From Efficacy to Effectiveness to Broad Implementation: Evolution of the Body Project

Carolyn Becker  
*Trinity University, cbecker@trinity.edu*

Eric Stice

Follow this and additional works at: [https://digitalcommons.trinity.edu/psych_faculty](https://digitalcommons.trinity.edu/psych_faculty)

Part of the [Psychology Commons](https://digitalcommons.trinity.edu/psych_faculty)

Publication Details  
*Journal of Consulting and Clinical Psychology*

Repository Citation  
From Efficacy to Effectiveness to Broad Implementation: Evolution of the Body Project

Carolyn B. Becker\textsuperscript{a} and Eric Stice\textsuperscript{b}

\textsuperscript{a}Trinity University

\textsuperscript{b}Oregon Research Institute

Abstract

\textbf{Objective}—At the turn of the millennium, eating disorders (ED) prevention was largely non-existent. No program had reduced future onset of EDs in even a single trial and most had not reduced ED symptoms. Sixteen years later, the ED prevention field has translated basic risk factor research into interventions with demonstrated efficacy and effectiveness in reducing ED risk factors and symptoms, as well as future ED onset in some trials. This paper reviews the aforementioned progress focusing on a model intervention (i.e., the Body Project (BP)).

\textbf{Method}—The paper is a qualitative review of the existing BP literature.

\textbf{Results}—Although clinical psychology has struggled with bridging the research practice gap and translating efficacy and effectiveness research into clinical implementation, researchers, clinicians, and community stakeholders working with the BP have made significant progress in addressing barriers to scalability, large-scale implementation, and sustainability, reaching 3.5 million girls and young women in 125 countries.

\textbf{Conclusion}—The successful evolution of the BP from basic research to traversing the efficacy-to-effectiveness-to-dissemination/implementation continuum has yielded significant public health impact and is thus a fitting case for this special issue commemorating the 125\textsuperscript{th} anniversary of the American Psychological Association. The BP example may help others broadly implement efficacious interventions for other mental health problems.

\textbf{Public Health Impact}—The Body Project (BP) has reduced eating disorder (ED) risk factors, symptoms, and future ED onset, though the latter effects were only observed in some trials; research also indicates that the BP can be broadly implemented using scalable strategies. To date, the BP has impacted the lives of over 3.5 million girls and young women in 125 countries thanks to strategic partnerships between researchers, clinicians, and community stakeholders.

\textbf{Keywords}

eating disorders; prevention; cognitive dissonance; risk factor; body dissatisfaction
Eating disorders (EDs) affect 13-15% of females and are marked by chronicity, relapse, distress, impairment, and increased risk for future obesity, depression, suicide, and mortality (Allen, Byrne, Oddy, & Crosby, 2013; Arcelus, Mitchell, Wales, & Nielsen, 2011; Stice, Marti, & Rohde, 2013). Unfortunately, 80% of individuals with EDs do not receive treatment (Swanson, Crow, Le Grange, Swendsen, & Merikangas, 2011), treatment only results in lasting symptom remission for 27-45% of patients (Agras, Walsh, Fairburn, Wilson, & Kraemer, 2000; Grilo, Crosby, Wilson, & Masheb, 2012; Lock et al., 2010) and treatment typically is less effective when delivered in real-world clinical settings (Weersing & Weisz, 2002) where dropout is often even higher (Merrill, Tolbert, & Wade, 2003). Thus, a public health priority is the development of effective ED prevention programs that can be broadly implemented (i.e., are cost-effective, palatable to community stakeholders, and scalable).

In 2000, the National Institute of Mental Health organized a roundtable to review the state of science regarding the reduction of ED risk factors (e.g., body dissatisfaction) and prevention of ED onset. In a summary of the findings, Pearson, Goldklang, and Striegel-Moore (2002) noted that no prevention program had reduced ED symptoms or prevented future onset of EDs. Sixteen years later, the state of the science in the ED prevention and body image field has radically changed (Watson et al., 2016). We now have programs that reduce ED risk factors and early stage ED pathology (Watson et al., 2016). We also have three programs that have significantly reduced future ED onset in at least one trial (Martinsen et al., 2014; Stice, Marti, Spoor, Presnell, & Shaw, 2008; Stice, Rohde, Shaw, & Gau, 2016; Stice, Rohde, Shaw, & Marti, 2013). In addition, although reviews suggest that many findings reported in psychology and other sciences fail to replicate (Ioannidis, Munafò, Fusar-Poli, Nosek, & David, 2014), most effects of the Body Project (BP) have generally replicated in most published studies. Finally, although the clinical psychology field often has struggled to translate efficacy/effectiveness research into large scale clinical implementation of empirically supported interventions (Kazdin, 2008), researchers working with the BP have made headway in addressing barriers to program scalability and global implementation with over 3.5 million girls/young women reached to date in 125 countries. As such, we consider a review of the successful evolution of the BP across the efficacy-to-effectiveness-to-dissemination/implementation continuum over the past 16 years a fitting contribution to this special issue commemorating the 125th anniversary of the American Psychological Association (APA). In addition to highlighting the specific successes associated with the BP, we believe there are general lessons that can be gleaned by other teams interested in translating clinical science into clinical impact.

This paper reviews five domains of success focusing on factors we believe facilitated progress as well areas where we see room for improvement. We first describe the development of the BP and review the empirical evidence base for the BP generated by the research team that developed it. We focus on the empirical support in reducing ED risk factors and ED onset, but also include basic risk factor research underpinning development of the BP and tests of the intervention theory. Next, we detail independent research from multiple teams that replicated findings in both efficacy and effectiveness trials. The subsequent two sections review two critical developments that allowed scale-up implementation of the BP, namely the adoption of task-shifting delivery of the program and
development of a train-the-trainers approach. Last, we describe examples of the broad implementation of the BP highlighting the importance of using community participatory approaches to create novel partnerships aimed at dissemination and implementation. Figure 1 provides an overview of the milestones in this research program.

It is important to note that one of the most robust ED risk factors, body dissatisfaction, is both a primary target of the BP in and of itself, as well a means to the end of reducing ED onset. Because body dissatisfaction is significantly more common than clinical EDs and associated with negative sequela and distress even in those who do not develop EDs (e.g., depression, suicidality after controlling for depression, unhealthy weight control behaviors and decreased quality of life: Bearman, Presnell, Martinez, & Stice, 2006; Crow, Eisenberg, Story, & Neumark-Stzainer, 2008; Jacobi & Fittig, 2010; Quick, Eisenberg, Buchianeri, & Neumark-Stzainer, 2013; Wilson, Latner, & Hayashi, 2013), many implementation partners are more interested in reducing body dissatisfaction than preventing EDs. We include implementation efforts that are aimed at either or both body dissatisfaction and ED prevention because many public health initiatives capitalize on community stakeholders’ multiple motivations for wanting to reduce a given risk factor (e.g., smoking) rather than just a single endpoint disorder (e.g., lung cancer).

The Body Project: Foundational Empirical Support

As noted, the state of the science in ED prevention was severely lacking 16 years ago. No program had yielded significant reductions in ED symptoms or future onset of EDs, two critical outcomes for ED prevention. Thus, researchers began to directly target empirically established risk factors found to predict future onset of ED symptoms/syndromes. If such risk factors truly were causal, programs that reduced said risk factors should result in a consequent reduction in ED onset. A handful of prospective studies at that time implicated the following factors as increasing risk for onset of binge eating, compensatory behaviors, and bulimia nervosa (BN): pressure to be thin, internalization of the thin-ideal standard of female beauty, body dissatisfaction, dieting, and negative affect (Killen et al., 1996; Patton, Johnson-Sabine, Wood, Mann, & Wakeling, 1990; Patton, Selzer, Coffey, Carlin, & Wolfe, 1999; Santonastaso, Friederici, & Favaro, 1999; Stice, Killen, Hayward, & Taylor, 1998; Stice & Agras, 1998).

Little consensus existed, however, regarding which risk factors were germinal versus resulting from upstream risk factors. Stice and colleagues, however, had conducted a) prospective studies investigating which risk factors predicted future increases in other risk factors (e.g., Stice, Mazotti, Krebs, & Martin, 1998) and b) randomized experiments that manipulated risk factors so as to examine effects on downstream risk factors and identify factors that amplified other risk factor effects (e.g., Stice & Shaw, 1994). This research suggested that sociocultural pressure for thinness promoted thin-ideal internalization, which in turn increased body dissatisfaction and consequent unhealthy weight control behaviors and negative affect. These, in turn, increased risk for emergence of ED symptoms, as proposed in the dual pathway model of ED development (Stice, 1994). A key implication of these findings was that it might be most fruitful to focus on reducing thin-ideal...
internalization, which was at the headwaters of the above risk factor cascade, and might be easier to change than sociocultural pressure for thinness.

Given the limited effects of first generation educational ED prevention programs, Stice decided to try non-educational approach; clinical experience suggested one promising avenue. After spending five sessions attempting to convince a client with anorexia nervosa (AN) to not engage in extreme dietary restriction without success, out of desperation, Stice switched roles with the client; he asked her to talk him out of having AN, parroting back arguments the client had provided to justify her ED. This exercise appeared useful because the client argued against herself, which seemed to reduce pursuit of thinness. Based on this experience, Stice created a dissonance-based ED prevention program wherein young women engaged in a series of verbal, written, and behavioral exercises aimed at exploring the negative effects of pursuing the thin ideal of female beauty. For instance, participants generate costs associated with pursuing the thin ideal, complete role-plays in which they talk intervention facilitators out of pursuing the thin ideal, and engage in acts of body activism that challenged the thin ideal (placing post-its saying “you are beautiful – no need to change your appearance” in diet books at a bookstore chain). These exercises putatively generate cognitive dissonance that prompts participants to reduce their pursuit of this unrealistic ideal because people are motivated to align their attitudes with their publically displayed behaviors (Festinger, 1957). According to the dual pathway model (Stice, 2001), reduced internalization of the thin ideal should decrease body dissatisfaction, unhealthy weight control behaviors, negative affect, and ED symptoms, as well future ED onset.

Three preliminary trials provided evidence that the dissonance-based program (i.e., the BP) produced significantly greater reductions in thin-ideal internalization, body dissatisfaction, dieting, negative affect, and ED symptoms relative to assessment-only/waitlist control conditions through 6-month follow-up, and in some cases relative to a credible comparison program that promoted healthy lifestyle modifications aimed at balancing energy intake and expenditure (Stice, Chase, Stormer, & Appel, 2001; Stice, Mazotti, Weibel, & Agras, 2000; Stice, Trost, & Chase, 2003). Based on these findings, Stice and colleagues initiated a fully powered, large-scale efficacy randomized controlled trial (RCT) in which young women with body image concerns were randomized to the BP, a refined version of the Healthy Weight program (HW), an expressive writing control group, or an assessment-only control group. Participants in the BP showed greater reductions in thin-ideal internalization, body dissatisfaction, dieting, negative affect, and ED symptoms relative to HW, expressive writing, and assessment-only controls by posttest, with many effects persisting through 1-, 2-, and 3-year follow-up (Stice et al., 2008; Stice, Shaw, Burton, & Wade, 2006). It is important to note that although the BP significantly outperformed the HW program acutely, the latter program produced significant reductions in ED risk factors and symptoms relative to the control conditions over longer-term follow-up. Both active groups yielded significantly lower onset of binge eating and compensatory behaviors relative to assessment-only, and significantly lower obesity onset over 1-year follow-up than expressive writing and assessment-only (Stice et al., 2006). The HW program may take longer to produce effects than the BP because participants make small refinements in diet and exercise and it takes time for these behavioral changes to impact ED risk factors.
Critically, among participants free of threshold or subthreshold EDs at baseline, onset of threshold/subthreshold EDs over 3-year follow-up was 6% in the BP condition and 6.5% in the HW condition, versus 15% in the assessment-only condition; this translates into a statistically significant 60% reduction in future ED onset respectively for both programs (Stice et al., 2008). Onset of EDs in the expressive writing control was 8%, which was not significantly different compared to either intervention. Overall, findings were notable because the BP appeared to be the first prevention program to reduce a composite measure of ED symptoms and future onset of EDs. It also was the first to produce stronger reductions in some outcomes relative to alternative interventions, which is important for establishing that the apparent intervention effects were not simply due to expectancy effects, demand characteristics, or non-specific effects inherent to RCTs.

A subsequent effectiveness trial found that the BP reduced ED risk factors and symptoms relative to educational brochure controls when clinicians at high schools recruited at-risk female students and delivered the BP under ecologically valid conditions (Stice, Rohde, Gau & Shaw, 2009; Stice, Rohde, Shaw, & Gau, 2011). A follow-up effectiveness trial also found that the BP reduced risk factors, ED symptoms, and functional impairment when college clinicians recruited students and delivered the BP: effects were 83% larger than in the high school effectiveness trial (Stice, Butryn, Rohde, Shaw, & Marti, 2013; Stice, Rohde, Butryn, Shaw, & Marti, 2015). The larger effects appeared to emerge because of use of an enhanced-dissonance version of the BP and improved procedures for selecting, training, and supervising clinicians. However, the BP did not significantly reduce future ED onset in these effectiveness trials, potentially because the facilitators only implemented the BP a few times compared to the facilitators in the efficacy trial.

In support of the intervention theory, reductions in thin-ideal internalization mediate the effects of the BP on symptom reductions (Stice, Marti, Rohde, & Shaw, 2011; Stice, Presnell, Gau, & Shaw, 2007). In support of the theory that dissonance-induction contributes to BP effects, participants assigned to high-dissonance versus low-dissonance versions of the BP, which contained the same content, showed greater reductions in ED symptoms (McMillan, Stice, & Rohde, 2011). For participants with initial elevations in thin-ideal internalization, the BP also produced larger effects, consistent with the thesis that they should experience the greatest dissonance induction (Stice, Marti, Shaw, & O’Neil, 2008). In addition, the BP reduced the risk conveyed by the most potent ED risk factor in one trial – denial of the costs of pursuing the thin-ideal (Stice, Rohde, Gau, & Shaw, 2012): among participants who denied these costs at pretest, those who completed the BP showed an ED incidence of 0% over 3-year follow-up. Finally, research indicates that the BP alters neural response to the ubiquitous thin-ideal stimuli thought to contribute to EDs: BP participants showed greater pre-post reduction in caudate (a key reward valuation region) responsivity to images of thin models relative to changes observed in educational brochure controls, implying that participants may not perceive the thin ideal as a desirable goal after completing the BP (Stice, Yokum, & Waters, 2015).
The Body Project and Independent Replication

Scholars have recently highlighted a “crisis” of replication within the field of psychology and other sciences (Ioannidis et al., 2014). Although clinical psychology intervention research maybe somewhat shielded against this secondary to the standard for independent replication in order to label an intervention as empirically supported, most ED prevention programs lack significant independent replication (c.f., Bird, Halliwell, Diedrichs, & Harcourt, 2013; Richardson & Paxton, 2010). Indeed, compared to other ED prevention programs, the BP has amassed the largest volume of independent research supporting its efficacy and effectiveness in reducing both ED risk factors and ED symptoms.

Matusek, Wendt, & Wiseman (2004) published the first independent replication only 4 years after Stice et al.’s initial publication. In this trial, Matusek et al. shortened the BP and HW scripts to one session and compared them to a waitlist control. Both interventions outperformed the control at one month in reducing ED risk factors. The following year, Becker, Smith and Ciao (2005) published a trial comparing the BP to a media advocacy intervention and waitlist control. Importantly, Becker et al. had no contact with Stice et al. prior to conducting this trial and developed their own intervention manual based on the BP description in Stice et al.’s earliest publication (2000), thus making this study as independent as possible. Both interventions significantly reduced body dissatisfaction, dieting and ED symptoms relatively to waitlist at 1-month follow-up; only the BP reduced thin-ideal internalization relative to waitlist. In the third study, Green, Scott, Diyankova, Gasser, and Pederson (2005) conducted an innovative trial that investigated whether dissonance played a role in intervention effects; they found that participants in the high-dissonance condition showed significantly lower ED symptoms than participants in the low-dissonance condition at posttest. They also found evidence that other elements of the program (e.g., content) contributed to effects. The short time period between the publication of the earliest Stice et al. papers and the emergence of independent trials is noteworthy as this was the first time an ED prevention program garnered independent replication support. It is likely that the rapid progression of the BP along the efficacy-effectiveness-dissemination/implementation continuum over the past 16 years was influenced by the relatively quick emergence of independent replication. In addition, although follow-up times were short, all trials had follow-up, used RCT design and had reasonable sample sizes (N range = 84-149). Further, Stice played no role in designing or running these trials, making them truly independent

Since these early replication studies, other published trials have supported the efficacy and effectiveness of the BP. For instance, Becker and colleagues conducted four additional trials with follow-ups of 8-14 months supporting the effectiveness of the BP when using community providers. Because these trials address scalability of the BP, in addition to the effectiveness of the BP when implemented by lay providers, they are described in detail below. Mitchell, Mazzeo, Rausch, & Cooke (2007) found that the BP produced greater reductions at post-intervention in body dissatisfaction, ED symptoms, drive for thinness, alexithymia and anxiety compared to both no-intervention and yoga control conditions. Seidel, Presnell, & Rosenfield (2009) replicated evidence that reductions in thin-ideal internalization mediated BP effects on symptom reductions. Halliwell & Diedrichs (2014) found that the BP not only reduced self-reported body dissatisfaction and thin-ideal
internalization in adolescent girls, but also eliminated the negative effect of supermodel exposure on body dissatisfaction observed in controls; the latter effects emerged in a trial that was ostensibly independent, making them unusually immune to demand characteristics. Other research groups also have conducted independent trials investigating various aspects and extensions of the BP (e.g., Cruwys, Haslam, Fox, & McMahon, 2015; Roehrig, Thompson, Brannick, & van den Berg, 2006; Serdar et al., 2014).

With regard to quantitative effects across trials, Table 1 describes studies that have evaluated variations of the BP in female-only samples along with effect sizes ranges. A description of the conservative approach used to calculate effect sizes is also described in Table 1. It should be noted that many of the studies that produced smaller effect sizes did not implement the Stice et al. (2000) version of the BP, but rather adapted versions that included only a subset of the dissonance-inducing exercises. In addition, some trials were effectiveness versus efficacy trials. Effectiveness trials often produce lower effect sizes due to their increased focus on external versus internal validity. For instance, in prevention effectiveness trials, less stringent enrollment criteria may result in lower baseline scores on relevant risk factors, which in turn increases floor effects. One trial (Atkinson & Wade, 2015) failed to find any significant effects for the BP for almost all ED risk factors. Four factors may account for these results. First, participant groups were markedly larger (i.e., 20-25) than is recommended (i.e., 6-10). Second, the program was implemented universally versus selectively. Although some other groups also have implemented on a more universal basis, this appears to result in lower effect sizes. Third, the intervention manual was adapted and made less scripted. Fourth, training in this trial was much lower than in other trials. Collectively, these factors may highlight what not to do when implementing the BP. With regards to reviews by other research groups, a recent independent meta-analytic review found that, averaging across all trials, dissonance-based programs produced the largest reductions in ED symptoms versus any other type of selective ED prevention program evaluated to date, with an average $d = 1.06$ when compared against assessment-only control conditions and an average $d = .30$ when compared against non-specific control conditions (Watson et al., 2016).

**Addressing the Scalability Problem: Task-shifting to Peer-Leaders**

Scalability refers to the degree that an intervention can be scaled-up to reach larger numbers of individuals without sacrificing effectiveness. Researchers encountered the scalability problem very early with the BP, when sororities at one university wanted to run all new members through the program at one time (Becker, Smith & Ciao, 2006). Although the actual scale of the plan was not large by many standards (i.e., the sororities wanted approximately 12 groups run simultaneously), because a) the university did not have a graduate psychology program, b) the research was not grant funded, and c) only one clinical psychologist was involved, the request produced the same problem encountered in many other global scalability conundrums (Kazdin & Blase, 2011). Specifically, Becker et al. lacked sufficient providers to deliver the program as requested by community stakeholders cost-effectively. Connecting this to the larger field of clinical psychology, Kazdin and Blase (2011) have noted that even if every licensed clinician delivered the most efficacious psychological treatment possible to each patient, we would still fail to adequately reduce the...
global burden of mental illness because we will never have enough licensed (i.e., expensive) providers for everyone who could use treatment. Although this is a significant problem in the United States (US: Kazdin & Blase, 2011), the scalability problem is even more glaring in resource poor countries (Fairburn & Patel, 2014). In sum, sorority demand at a small university pushed BP researchers to confront the same scalability issue faced by researchers working quite different settings, and to confront it early in the testing of the BP.

To address the scalability problem, Becker et al. (2006) tested whether BP delivery could be task-shifted to undergraduate peers, who were abundant and inexpensive. Task-shifting has been used elsewhere in healthcare to address scaling problems (e.g., providing mental health services and HIV treatment in resource poor countries: Patel, Chowdhary, Rahman & Verdeli, 2011; Vamos et al., 2014), but had not been used previously with the BP or any similar program. Several studies supported task-shifting delivery of the BP (Becker et al., 2006; Becker, Bull, Schaumberg, Cauble, & Franco, 2008; Becker et al., 2010), with 8- and 14-month within-group effect sizes for ED risk factors and symptoms being similar in magnitude to those reported in the Stice et al. (2006; 2008) trials using benchmarking (McFall, 1996). Importantly, using RCT methodology, Becker et al. (2006) also found that peer-led BP outperformed a peer-led media advocacy intervention in reducing dieting, thin-ideal internalization, and body dissatisfaction at 8-month follow-up; both interventions reduced ED symptoms. This supported the contention that effects for the peer-led BP were not merely due to expectancy effects, demand characteristics, or non-specific factors. Further, Becker et al. (2010) found a similar pattern of findings to Stice et al.’s (2006) clinician-led study: in a RCT, peer-led BP outperformed peer-led HW at post-test, although both programs yielded similar significant decreases in outcomes by 14 months. Perez, Becker & Ramirez (2010) also found that effects at 5-month follow-up largely held when the task-shifting model was transported to a different campus with a different trainer. Combined, these results laid the foundation of empirical support for task-shifting BP delivery to peers and provided additional evidence for the effectiveness of the BP with longer follow-ups.

Subsequent research by Stice and colleagues also provided support for task shifting to peer-leaders as long as materials and training were sufficiently adapted to meet peers more limited facilitator skills. Without such support, peers yielded somewhat smaller effects than clinicians (Stice, Rohde, Durant, Shaw, & Wade, 2013). More recently, in the largest BP trial (N = 680), adolescent girls and young women were randomly assigned to clinician- or peer-led BP groups, an Internet-delivered dissonance-based intervention, or an ED education video condition (Stice, Rohde, Shaw & Gau, 2016). Both clinicians and peers produced similar effects, again supporting task-shifting. Both clinicians- and peers also outperformed the video control yielding greater reductions in thin-ideal internalization, body dissatisfaction, dieting, and ED symptoms at post, with some effects persisting through 7-month follow-up. Finally, among participants free of threshold/subthreshold EDs at baseline, 2.2% of those in peer-led groups showed future onset of a threshold/subthreshold ED at 7-months compared to 8.4% of those in the Internet condition, which translated into a significant 74% reduction in future ED onset. There was a 64% reduction in future ED onset in peer-led groups compared to the video control, but this effect did not reach significance (p = .115). This trial is important for several reasons. First, it provides randomized evidence that BP delivery can be task-shifted to peers without loss of effectiveness relative to
clinicians. Second, although Becker provided some early consultation on Stice et al.’s first peer-led trial, the second peer-led trial was run completely independent from the developer of the peer-led approach (i.e., Becker) providing independent replication for task shifting. Third, this trial provides the first evidence that peer delivery not only reduces ED risk factors including body dissatisfaction and early ED pathology, but also can reduce future onset of EDs. Finally, the fact that effects were stronger for participants in the group-based BP conditions versus the Internet condition suggests that the group-based nature of BP contributes to its effects, possibly because the public accountability afforded by the group promotes greater dissonance-induction and because participants provide social support for each other (Shaw, Rohde, & Stice, 2016).

As highlighted below, development of the task-shifting approach to delivery of the BP has played a critical role in successful dissemination and implementation of the BP on a large scale throughout the US. Over 135 universities have implemented the BP since 2008 with the vast majority utilizing task-shifting. Universities have indicated they prefer task-shifted BP delivery for two primary reasons. First, task shifting addresses the fact that most college counseling centers do not have enough staff to implement prevention programs, and instead have to focus on crisis management (Gallager, 2014). Thus, a typical university cannot afford to scale-up the BP by relying on clinician delivery of the program. Second, task-shifting facilitates the creation of leadership opportunities for undergraduates, also filling part of the educational mandate of universities. This allows BP implementation teams to tap into financial resources from non-counseling center departments (e.g., student affairs, wellness centers, peer educator programs).

Scaling-up Further: Development of a Train-the-Trainers Model

In 2005, the Tri Delta national sorority contacted Becker after learning about her work with the BP and sororities at Trinity University. Tri Delta was interested in the existing evidence-based for the BP, its focus on body image, and the task-shifting approach. Tri Delta was drawn to task-shifting for two reasons. First, consistent with other university-based partners, Tri Delta staff knew they would be facing issues of scalability given that they typically approximate 13,500 collegiate members distributed over 130+ university chapters. Second, task-shifting fit their values. The four pillars of sorority life include service, leadership, scholarship and sisterhood. Thus, the peer-led approach resonated with Tri Delta’s core values in providing leadership opportunities via a program that had the potential to enhance sisterhood by addressing a problem (i.e., body dissatisfaction) that was viewed as decreasing quality of life, increasing risk for mental illness (i.e., depression and EDs), and causing competition between some members.

After two years of collaborative pilot testing, Tri Delta and Becker embarked on an unprecedented implementation effort for that time: they set the goal of reaching 20,000 young women with the BP over 5-year period. After re-branding the peer-led BP as the Reflections: Body Image Program™, Tri Delta and Becker launched their plan in 2008 in conjunction with a marketing advocacy campaign called Fat Talk Free Week™. Between 2008 and 2012, 15,000 workbooks were distributed and Reflections was implemented at approximately 100 university campuses. As the BP is the most commonly used name for the BP.
dissonance-based intervention, for the remainder of this paper we use the term the **BP** instead of Reflections.

During the development of the Tri Delta implementation plan, however, it became apparent that task-shifting alone was insufficient to scale up delivery as envisioned. Specifically, although Becker identified three additional peer-leader trainers, all four had regular jobs and did not have sufficient time to reach all of the universities targeted by Tri Delta. To resolve this scalability problem, Becker and colleagues turned to a train-the-trainer (TTT) approach to enhance reach. Again, scalability concerns for the **BP** mirrored those that have emerged elsewhere in clinical psychology. Specifically, lack of clinical trainers/supervisors has been identified as a key barrier to scalability of psychological treatments (Fairburn & Patel, 2014).

Becker and colleagues conducted two studies investigating the blending of the TTT approach with task-shifting. Kilpela et al. (2014) randomly assigned participants to the **BP** delivered by peer-leaders who either had been trained by Becker or by undergraduate RAs who had been trained to train their fellow peers as peer-leaders. At 14-month follow-up few differences emerged between groups providing support for the TTT approach.

Because the specific TTT model used by Tri Delta relied on an intensive 2-day training model, Greif, Becker, & Hildebrandt (2015) explored whether a graduate student could be trained to train peer leaders in that condensed 2-day time period in an open trial that benchmarked findings back to previous peer-led trials. Results indicated that peer-led **BP** produced similar and significant within-group reductions in ED symptoms, body dissatisfaction and thin-ideal internalization at 5-month follow-up using the 2-day TTT training.

**Global Implementation Examples**

A global community of researchers, clinicians, community stakeholders currently supports implementation of the **BP**. Dating back to the early work by Becker and colleagues with sororities, the **BP** community heavily relies on principles drawn from community participatory research methods (CPR: see Becker, Stice, Shaw, & Woda, 2009). We use principles from CPR to inform how we approach all partnerships regardless of whether or not implementation efforts include research. Some common components of CPR methods include a reliance on sharing decision making and power, building on partner strengths, striving to create equitable and collaborative partnerships; promoting capacity building and co-learning for all; balancing the joint aims of research with intervention implementation; using a collaborative, cyclical and iterative process; sharing results in a way that respects all parties; and developing a commitment to sustainability, partners, and the project (Israel, Eng, Shculz, & Parker, 2005). In this section we describe four ongoing implementation initiatives: two based in the US and Canada, one is based in Mexico with expansion to other parts of Latin America, and one global.
Eating Recovery Foundation (ERF)

ERF is the non-profit arm of the for-profit residential Eating Recovery Center, which treats EDs. This partnership piloted in 2013. The ERF sought to partner with the BP for two reasons. First, the BP's strong empirical base and scalable delivery models suggested the donated financial resources would not be wasted and would yield a positive and significant impact. Second, it appeared that, pursuant to reason one, the BP had the potential to increase the brand equity for both the non-profit foundation and the for-profit company.

The ERF partnership targets task-shifted implementation at the university-level in both the US and Canada by providing universities with grants to cover the majority of the cost of a 2-day, on-site, TTT training. These trainings can train approximately 12 peer-leaders and 4-6 staff trainers. During the training a starter cohort of peer-leaders is trained simultaneously in conjunction with staff, who then are positioned to sustainably train subsequent generations of peer-leaders. Thus, use of the TTT model means that a university does not need to pay for repeat trainings as long as at least one trained staff person remains, improving cost efficiency.

The initial 18-month, 2-phase pilot facilitated expansion of the BP to 10 new universities. At the end of the pilot, the ERF agreed to fund 85 new grants over a 5-year period. To date, across both the pilot and first year of the 5-year period, 355 student peer-leaders and 223 staff from 41 universities have been trained to implement the BP at their respective universities. Data collection regarding numbers of participants actually reached is ongoing as is a study investigating potential benefits of completing the program training with both staff and peer-leaders. Qualitative feedback gathered at trainings indicate overwhelmingly positive responses. It is important to note that the ERF grants facilitate dissemination of the BP to universities that either lack financial resources for the initial training or for whom the training costs reduce enthusiasm for adoption. Many other universities have both the financial resources and motivation to absorb the initial training expense.

National Eating Disorders Association (NEDA)

NEDA is the largest grassroots ED association in the US. The partnership with NEDA launched in 2012 with the aim of using NEDA's existing infrastructure to facilitate implementation of the BP. NEDA was attracted to the BP secondary to a longstanding goal of advancing prevention and because it believes that the BP can help create more pro-recovery environments for people with EDs. To reduce long term costs for NEDA, the TTT approach was adapted to train not only facilitator trainers but also two NEDA “master trainers,” who in turn could train more trainers. To date, NEDA has trained 165 undergraduate peer leaders to implement the BP in low SES high schools throughout the state of New York. One indication of program acceptability in this population is the fact that the students have asked for both a longer version of the program and follow-up activities for monthly booster meetings. NEDA is currently piloting the use of both. NEDA also offers the BP trainings (both regular trainings for clinicians and the TTT training for those who want to train peer-leaders) prior to the start of their annual convention. This has increased the national profile of the BP to the point that new partnerships are starting to arise via clinicians trained by NEDA.
Comenzar de Nuevo (CdN)

CdN is a non-profit residential ED treatment program located in Mexico. In 2014, CdN decided to spearhead dissemination of the BP throughout Mexico and other parts of Latin America. They also translated the manual into the variant of Spanish commonly used in Mexico. In January 2015, they invested in 5 days of training with Becker and Perez with the aim of building marked depth in terms of trainers and group facilitators. The 5-day training also was used to train three master trainers. The goal in providing knowledgeable organizational partners (e.g., NEDA and CdN) with master trainers is to help them become self-sufficient in scaling up the BP. This is important for two reasons. First, the more organizations can effectively scale up the BP on their own, the greater the global reach of the program. Second, by not requiring partner organizations to constantly pay developers for more training, we demonstrate that our joint goal is greater implementation of the BP versus greater income. Multiple partners have stated that by focusing on financially sustainable (versus highly profitable) training models, they feel more motivated to invest resources in helping to disseminate and implement the BP.

In total, CdN started with almost 30 group facilitators, 5 trainers and 3 master trainers. They spent the first 6-9 months after training developing comprehensive business and marketing plans. This was key because CdN recognized that many schools in Latin America would not be able to pay a training fee, yet CdN needed a sustainable business model. To date, CdN has launched a website (over 1,300,000 hits to date) as well as social media campaigns aimed at building interest in both the program and fundraising initiatives; their Facebook page has over 35,000 likes and currently is adding approximately 2000 per month. CdN has trained 2 self-paying private universities (25 group facilitators and 8 trainers). Both universities agreed to collect effectiveness data, and this is ongoing. The two universities each aim to reach 300 students per year. In addition, CdN trained facilitators and trainers at a public university via their “Godfather Program” which allows businesses to sponsor less financially able institutions. Via this funding program, CdN has arranged for an additional 5 university trainings. CdN is also outreaching to area high schools. Finally, they have conducted a training in Chile and the Dominican Republic; trainings are scheduled in 4 new cities in Mexico, as well as Argentina, Ecuador, San Salvador.

Dove and the World Association of Girl Guides and Girl Scouts

The Dove Self Esteem Project's global social mission to promote body confidence in girls was launched in 2006. In 2012, Dove began partnering with the Center for Appearance Research at the University of West England. A key component of this partnership was Dove's desire to strongly pivot towards an evidence-based approach, using an empirically supported body image intervention as the foundation for its body image efforts. Spearheaded by Diedrichs, and in partnership with the World Association of Girl Guides and Girl Scouts (WAGGGS), Dove commissioned the adaptation of the BP into a manualized intervention that could be delivered by troop leaders. Developed by Diedrichs, Stice and Becker in collaboration with WAGGGS staff, Free Being Me (http://www.free-being-me.com/downloads/) was created with the aim of bringing the BP to millions of girls globally with delivery task-shifted to existing WAGGGS troop leaders. Free Being Me also uses the TTT
model. In sum, this global endeavor relied heavily on previous research supporting the efficacy, effectiveness, and scalability of the BP.

_Free Being Me_ truly represents the three-legged stool of evidence-based psychological practice (Spring, 2007) applied globally. With respect to the first leg, which many people argue is the most important (Lilienfeld, Ritschel, Lynn, Cautin, Latzman, 2013), _Free Being Me_ is, at its core, simply an adaptation of the empirically supported BP. Pursuant to the second leg, clinical expertise, Diedrichs, Stice and Becker relied on thousands of hours of experience implementing the BP in providing advice to WAGGGS and Dove regarding which elements of the manual could be adapted and which could not. Becker and Diedrichs also brought their training experience to the table in helping WAGGGS host the first global _Free Being Me_ training summit at their world headquarters in the United Kingdom. Finally, regarding the third leg, client values, WAGGGS collaborators ensured that _Free Being Me_ “looked like a WAGGGS program” so that it would be better embraced by its member organizations (national Guiding/Scouting organizations). WAGGGS also worked with Becker and Diedrichs in designing the plan for adapting the TTT model to their organization.

To date, WAGGGS and Dove report reaching 3.5 million girls in 125 countries with the _Free Being Me_ over the first 3 years of implementation. _Free Being Me_ has been translated into over 19 languages and implemented in countries previously overlooked in academic body image research (e.g., Rwanda). Diedrichs and Stice are currently collaborating with WAGGGS and Dove to evaluate the adoption process, with the goal of improving future implementation efforts, and the effectiveness of _Free Being Me_ when implemented on this unprecedented scale. Preliminary results indicate that virtually all Member Organizations (e.g., Girl Guide Chapters) in the 55 countries evaluated were able to complete all stages of the implementation process, which included committing to implement _Free Being Me_, holding national leader trainings, conducting local trainings, implementing _Free Being Me_ to girl guides/scouts, and discussing expansion and sustainability plans. Equally important, data collected from girl guides/scouts in 12 countries from an ongoing study has confirmed that _Free Being Me_ is producing significant reductions in thin-ideal internalization, negative mood, self-esteem, body dissatisfaction, disordered eating, avoidance of life activities because of body image concerns, and significant improvements in social adjustment from pre-to-post compared to waitlist control girls, though effects vary by country (Stice & Diedrichs, 2016). Based on these positive findings, Dove and WAGGGS has committed to implementing _Free Being Me_ for another 4 years.

**Critical Lessons**

It is our impression that our experiences may offer learning opportunities of value for researchers interested in translating clinical science into clinical impact. First, “giving away” the intervention script for free to other researchers and clinicians was critical in promoting broad BP implementation and in encouraging independent replication. Second, training people in other organizations to sustainably train facilitators and trainers, a practice which ultimately cuts us out of the financial loop, appears to have been key. Third, the use of rigorous research methods, such as random assignment, comparing the BP to other active
interventions, use of diagnostic interviews and objective outcomes, and long-term follow-up used in several trials created a strong evidence-base, which has promoted adoption and implementation. Fourth, in line with CPR tenets, we have solicited input from BP group leaders and participants to continuously refine and improve the BP over the past 16 years, which has played a key role in the acceptability and efficacy of this program. Another part of CPR entailed soliciting input from implementation partners regarding how best to scale up delivery. Fifth, encouraging tailored adaptations of the BP for targeted populations, such as sorority members and WAGGGS, has facilitated broad implementation. Importantly, all of these adaptations are made available for free via a website. Providing materials in this manner allows the development of as many iterations as our community partners feel they need and shows that we are responsive to requests for tailored materials. It also overcomes language barriers. Fifth, identifying ways in which the BP could address multiple goals (e.g., reducing body dissatisfaction and provide leadership opportunities) appears to help partners access more resources and enhances motivation for implementation. Finally, it is commonly argued that technology offers the best path forward in dissemination and implementation (e.g., Fairburn & Patel, 2014). Yet our experiences with the BP suggest that the combined benefit of task-shifting and the TTT approach may be underappreciated. Our community partners also regularly tell is that they appreciate the low tech group approach, which brings people together face-to-face.

Limitations, Unanswered Questions, and Future Directions

There are several limitations with the evidence-based reviewed above. First, replication trials conducted by independent groups often do not include long-term follow-ups and were not designed to detect reductions in future ED onset. It would be useful for future studies to use follow-up periods longer than 3-years, which is the longest follow-up to date. Second, it is unclear why the BP significantly reduced future ED onset in the large efficacy and implementation trials, but not in the two large effectiveness trials. One possible explanation is that the facilitators in the effectiveness trials did not develop deep expertise in implementing the BP because they implemented it fewer times than the facilitators in the other trials. Third, the vast majority of BP trials have not included objective biological measures, which is needed to rule out the possibility that demand characteristics were responsible for intervention effects on the self-report measures. Fortunately, the fact that the BP reduced obesity onset (Stice et al., 2006), fMRI-assessed reward region response to thin models (Stice, Yokum, & Waters, 2015) and EKG-assessed markers of cardiac abnormalities (Green et al., 2016) provides some evidence that the BP affects objective outcome biological outcomes. It is important to note, however, that the Green et al. EKG trial was conceived as indicated not selective prevention. Third, most of the broad implementation efforts are not collecting data that would allow them to confirm if the implemented variants of the BP are producing all of the intervention effects observed in highly controlled trials, such as a reduction in future ED onset. One barrier to this is that some of our implementation partners do not view reduction of future ED onset a major goal (i.e., they care more about reducing body dissatisfaction), and detecting reductions in ED onset requires more expensive studies with control groups, longer follow-ups, and substantial research funding.
In terms of unanswered questions, although the BP has been compared to several alternative interventions (e.g., HW and media advocacy), it has yet to be compared to cognitive behavioral therapy-based ED prevention programs, which also have a broad evidence-base. Further, it would also be useful for future trials to evaluate alterations that might increase the efficacy of the BP, such as adding on intervention elements that focus on reducing negative affect, as this risk factor is not directly addressed in the current version of the BP.

With regard to future directions, more robust testing of the BP during large scale implementation is needed, with longer follow-ups. Also, prevention programs that expressly target AN are needed, as the risk factors targeted in the BP predict future onset of BN, binge eating disorder, and purging disorder, but not future onset of AN (Stice, Gau, Rohde, & Shaw, 2016). Another direction would be to design programs that target trans-diagnostic risk factors that predict onset of AN and other EDs, such as impaired psychosocial functioning (Stice, Gau et al., 2016). Finally, it might be useful to design dissonance-based interventions for other health and mental health problems, as this modality appears to produce robust and enduring effects.

Conclusion

The past 16 years have proven to be remarkable ones for the ED prevention and body image fields. Building on basic risk factor research, researchers have now identified strategies that effectively and reliably reduce ED risk factors, including body dissatisfaction and early stage ED symptoms, and have prevented the onset of EDs in some trials. Moving from basic research to traversing the efficacy to effectiveness continuum alone would have been a significant accomplishment in less than 2 decades. Yet, by working collaboratively towards a common goal with a single intervention, a global community of researchers, clinicians, community stakeholders also demonstrated, in the same time period, that it is possible to translate clinical psychological science into tremendous clinical impact by attending to issues of scalability and CPR methods. Indeed, it is even possible that engaging entire organizations/communities in these implementation projects has increased awareness that narrowly defined beauty ideas contribute to body dissatisfaction and EDs across a much wider spectrum of individuals than just the participants who have completed variants of the BP. This evolving implementation story provides evidence that clinical psychology can overcome the research practice gap, produce effects that are replicable, and even enact global-level evidence-based practice, but we will need many more partners at the table to do so. With that in mind, we want to end by thanking everyone, from the BP participants and group leaders to the funders, visionaries, scholars and clinicians, who made this journey possible. Everything that has been accomplished and described in this paper belongs to and exists because of the entire the BP community.

Acknowledgments

This program of research has been supported by NIH grants K01MH01708, R01MH/DK061957, R01MH070699, R01MH086582, R01MH097720, funding from the Hogg Foundation, Dove/World Association of Girl Guides and Girl Scouts, The Eating Recovery Foundation, Comenzar de Nuevo, Delta Delta Delta Fraternity, National Eating Disorders Association, New York State, Stavros Niarchos Foundation. The authors would also like to thank all members of the Body Project community described in this paper.
References


J Consult Clin Psychol. Author manuscript; available in PMC 2018 August 01.


McFall, RM. Consumer satisfaction as a way of evaluating psychotherapy: Ecological validity and all that versus the good old randomized trial (panel discussion). 6th Annual Convention of the American Association of Applied and Preventative Psychology; 1996; San Francisco. 1996.


J Consult Clin Psychol. Author manuscript; available in PMC 2018 August 01.


J Consult Clin Psychol. Author manuscript; available in PMC 2018 August 01.


### Figure 1. Timeline of the Development of the Body Project

<table>
<thead>
<tr>
<th>Year</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>Foundational support and empirical research are based on publication dates.</td>
</tr>
<tr>
<td>2006-2008</td>
<td>Large scale efficacy trial (2006-2008)</td>
</tr>
<tr>
<td>2006-2010</td>
<td>Becker lab scalability task-shifting peer leader trials (2006-2010)</td>
</tr>
</tbody>
</table>

**Key Dissemination and Implementation Events and Partnerships:**

- Partnership with local Sororities at Trinity University (2001-2016)
- Partnership with Tri Delta National Sorority (2005-2012)
- NEDA partnership (2012-present)
- EER partnership (2014-present)
- CNB partnership (2014-present)
- Dove/WAGGGS partnership (2014-present)

Note: Dates for foundational empirical support and independent research are based on publication dates. Dates for partnerships indicate start and end of partnership. TTT = train-the-trainer; NEDA = National Eating Disorders Association; EER = Eating Recovery Foundation; CNB = Comenzar de Nueve; WAGGGS = World Association of Girl Guides and Girl Scouts.
Table 1
Summary of the Sample, Intervention Content, and Findings from Female-Only Dissonance-Based Program Trials with Effect Sizes

<table>
<thead>
<tr>
<th>Study</th>
<th>Sample</th>
<th>Intervention</th>
<th>Findings &amp; Effect Sizes (Cohen’s d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atkinson &amp; Wade (2016)</td>
<td>347 adolescent girls</td>
<td>Randomized trial comparing 3-session dissonance-based intervention (DBI) to mindfulness-based intervention or control</td>
<td>DBI showed significant reductions in sociocultural pressures compared to mindfulness-based intervention and an assessment only control (effects collapsed at all time points) DBI vs control post $d = -0.06-0.32$; 1-mo $d = -0.10-0.32$; 6-mo $d = -0.03-0.41$; DBI vs mindfulness intervention post $d = -.03-0.41$; 1-mo $d = -0.09-0.33$; 6-mo $d = -0.20-0.03$.</td>
</tr>
<tr>
<td>Becker, Smith, &amp; Ciao (2006)</td>
<td>90 sorority members</td>
<td>RCT comparing a 4 hr. peer-led DBI to a MA intervention</td>
<td>DBI participants showed greater reductions in TII, BD, and dieting, but not ED symptoms, than media advocacy participants at 7-wk and 8-mo follow-up. Post $d = -.07-.36$; 1.75-mo $d = 0.88-0.50$.</td>
</tr>
<tr>
<td>Becker, Bull et al. (2008)</td>
<td>188 sorority members</td>
<td>RCT comparing a 4 hr. peer-led version of DBI to MA intervention</td>
<td>DBI participants did not show greater reductions in TII, BD, dieting, or ED symptoms than MA participants. Post $d = .02-.19$; 1.75-mo $d = -0.02-.11$; 8-mo $d = -0.13-.17$.</td>
</tr>
<tr>
<td>Becker, Wilson et al., (2010)</td>
<td>106 sorority members</td>
<td>RCT comparing a 3.5 hr. peer-led version of the DBI to a modified version of the Healthy Weight (HW) management intervention</td>
<td>DBI showed greater pre-post reductions in TII, negative affect, and ED symptoms, but not BD and dieting, than modified HW participants, but no between group effects persisted through 8-wk, 8-mo, or 14-mo FU. Post $d = .19-.44$; 2-mo $d = .06-.13$; 8-mo $d = -.04-.36$; 14-mo $d = 0.08-.50$.</td>
</tr>
<tr>
<td>Becker, McDaniel, et al. (2012)</td>
<td>157 female athletes</td>
<td>RCT that compared the 4 hr. peer-led version of the DBI to an athlete-modified version of the HW intervention</td>
<td>DBI participants did not show greater reductions in TII, BS, dieting, negative affect, or ED symptoms than HW participants. Post $d = .13-.21$; 2-mo $d = .06-.13$; 8-mo $d = -.04-.36$; 14-mo $d = 0.08-.50$.</td>
</tr>
<tr>
<td>Ciao et al. (2015)</td>
<td>51 high school girls</td>
<td>Non-randomized trial comparing a 3 hr. peer-led version of DBI to a waitlist control with high school girls participating on a mandatory basis within school curriculum</td>
<td>DBI participants did not show significantly greater reductions in outcomes than waitlist controls, though this study was only powered to detect large effects. Between group post $d = .08-.26$; within group 3 mo follow-up $d = .02-.15$.</td>
</tr>
<tr>
<td>Corning et al. (2010)</td>
<td>31 middle school girls and their mothers</td>
<td>RCT comparing 6 hr. version of DBI based on Trost (2006) with a waitlist control with middle school girls with body image disturbance and their mothers</td>
<td>DBI did not show significantly greater reductions in outcomes than waitlist control, though this study was only powered to detect large effects. Post $d = .17-.32$; within group 3-mo $d = .48-.55$.</td>
</tr>
<tr>
<td>Cruwys et al. (2015)</td>
<td>112 girls and women</td>
<td>Uncontrolled study examining mechanism of action in group therapy via a test of the DBI</td>
<td>DBI participants showed a significant pre-post decrease TII, BD, and dieting. Within group post $d = .68-.144$.</td>
</tr>
<tr>
<td>Danielsdottir et al. 2012</td>
<td>62 high school girls</td>
<td>RCT comparing the 4 hr Body Project DBI (2007) with the Reflections DBI (2011)</td>
<td>Both DBI conditions showed significant reductions in TII, BD, dieting, and ED symptoms; effects did not differ between groups. Study 1 within group post $d = .21-.72$; study 2 post $d = .04-.81$.</td>
</tr>
<tr>
<td>Green et al. (2005)</td>
<td>155 college females</td>
<td>RCT comparing 4 hr. high-dissonance DBI to a low-dissonance DBI and control condition</td>
<td>High dissonance DBI showed lower ED symptoms compared to low-dissonance DBI at post. Post $d = .15$; 1-mo $d = .17$.</td>
</tr>
<tr>
<td>Study</td>
<td>Sample</td>
<td>Intervention</td>
<td>Findings &amp; Effect Sizes (Cohen's $d$)</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Greif et al. (2015)</td>
<td>64 sorority members</td>
<td>Uncontrolled study examining effects of a 4 hr. peer-led DBI</td>
<td>DBI showed significant reductions in TII, BD, and ED symptoms at post, significant reductions in TII and ED symptoms at 5-months, no improvements in negative affect. Post $d = .17-.93$; 5-mo $d = -.43-.39$.</td>
</tr>
<tr>
<td>Halliwell &amp; Diedrichs (2014)</td>
<td>106 middle school girls</td>
<td>Non-randomized trial comparing a 1.3 hr. version of the DBI to a waitlist control condition</td>
<td>DBI showed greater reductions in TII and BD, but not dieting, than controls at post-test, and less negative effects of exposure to thin models relative to control at 1-mo follow-up. 1-mo $d = .14-.55$.</td>
</tr>
<tr>
<td>Kilpela, DeBoer et al. (2015)</td>
<td>73 female undergraduates</td>
<td>Uncontrolled trial comparing 4 hr session over 4 week version of peer-led DBI to two 2 hr session version of peer-led DBI</td>
<td>Participants in both DBIs showed similar rates of improvement in ED risk factors and symptoms through 12 mo-follow-up except for TII, where participants in 4 week version had lower scores at 6-and 12-mo-follow-ups. 4 hr sessions; within group post $d = .04-.56$; 6-mo $d = .07-.51$; 12-mo $d = .05-.80$; 2 2-hr sessions; within group post $d = .25-.41$; 6-mo $d = .03-.62$; 6-mo $d = .12-.61$.</td>
</tr>
<tr>
<td>Kilpela, Hill et al. (2014)</td>
<td>285 sorority members</td>
<td>RCT tested if undergraduate facilitators trained by other undergraduates to lead DBI produced weaker effects than facilitators trained by doctoral-level trainer.</td>
<td>Participants in both DBI conditions showed significant reductions in TII, BD, dieting, negative affect, and ED symptoms; effects did not differ between groups. Within group post $d = .15-.72$; 2-mo FU $d = .09-.39$; 8-mo $d = .12-.25$; 14-mo $d = .14-.40$.</td>
</tr>
<tr>
<td>Linville, Cobb, Bluhm, et al. (2015)</td>
<td>66 females w/ body image concerns (BIC)</td>
<td>Randomized effectiveness trial comparing 4 hr. educational video control condition in 2 primary care clinics</td>
<td>DBI participants showed significantly greater reductions in TII, pressure to be thin, dieting, ED symptoms, body dissatisfaction, and negative affect at posttest and 3-moth follow-up. Post $d = .45-.82$; 3-mo $d = .18-.75$.</td>
</tr>
<tr>
<td>Matusek et al. (2004)</td>
<td>84 females w/ body image concerns</td>
<td>RCT comparing a 2 hr. version of DBI to a psychoeducational or waitlist control condition</td>
<td>DBI participants showed greater decreases in TII and ED symptoms, but not BD, than psychoeducational and waitlist control participants. Waitlist 1-mo FU $d = .09-.39$; psychoeducation 1-mo $d = .02-.33$.</td>
</tr>
<tr>
<td>McMillan et al. (2011)</td>
<td>124 female students w/ BIC</td>
<td>RCT comparing a 4 hr. high-dissonance to a low-dissonance version of the DBI and a waitlist control</td>
<td>High-dissonance showed greater reductions in TII, BD, dieting, and ED symptoms than control at post; low-dissonance showed greater reductions in TII, BD, and dieting, but no ED symptoms than control; high-dissonance produced showed greater reductions in ED symptoms than low-dissonance. Post $d = .28-.1.08$; 3-mo $d = .05-.81$.</td>
</tr>
<tr>
<td>Mitchell et al. (2007)</td>
<td>93 female students w/ BIC</td>
<td>RCT comparing a 4.5 hr. version of DBI to a yoga and control condition</td>
<td>DBI showed greater reductions in TII, BD, ED symptoms, alexithymia, and anxiety than yoga and control groups. Control condition post $d = -.01-.94$; yoga condition post $d = -.12-.39$.</td>
</tr>
<tr>
<td>Perez et al. (2010)</td>
<td>182 sorority members</td>
<td>Non-randomized study examining effects of 4-hr peer-led version of DBI on ED risk factors</td>
<td>DBI participants showed reductions in BD, TII, dieting, media use as source of beauty information at post-test, 5-mo and 1-yr follow-up. Post $d = .08-.44$; 5-mo $d = -.09-.18$; 12-mo $d = -.02-.51$.</td>
</tr>
<tr>
<td>Presnell et al. (2008)</td>
<td>133 females w/ BIC</td>
<td>RCT comparing a 4 hr. DBI with an assessment–only control condition</td>
<td>DBI participants showed greater reductions in BD, dieting, and negative affect, but not TII and ED symptoms than controls. Post $d = .07-.20$; 3-mo $d = -.02-.37$.</td>
</tr>
<tr>
<td>Ramirez et al. (2012)</td>
<td>209 dating couples</td>
<td>Non-randomized trial of a 2 hr. couples-based DBI compared to an assessment–only control condition</td>
<td>Dissonance group reduced pressures to be thin, thin and athletic ideal internalization, BD, and actual-ideal body discrepancy. Post $d = .03-.59$; 1-mo $d = .04-.17$.</td>
</tr>
<tr>
<td>Study</td>
<td>Sample</td>
<td>Intervention</td>
<td>Findings &amp; Effect Sizes (Cohen's $d$)</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Roehrig, et al. (2006)</td>
<td>78 at-risk college women</td>
<td>RCT comparing the 3 hr. version of the DBI with a dismantled version of the DBI with only pure dissonance induction activities</td>
<td>Participants in both DBI conditions showed significant reductions in TII, BD, dieting, and ED symptoms; effects did not differ between groups. Within group post $d = .38-.79$; 1-mo $d = .49-.79$.</td>
</tr>
<tr>
<td>Rohde et al., (2014)</td>
<td>81 (study 1) and 52 (study 2) middle school girls</td>
<td>Randomized pilot trials testing 4.5 hr version of DBI versus a psychoeducational control condition for middle school girls with BIC</td>
<td>DBI participants showed greater pre-post reductions in pressure to be thin TII, BD, dieting, negative affect, and ED symptoms than educational brochure controls in the two trials, but effects did not persist through 3-mo follow-up in either trial. Study 1 post $d = .17-.89$; study 2 post $d = .24-.87$; Study 2 3-mo $d = -.23-.33$.</td>
</tr>
<tr>
<td>Serdar (2006)</td>
<td>343 college women</td>
<td>RCT comparing 3 hr. version of DBI to 3 hr. internet version of the DBI assessment-only control group</td>
<td>Participants in both DBI variants showed significant reductions in BD than controls, but no differences for TII, dieting, negative affect, and ED symptoms; effects did not differ between the two DBI conditions. Group DBI post $d = .09-.28$; Internet DBI post $d = .00-.17$.</td>
</tr>
<tr>
<td>Stice, Butryn et al. (2013); Stice, Butryn et al. (2015)</td>
<td>408 college women with BIC</td>
<td>RCT comparing enhanced 4 hr version of DBI to educational brochure control condition</td>
<td>DBI participants showed significantly greater decreases in TII, BD, dieting, negative affect, and ED symptoms than controls at posttest, and at 1-, 2-, and 3-year FU. Post $d = .42-.78$; 12-mo $d = .15-.55$.</td>
</tr>
<tr>
<td>Stice, Chase, et al. (2001)</td>
<td>87 college females w/ BIC</td>
<td>RCT comparing 3 hr. DBI to HW</td>
<td>DBI participants showed greater reductions in TII and BS, but not dieting, negative affect, or ED symptoms than controls. Post $d = .03-.62$; 1-mo $d = .05-.91$.</td>
</tr>
<tr>
<td>Stice, Mazotti, et al. (2000)</td>
<td>30 college females w/ BIC</td>
<td>Non-randomized pilot study comparing 3 hr. DBI to delayed-intervention control condition</td>
<td>DBI showed greater pre-post reductions in TII, BS, negative affect, and ES symptoms, but not dieting, than waitlist controls; most effects persisted at 1-month follow-up. Post $d = .22-.50$; 1-mo $d = .14-.67$.</td>
</tr>
<tr>
<td>Stice, Rohde, Durant, &amp; Shaw (2012); Stice, Durant, Rohde, Shaw (2014)</td>
<td>107 college females w/ BIC</td>
<td>RCT comparing a prototype internet (I) version of the DBI with a group (G) intervention, educational video condition, or educational brochure condition</td>
<td>Participants in both the Internet and group DBI interventions showed greater pre-post reductions in TII, BD, dieting, negative affect, and ED symptoms than educational video and educational brochure controls. G DBI vs brochure control post $d = .61-.97$; 12-mo $d = .51-.75$; 24-mo $d = .19-.71$; I DBI vs brochure control post $d = .41-.80$; 12-mo $d = .12-.79$; 24-mo $d = .37-.61$; G DBI vs video control post $d = .35-.54$; 12-mo $d = .25-.62$; 24-mo $d = .27-.58$; I DBI vs video post $d = .14-.53$; 12-mo $d = .12-.47$; 24-mo $d = .20-.60$.</td>
</tr>
<tr>
<td>Stice, Rohde, et al. (2009); Stice, Rohde, et al. (2011)</td>
<td>306 high school females w/ BIC</td>
<td>Randomized effectiveness trial comparing 4 hr. DBI to psychoeducational brochure control condition</td>
<td>DBI participants showed greater pre-post decreases in TII, BD, dieting, and ED symptoms than controls; certain effects remained significant at 1-, 2-, and 3-yr follow-up. Post $d = .21-.54$; 6-mo $d = .02-.33$; 12-mo $d = .04-.32$; 24-mo $d = .15-.33$; 36-mo $d = .00-.32$.</td>
</tr>
<tr>
<td>Stice, Rohde et al. 2016</td>
<td>680 females</td>
<td>Randomized effectiveness trial testing an Internet (I) version of a dissonance intervention to a peer-led (PL) and clinician-led (CL) versions, compared to an educational video control condition</td>
<td>All DBI conditions showed greater reductions in TII, BS, dieting, negative affect, and ED symptoms than educational video control at post with most effects holding for CL and PL at FU. At post, CL and PL showed greater reductions on some variables compared to internet DBI; this held for TII and BD (PL only) at FU. Reference group is first group for effect sizes (i.e., negative effects means first group yielded greater effects). CL vs control post $d = -.38-.70$; 6-mo $d = -.20-.43$; PL vs control post $d = -.35-.71$; 6-mo $d = -.17-.52$; I vs control post $d = -.27-.45$; 6-mo $d = -.07-.26$; CL vs I post $d = -.16-.30$; 6-mo $d = -.20-.43$.</td>
</tr>
<tr>
<td>Study</td>
<td>Sample</td>
<td>Intervention</td>
<td>Findings &amp; Effect Sizes (Cohen's $d$)</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>---------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Stice, Shaw et al., (2006); Stice, Marti, et al. (2008)</td>
<td>481 female adolescents w/ BIC</td>
<td>RCT comparing 3 hr. DBI to HW, expressive writing (EW), or assessment (AO) control group</td>
<td>DII showed greater decreases in TII, BD, negative affect, ED symptoms, and psychosocial impairment and lower risk for ED onset through 3-yr follow up than assessment only control, with some effects also emerging compared to HW and expressive writing participants. DBI vs AO post $d = .28-.84$; 6-mo $d = .23-.53$; 12-mo $d = .16-.34$; 24-mo post $d = .09-.38$; 36-mo $d = .11-.39$; DBI vs EW post $d = .40-.73$; 6-mo $d = .11-.49$; 12-mo $d = .09-.24$; 24-mo $d = .04-.36$; 36-mo $d = .02-.29$; DBI vs HW post $d = .16-.38$; 6-mo $d = .00-.27$; 12-mo $d = .13-.22$; 24-mo $d = .11-.18$; 36-mo $d = .21-.15$.</td>
</tr>
<tr>
<td>Stice, Rohde et al. (2013)</td>
<td>171 in study 1 &amp; 148 in study 2; college females w/ BIC</td>
<td>RCT comparing DBI's led by either peer-led or clinician-led groups to an educational brochure control condition (Study 1), and comparing an immediate peer-led group to a waitlist control condition (Study 2)</td>
<td>Peer and clinician-led DBIs showed significantly greater pre-post reductions in TII, BD, dieting, negative affect, and ED symptoms than the control (Study 1); peer-led DBI showed greater pre-post reductions in all outcomes than waitlist (Study 2). Study 1: CL vs control post $d = .55-1.07$; 12-mo $d = .40-.53$; PL vs control post $d = .25-.87$; 12-mo $d = .21-.41$. Study 2: PL vs control post $d = .38-1.12$.</td>
</tr>
<tr>
<td>Stice, Trost, &amp; Chase (2003)</td>
<td>148 college females w/ BIC</td>
<td>RCT comparing 3 hr. DBI to HW and waitlist control groups</td>
<td>DBI showed greater reductions in TII, negative affect and ED symptoms, but not BD and dieting, than controls; DBI did not show greater reductions in outcomes than HW. DBI vs control post $d = .23-.62$; 1-mo $d = .06-.23$; 3-mo $d = .10-.78$; DBI vs HW post $d = .39-.12$; 1-mo $d = .06-.23$; 3-mo $d = .09-32$; 6-mo $d = .08-16$.</td>
</tr>
<tr>
<td>van Diest &amp; Perez (2013)</td>
<td>177 sorority members</td>
<td>Uncontrolled trial examining the effects of a 4-hr. peer-led DBI</td>
<td>DBI participants showed reductions in TII, self-objectification, BD, and ED symptoms, with most effects persisting through 1-yr. follow-up. Post $d = .21-.67$; 5-mo $d = .08-.19$; 12-mo $d = .09-.91$.</td>
</tr>
<tr>
<td>Wade, George, &amp; Adkinson (2009)</td>
<td>100 college females</td>
<td>Randomized trial comparing DBI to control, attentional control, acceptance and distraction conditions</td>
<td>DBI, acceptance, and distraction conditions were superior to both control groups in increasing weight satisfaction and to the basic control group for appearance satisfaction. DBI versus control post $d = .35-.40$; DBI versus rumination control post $d = .37-.43$; DBI versus acceptance post $d = .24-.01$; DBI versus distraction post $d = .70-.23$.</td>
</tr>
</tbody>
</table>

*Note: Only published or conference presented studies with at least 15 participants per cell included. Effect sizes are between group unless there was no control group, in which case effect sizes are within group and corrected (see below). Per Morris and DeShon (2002) recommendations for computing effect sizes for meta-analyses that combine repeated measure studies, independent-groups, and independent-group repeated measures designs, we computed effect sizes in a raw-score metric (i.e., standard deviations [SD] based on raw scores as opposed to change scores). Effect sizes were adjusted with a bias-correction function (Hedges, 1982) using the appropriate degrees of freedom (Morris & DeShon, 2002). Sampling variances for effect sizes based on repeated measures requires a pretest-posttest correlation that is typically unreported. Thus, we estimated pretest-posttest correlations based on raw data from two studies conducted Stice et al. (Stice, Rohde, Butryn et al., 2015; Stice, Rohde, Shaw et al., 2016) that was used for all repeated measures studies. Studies with a single-group repeated measures design (i.e., all participants received the intervention and no control group) potentially overestimate effect size because the effect does not contain an adjustment for longitudinal change that would have occurred in the control group. We followed Becker (1988) recommendations in which single-group repeated measures effect sizes are adjusted by first conducting a meta-analysis using only control groups, then subtracting the average control change from single-group repeated measures effect sizes. Mo = Month; FU = Follow-up; TII = Thin-ideal internalization; BD/BS = Body (Dis) Satisfaction; ED = Eating Disorder.