and each active treatment requires its own control, we hope that our template of a comparison treatment design process will raise the standards for other research groups in the attempt to close the gap between theory and practice.

References

- Abraham SF, Beumont PJ. How patients describe bulimia or binge eating. *Psychological Medicine* 1982;12(3):625–635. [PubMed: 6957905]
- Agras WS, Telch CF, Arnow B, Eldredge K, Wilfley DE, Raeburn SD, et al. Weight-loss, cognitive-behavioral, and desipramine treatments in binge eating disorder: An additive design. *Behavior Therapy* 1994;25:225–238.
- Altshuler KZ. Will the psychotherapies yield different results?: A look at assumptions in therapy trials. *American Journal of Psychotherapy* 1989;53:310–320. [PubMed: 2683808]
- Baskin TW, Tierney SC, Minami T, Wampold BE. Establishing specificity in psychotherapy: a meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology* 2003;71:973–979. [PubMed: 14622072]
- Bowers TG, Clum GA. Relative contribution of specific and nonspecific treatment effects: meta-analysis of placebo-controlled behavior therapy research. *Psychological Bulletin* 1988;103:315–323. [PubMed: 3289070]
- Carter WP, Hudson JI, Lalonde JK, Pindyck L, McElroy SL, Pope HG. Pharmacologic treatment of binge eating disorder. *International Journal of Eating Disorders* 2003;34:S74–S88. [PubMed: 12900988]
- Chambless DL, Hollon SD. Defining empirically supported therapies. *Journal of Consulting and Clinical Psychology* 1998;66:7–18. [PubMed: 9489259]
- Ciano R, Rocco PL, Angarano A, Biasin E, Balestrieri M. Group-analytic and psychoeducational therapies for binge eating disorder: An exploratory study on efficacy and persistence of effects. *Psychotherapy Research* 2002;12:231–239.
- Classen, C., Diamond, S., Soleman, A., Fobair, P., Spira, J., & Spiegel, D. (1993). *Brief Supportive-Expressive Group Therapy for Women with Primary Breast Cancer: A Treatment Manual* Unpublished manuscript.
- Critelli JW, Neumann KF. The placebo. Conceptual analysis of a construct in transition. *American Psychology* 1984;39(1):32–39.
- Fairburn, C. G. (1995). *Overcoming Binge Eating* New York: The Guilford Press.
- Frank, J. D., & Frank, J. B. (1991). *Persuasion and healing: A comparative study of psychotherapy* (3rd ed.). Baltimore: Johns Hopkins University Press.
- Grencavage LM, Norcross JC. What are the commonalities among therapeutic common factors? *Professional Psychology: Research and Practice* 1990;21:372–378.
- Horvath P. Placebos and common factors in two decades of psychotherapy research. *Psychological Bulletin* 1988;104:214–225. [PubMed: 3054995]
- Kazdin AE. Reactive self-monitoring: the effects of response desirability, goal setting, and feedback. *Journal of Consulting and Clinical Psychology* 1974;42:704–716. [PubMed: 4427011]
- Kenardy J, Mensch M, Bowen K, Green B, Walton J. Group therapy for binge eating in Type 2 diabetes: A randomized trial. *Diabetic Medicine* 2002;19:234–239. [PubMed: 12162246]
- Klerman, G. L., & Weissman, M. M. (Eds.). (1993). *New Applications of Interpersonal Therapy* Washington, D. C.: American Psychiatric Press.

- Linehan, M. M. (1993). *Cognitive behavioral therapy of borderline personality disorder* New York: Guilford Press.
- Lohr M, DeMaoi C, McGlynn FD. Specific and nonspecific treatment factors in the experimental analysis of behavioral treatment efficacy. *Behavior Modification* 2003;27:322–368. [PubMed: 12841588]
- Luborsky L, Barber JP, Siqueland L, Johnson S. The revised helping alliance questionnaire (HAQ-II). *Journal of Psychotherapy Practice & Research* 1996;5:260–271.
- Luborsky L, DeRubeis RJ. The use of psychotherapy treatment manuals: a small revolution in psychotherapy research style. *Clinical Psychology Review* 1984;4:5–14.
- Luborsky L, Diguer L, Seligman D, Rosenthal R, Krause ED, Johnson S, et al. The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice* 1999;6:95–106.
- Marcus, M. D. (1997). Adapting treatment for patients with binge-eating disorder. In D. M. Garner & P. R. Garfinkel (Eds.), *Handbook of Treatment for Eating Disorders* (pp. 484–493). New York: The Guilford Press.
- Marcus MD, Wing RR, Fairburn CG. Cognitive behavioral treatment of binge eating vs. behavioral weight control on the treatment of binge eating disorder. *Annals of Behavioral Medicine* 1995;17:S090.
- Markowitz, J., & Sacks, M. (2002). *Manual for Brief Supportive Psychotherapy* Unpublished manuscript.
- Messer, S. B., & Warren, C. S. (1995). *Models of Brief Psychodynamic Therapy* New York: The Guilford Press.
- Parloff M. Placebo controls in psychotherapy research: a sine qua non or a placebo for research problems? *Journal of Consulting and Clinical Psychology* 1986;54:79–87. [PubMed: 3958304]
- Polivy, J., & Herman, C. P. (1993). Etiology of binge eating: Psychological mechanisms. In F. C. G & W. G. T (Eds.), *Binge eating: Nature, assessment, and treatment* (pp. 173–205). New York: Guilford Press.
- Rogers, C. R. (1951). *Client-centered Therapy* Boston: Houghton Mifflin.
- Stevens SE, Hynan MT, Allen M. A meta-analysis of common factor and specific treatment effects across the outcome domains of the phase model of psychotherapy. *Clinical Psychology: Science and Practice* 2000;7:273–290.
- Telch CF, Agras WS, Linehan MM. Group dialectical behavior therapy for binge eating disorder: A preliminary uncontrolled trial. *Behavior Therapy* 2000;31:569–582.
- Telch CF, Agras WS, Linehan MM. Dialectical behavior therapy for binge eating disorder. *Journal of Consulting and Clinical Psychology* 2001;69(6):1061–1065. [PubMed: 11777110]
- Weinberger, J. (1993). Common Factors in Psychotherapy. In G. Stricker & J. R. Gold (Eds.), *Comprehensive Handbook of Psychotherapy Integration* (pp. 43–56). New York: Plenum Press.
- Weisz JR, Hawley KM, Pilkonis PA, Woody SR, Follette WC. Stressing the (other) three Rs in the search for empirically supported treatments: review procedures, research quality, relevance to practice and the public interest. *Clinical Psychology: Science and Practice* 2000;7:243–258.
- Westen D, Novotny C, Thompson-Brenner H. The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin* 2004;130:631–663. [PubMed: 15250817]
- Wilfley DE, Agras WS, Telch CF, Rossiter EM, Schneider JA, Cole AG, et al. Group cognitive-behavioral therapy and group interpersonal psychotherapy for the nonpurging bulimic individual: A controlled comparison. *Journal of Consulting and Clinical Psychology* 1993;61:296–305. [PubMed: 8473584]
- Wilfley DE, Welch RR, Stein RI, Spurrell EB, Cohen LR, Saelens BE, et al. A randomized comparison of group cognitive-behavioral therapy and group interpersonal therapy for the treatment of overweight individuals with binge eating disorder. *Archives of General Psychiatry* 2002;59:713–721. [PubMed: 12150647]
- Wilfley DE, Wilson GT, Agras WS. The clinical significance of binge eating disorder. *International Journal of Eating Disorders* 2003;34:S96–S106. [PubMed: 12900990]

Table 1**Common Factors in Psychotherapy**

Common factor	Definition[*]
Therapeutic Alliance	Therapist(s) empathic support addresses patients' demoralization, helping them feel bolstered and understood. Alliance between patients with fellow group members adds additional support.
Rituals to be Observed/Healing Setting	These provide a focus or direction to follow as a group; can increase a sense of belonging.
Systems of Explanation/Therapeutic Rationale	Provision of a plausible explanation for patients' problems as well as procedures for resolving them. Utilization of theoretical framework supports therapists' sense of comprehension and enhances their view of themselves as therapeutic agents.
Opportunity to Express Emotions/Ventilate Problems	Patients are encouraged to become aware of and express their wishes, impulses and feelings in an accepting atmosphere.
Acquisition and Practice of New Behaviors	Patients attempt behaviors they had heretofore feared or not thought of. These likely include success experiences.
Positive Expectancy for Change/Therapeutic Optimism	Patients respond to treatment with positive expectation for symptom relief and arousal of hope.

* compiled from: Altshuler, 1989; Frank & Frank, 1991; Grencavage & Norcross, 1990; Luborsky & DeRubeis, 1984; Weinberger, 1993

Table 2
Similarities between Specific and Non-specific Treatments

Common Factors	DBT for BED	Non-specific Treatment for BED
Therapeutic Alliance	Same two co-therapists	Same two co-therapists
Rituals To Be observed/Healing Setting	Group format 20 sessions 2 hours/week Same time of day (e.g. Thurs 11am–1pm)	Group Format 20 sessions 2 hours/week Same time of day (e.g. Tues 11am–1pm)
Opportunity to Express Emotions/ Ventilate Problems	Group Format (including time for check-in, discussion, etc)	Group Format (including time for check-in, discussion, etc)
Therapeutic Rationale	Provided	Provided
Acquisition and Practice of New Behaviors	Monitoring: DBT diary card	Monitoring: Self-esteem diary card
Positive Expectancy for Change/ Therapeutic Optimism	Treatments presented with equal enthusiasm. Interdependent with other common factors (e.g. rituals).	Treatments presented with equal enthusiasm. Interdependent with other common factors (e.g. rituals).

Table 3

Differences between Dialectical Behavior Therapy (DBT) for Binge Eating Disorder (BED) and the Comparison Control Treatment

Common Factors	DBT for BED	Comparison Control Treatment for BED
Therapeutic Rationale	Explicit link between emotional dysregulation & binge eating.	No such systematic linkage.
Rituals To Be Observed/Healing Setting	Focused homework review. Didactic instruction of new skills.	General linkage between self-esteem and binge eating. Sufficient focus to keep group running smoothly (e.g. group "rules" about not interrupting, etc); No systematic homework review or didactic instruction of behavioral skills, etc.
Opportunity to Express Emotions/Ventilate Problems	Explicit focus on emotional dysregulation and binge eating. Didactic material presented.	General focus on self-esteem and binge eating. Expression of feelings and problems is encouraged. No systematic or explicit linkage between emotions and binge eating.
Acquisition and Practice of New Behaviors	Behavioral Chain Analysis Systematic review of diary card. Systematic teaching of behavioral skills. Group commitment to abstinence.	No Behavioral Chain Analysis No systematic review of diary card. No systematic teaching of behavioral skills. No group commitment to abstinence.
Therapeutic Rationale	Linkage between emotional dysregulation and binge eating.	General linkage between self-esteem and binge eating.
Positive Expectations and Hope for Improvement/Therapeutic Optimism	Therapists specifically rehearse positive expectations for change through eliciting patients' verbal commitment to eliminate binge eating and practice skills.	Therapists present goal of eliminating binge eating but do not elicit verbal commitment.

Specific Measurement of Common Factors

Table 4

Common Factor	Assessment Instruments
Therapeutic Alliance	<p>Helping Alliance Questionnaire-Patient (HAQ-P) (Luborsky, Barber, Sigueland, & Johnson, 1996). (Rated sessions 6, 12) This 19-item self-report questionnaire measures two main aspects of the therapeutic relationship--the experience of being understood/receiving a helpful attitude and the experience of being involved in a collaborative effort with the therapist. These scales predict therapeutic outcome in focal psychotherapy (Luborsky & DeRubeis, 1984). Rated on a 6 point Likert scale, scores range from a low of 19 to a maximum of 112. Patients rate the two co-therapists as a unit.</p> <p>Helping Alliance Questionnaire-Group (HAQ-G) (Rated sessions 6, 12) The 19 items on the HAQ-P were modified to assess the relationship between the patient and his/her fellow group members along the same main aspects of the HAQ-P in terms of being understood and involved in collaborative effort with fellow group members. Rated on a 6 point Likert scale, scores range 19 to 112. In addition, to assess for group support as a possible mediator, patients rate a weekly question from 1-6 ("A good relationship has formed with my fellow-group members").</p> <p>Therapist's Alliance with the Patient: (Rated sessions 4, 12) Therapists answer the question: "How would you characterize your alliance with the patient?" as (1="Very weak;" to 5 = "Very strong"). Lead therapist ratings are used.</p> <p>Diary Cards: Diary cards for both DBT and comparison therapy assess number of days patients filled them out ("once" to "daily"). Also tracked is not turning in a card at all.</p> <p>Patient rating of suitability of treatment (pre-treatment, Week 1, Post-Treatment): Patients rate "How suitable do you think this treatment is for your problems?" using a 10 point visual analogue scale from "Not suitable" to "Extremely suitable." Administered after pre-treatment orientation wherein rationales for both treatments are presented.</p> <p>Therapists' view of assigned treatment's suitability for patient (Week 4, 12) Therapists rate "Predict the likelihood that this patient will benefit from the particular treatment you are offering" from 1 = "Not very likely" to 5 = "Very likely"</p> <p>The HAQ-P and HAQ-G capture differences between the two groups in terms of ability to express pertinent emotional material with therapists and fellow group members</p>
Rituals to be Observed	<p>Diary Cards: Patients report the number of days they filled out their diary card each week</p>
Therapeutic rationale	<p>Patient expectancy of therapy (pre-treatment, Week 1, post-treatment): Patients rate: "How successful do you think your treatment here will be?" from 1 ("not at all successful") to 10 ("extremely successful")</p>
Opportunity Express Emotions	<p>Therapists' optimism about patients (Sessions 4, 12) "How much do you expect this patient to improve by the end of this 20 session treatment?" from 1 ("Not at all") to 5 ("Very much")</p>
Acquisition New behaviors Positive expectations and hope for improvement/Therapeutic optimism	

Table 5

Comparison of Credible Comparison Treatment with Dialectical Behavior Therapy (DBT) on Common Factors:
Means (and Standard Deviations)

Common Factor	Comparison	DBT
Therapeutic Alliance		
Session 6 (Patient Rating of Therapists) ^a	89.3 (9.3)	94.9 (13.6)
Session 6 (Patient Rating of Fellow Patients) ^a	4.9 (0.7)	5.1 (0.7)
Session 12 (Patient Rating of Therapists) ^b	94.6 (14.4)	96.8 (9.8)
Session 12 (Patient Rating of Fellow Patients) ^b	5.1 (0.9)	5.0 (0.7)
Session 4 (Therapist Rating of Alliance)	4.3 (0.8)	4.2 (0.9)
Session 12 (Therapist Rating of Alliance) ^c	4.0 (1.1)	4.3 (.9)
Systems of Explanation/Therapeutic Rationale		
Pretreatment (Patient Rating)	8.8 (1.4)	9.2 (1.4)
Session 1 (Patient Rating)	7.7 (2.3)	8.8(1.3)
Session 4 (Therapist Assessment of Suitability)	4.1 (0.8)	4.2 (0.8)
Session 12 (Therapist Assessment of Suitability)	3.9 (0.9)	4.3 (0.9)
Therapeutic Optimism		
Pretreatment (Patient Rating of Expectancy)	7.3 (1.6)	7.9 (2.1)
Session 1 (Patient Rating of Expectancy) ^d	7.0 (1.7)	7.7 (1.7)
Session 4 (Therapist Optimism about Patient)	4.0 (0.7)	4.1 (0.8)
Session 12 (Therapist Optimism about Patient)	3.7 (1.0) [‡]	4.3 (0.9) [‡]

Note. Refer to Table 4 for description of assessments.

* Difference between means $p=.0015$. No other comparisons reached statistical significance $p<.05$. $n=30$ per condition unless noted

^a $n=28$ (Comparison), $n=30$ (DBT)

^b $n=26$ (Comparison), $n=30$ (DBT)

^c $n=29$ (Comparison), $n=30$ (DBT)

^d $n=21$ (Comparison), $n=21$ (DBT)