Behaviour Research and Therapy 61 (2014) 142–149

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Do participant, facilitator, or group factors moderate effectiveness of the *Body Project*? Implications for dissemination

Meghan L. Butryn [a, *], Paul Rohde [b], C. Nathan Marti [b], Eric Stice [b]

[a] Department of Psychology, Drexel University, 3201 Chestnut St., Philadelphia, 19104, PA, USA
[b] Oregon Research Institute, 1776 Millrace Drive, Eugene, OR, 97403, USA

ARTICLE INFO

Article history:
Received 30 May 2014
Received in revised form
28 July 2014
Accepted 11 August 2014
Available online 20 August 2014

Keywords:
Eating disorders
Prevention
Dissemination

ABSTRACT

The *Body Project* is a dissonance-based selective eating disorder prevention program with a broad evidence-base. The study sought to determine if previous findings regarding participant moderators replicate in an effectiveness trial under more real-world conditions. This study also had the novel aim of examining facilitator characteristics and group-level variables as potential outcome predictors. These aims are critical for understanding when the intervention is most effective and for whom. Participants were 408 young women with body image concerns recruited from seven universities. Change in eating disorder symptoms at 1-year follow-up was the primary outcome. Intervention effects were significant for both participants who had low or high baseline symptom levels, but the effect size was approximately twice as large for participants with high initial symptom levels ($d = 0.58$ vs. 0.24). Intervention effects were not predicted by facilitator factors (education, age, BMI, sex) or by group size or attendance rate. This study demonstrates that participants with either low or high eating disorder symptoms will benefit from the intervention but if resources are limited, targeting those with elevated eating disorder symptoms may be sensible. Results also suggest that a wide variety of facilitators can effectively deliver the *Body Project*, which has encouraging implications for dissemination.

A large proportion (13%) of young women experience DSM-5 eating disorders (Stice, Marti, & Rohde, 2013). Eating disorders are marked by chronicity, relapse and suffering in many domains, including emotional distress, functional impairment, psychiatric comorbidity, and early mortality (e.g., Arcelus, Mitchell, Wales, & Nielsen, 2011; Ben-Tovim et al., 2001; Fairburn, 2008; Swanson, Crow, Le Grange, Swendsen, & Merikangas, 2011). Prevention is critical not only because of the inherent importance of preventing this impairment and suffering, but also because treatment has both limited reach and efficacy (Bulik, 2013).

The *Body Project*, a selective prevention program targeting women who report body image concerns, is one of only two prevention interventions shown in efficacy trials to reduce risk for future onset of eating disorders over 2- and 3-year follow-up (the other being the *Healthy Weight Intervention,* a brief selective prevention program that has significantly reduced eating disorder onset through promoting small but sustainable improvements to dietary intake and physical activity; Stice, Marti, Spoor, Presnell, & Shaw, 2008; Stice, Rohde, Shaw, & Marti, 2012). *The Body Project* uses cognitive dissonance strategies to reduce internalization of the thin-ideal through a series of written, behavioral, and verbal exercises during and between the four weekly group sessions. The *Body Project* has extensive evidence supporting its efficacy, both compared to control participants receiving no intervention and those in a time-matched alternate intervention (*Healthy Weight*), in terms of significantly greater reductions in eating disorder risk factors and symptoms, with some effects persisting 3 years (Stice, Marti, Spoor, et al., 2008; Stice, Rohde, Durant, & Shaw, 2012; Stice, Shaw, Burton, & Wade, 2006). Effects have been independently replicated (Becker, Smith, & Ciao, 2005; Halliwell & Diedrichs, 2014; Matusek, Wendt, & Wiseman, 2004; Mitchell, Mazzeo, Rausch, & Cooke, 2007), which increases confidence in the findings.

Two large effectiveness trials of the *Body Project* have been conducted (Stice, Butryn, Rohde, Shaw, & Marti, 2013; Stice, Rohde, Gau, & Shaw, 2009; Stice, Rohde, Shaw, & Gau, 2011), in which the program was delivered using endogenous providers under ecologically valid conditions. In the first effectiveness study, clinicians in high school settings were responsible for recruitment and intervention delivery. Participants randomly assigned to the *Body Project* demonstrated greater reductions than educational brochure

* Corresponding author. Tel.: +1 215 553 7108.
E-mail address: mlb34@drexel.edu (M.L. Butryn).

http://dx.doi.org/10.1016/j.brat.2014.08.004
0005-7967/

control participants in eating disorder symptoms through 3-year follow-up (Stice et al., 2009, 2011), although differences in eating disorder onset were nonsignificant, perhaps because risk of onset was relatively lower than in the efficacy trial. The smaller effect size observed in the high school effectiveness trial, versus the efficacy trial, raises the question of how to maximize the impact of the *Body Project* when it is delivered in community settings. In the second effectiveness trial, clinicians at colleges delivered the intervention. Significant differences were observed at 1-year follow-up between intervention and brochure control participants in both risk factors and eating disorders symptoms (Stice, Butryn, et al., 2013; 2- and 3-year follow-up data collection is ongoing). Effects were 83% larger than those observed in the high school effectiveness trial. Possible explanations for this include use of an enhanced-dissonance version of the intervention script; improved selection, training, and supervising of clinicians; and the higher level of body dissatisfaction in the university sample, providing more opportunity for reductions in outcomes.

The present report examined potential factors that moderate the effects of the *Body Project* using data from this second effectiveness trial. Three categories of moderators were examined: participant-, facilitator-, and group-level factors. Examining moderators of intervention effects is critical for understanding under what circumstances the intervention is most effective, and for which individuals (Kraemer, Wilson, Fairburn, & Agras, 2002), which can guide refinement of inclusion and exclusion criteria to maximize the yield of prevention efforts, as well as to inform the design of alternative interventions those who do not benefit from the original intervention.

The first aim of this study was to determine if participant characteristics moderated the effectiveness of the intervention. We hypothesized that intervention effects would be stronger for participants with the highest levels of thin-ideal internalization, body dissatisfaction, and eating disorder symptoms, because these factors could provide greater motivation for change and facilitate the learning of intervention skills as they can be applied to address current risk factors. We also hypothesized that *Body Project* effects would be weaker for those with higher negative affect, as they had lower motivation to engage in the program, higher social anxiety, or less optimism for change. Previous research has identified significant participant moderators for eating disorder prevention programs. Moderator analyses of the *Healthy Weight* prevention program found greater effects for participants with elevated eating disorder symptoms at baseline (Stice, Rohde, Shaw, et al., 2012; 2013). A meta-analysis evaluation of *Student Bodies*, an eating disorder prevention program that focuses in improving body image and healthy dietary practices, found that changes in weight and shape concerns were larger in higher-risk groups than in lower-risk groups across multiple trials (Beintner, Jacobi, & Taylor, 2012). Moderator analyses from an efficacy trial of the *Body Project* found that participants with higher baseline levels of body image distress, eating disorder symptoms, and thin-ideal internalization benefited most from the intervention (Stice, Marti, Shaw, et al., 2008). Moderation analyses also have been conducted after combining data from three trials evaluating the *Body Project*. In those analyses, participants who had a DSM-5 eating disorder at baseline showed significantly greater pre-post reductions in eating disorder symptoms compared to those not meeting DSM-5 criteria at baseline ($d = .71$ and .18 respectively; Müller & Stice, 2013). Replication of these effects in an effectiveness study with college students is particularly important because such data would most directly inform dissemination efforts in those settings and because moderation effects are more difficult to detect than main intervention effects and hence more difficult to replicate (Brookes et al., 2004).

Participant demographic factors were also examined as part of Aim 1. Analyses examined participant age, ethnicity, and body mass index (BMI) as potential moderators. It is the responsibility of intervention developers to examine the generalizability of intervention effects, ideally ensuring that programs are effective for a broad range of individuals. Though we had no directional hypothesis regarding age effects, this effectiveness trial contained participants with a broader age range than previous efficacy research. Similarly, we anticipated no intervention effects as a function of race/ethnicity, given that none were detected in prior efficacy research (Rodriguez, Marchand, Ng, & Stice, 2008; Stice, Marti, & Cheng, 2014). Higher BMI was found to predict stronger eating disorder effects for the *Healthy Weight* intervention but not for the *Body Project* in prior efficacy research (Stice, Marti, Shaw, et al., 2008).

The second aim of the study was to determine if the following facilitator variables predicted participant response to the intervention: facilitator education level, facilitator age, facilitator BMI, and the presence of a male facilitator. As a broader range of clinicians deliver interventions in effectiveness compared to efficacy research, determining whether provider characteristics predict stronger versus weaker responses to the intervention is important. Very little research has examined moderators at the facilitator-level for this intervention, or more generally in eating disorder prevention or treatment research. It is important to understand if the effectiveness of intervention delivery depends on facilitator education level; one could hypothesize that facilitators with higher levels of education produce better outcomes because they have a higher level of competence delivering the material. It also is important to understand if the age, BMI, or sex of the facilitator is related to the effectiveness of intervention delivery, as it is unknown whether participants may react to those features of a facilitator in a way that impacts their response to the intervention. Previous research has found that health promotion messages to be more persuasive when they are delivered by individuals who are more similar to the average participant (e.g., Cialdini, 2008).

The third aim examined whether group size or average group attendance rate in that specific group predicted the improvements observed in individual participants in that group. It is unknown if groups that are smaller or larger are generally more or less effective in producing symptoms reduction effects. One could hypothesize that the extra opportunities for active participation (which is critical for cognitive dissonance induction) in a small group could be advantageous. Conversely, one could hypothesize that in a large group greater cognitive dissonance occurs because of the greater accountability of having a larger audience observe each individual speak out against the thin-ideal. The average group attendance rate could also impact effectiveness of the intervention for individual participants. If there is large drop-out in a group, the remaining participants may not benefit as much from discussions regarding costs of pursing the thin ideal. Low attendance rates could also undermine group cohesion. Conversely, one could hypothesize that the voluntary nature of participation would be heightened in groups with poor attendance, which should theoretically maximize cognitive dissonance (e.g., “I must really care about this issue and want to change because I'm continuing to attend this group while others have dropped out”) and subsequently produce greater symptom reductions for individuals in those groups.

This is the first study to examine the degree to which facilitator- and group-level factors predict effects of the *Body Project* prevention program. As dissemination of this intervention becomes more widespread, there is likely to be more potential variability in participant, facilitator, and group-level factors, and thus examining whether these factors influence the effectiveness of the intervention is critical. Of note, this study was powered such that null effects

would be meaningful; we had a power = .80, assuming a 2-tailed $\alpha = .05$, to detect $r = .14$ with the full sample and $r = .19$ for analyses examining only the *Body Project* participants, which are both small effects. This study also is unique in using change over 1-year follow-up rather than changes only present during receipt of the intervention.

Methods

Participants and procedure

Participants were 408 young women with body image concerns recruited from seven universities in Oregon, Texas, and Pennsylvania. Recruitment e-mails and posters invited women with body image concerns to participate in a trial designed to improve body acceptance. Participants provided interview and survey data at pretest, posttest, and 1-year follow-ups. See Stice, Butryn, et al. (2013) for details about the sample demographics, informed consent, random assignment, content of the Body Project intervention and the educational brochure condition, facilitator training and supervision, assessor training, quality assurance, exclusions due to current eating disorder diagnosis, and participant compensation.

Outcome assessment

Eating disorder symptoms

The semi-structured Eating Disorder Diagnostic Interview (EDDI) assessed change in DSM-5 eating disorder symptoms as the outcome. Items assessing all symptoms of anorexia nervosa, bulimia nervosa, and binge eating disorder in the past month were summed to form a composite. Baseline level of eating disorder symptoms, as measured by the EDDI, also was examined as a moderator. The EDDI has demonstrated internal consistency ($\alpha = .92$), inter-rater agreement (ICC $r = .93$), 1-week test-retest reliability (ICC $r = .95$), sensitivity to prevention and treatment interventions, and predictive validity for future onset of depression (Burton & Stice, 2006; Stice et al., 2009). The symptom composite exhibited internal consistency at pretest ($\alpha = .74$), inter-rater agreement for 77 randomly selected participants (ICC = .84), and 1-week test-retest reliability for 75 randomly selected participants (ICC = .95).

Participant-level moderators

Thin-ideal internalization

The Ideal-Body Stereotype Scale-Revised assessed thin-ideal internalization (Stice et al., 2006). Items had a response format ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Based on our prior observation that one item, *Shapely women are more attractive*, did not elicit responses consistent with the other items, that item was dropped and ratings for the remaining five items were averaged. The remaining 5 items were found to exhibit internal consistency ($\alpha = .78$). The scale has shown 2-week test-retest reliability ($r = .80$), predictive validity for bulimic symptom onset, and sensitivity to detecting intervention effects (Stice, Marti, Shaw, et al., 2008; Stice, Marti, Spoor, et al., 2008).

Body dissatisfaction

The Satisfaction and Dissatisfaction with Body Parts Scale (Berscheid, Walster, & Bohrnstedt, 1973) was used to assess body dissatisfaction. Participants were asked to rate their satisfaction with 9 body parts using responses ranging from 1 = *extremely satisfied* to 6 = *extremely dissatisfied*. The scale has shown internal consistency ($\alpha = .94$), 3-week test-retest reliability ($r = .90$), predictive validity for bulimic symptom onset, and sensitivity to intervention effects (Stice, Marti, Shaw, et al., 2008; Stice, Marti, Spoor, et al., 2008); $\alpha = .89$ at pretest.

Negative affect

The Beck Depression Inventory (BDI; Beck, Steer, & Carbin, 1988) assessed negative affect over the previous two weeks. Participants rated each of the 21 items on a 4-point scale ranging from 0 = no symptoms present to 3 = severe symptoms. This measure has demonstrated internal consistency ($\alpha = .73$ to .95), test-retest reliability ($r = .60$ to .90), and convergent validity with both clinician ratings of depressive symptoms ($M\ r = .75$; Beck et al., 1988) and broader self-report measures of negative affect ($r = .56$ to .58; Watson, Clark, & Tellegen, 1988); $\alpha = .92$ at pretest.

Body mass index

As part of the EDDI interview, height and weight data were collected, which was used to compute BMI (kg/m^2). After removal of shoes and coats, height was measured to the nearest mm using stadiometers and weight was assessed to the nearest 0.1 kg using digital scales. Two measures of each were obtained and averaged.

Facilitator-level predictors

Education level, age, and BMI represented the maximum value of the two facilitators for each group; for groups with a single facilitator, that facilitator's characteristics were used. Education level was ordinally coded from 1 to 5 to represent the following education levels: high school, BA/BS, Nursing, MA/MS, and PhD. Age was ordinally coded from 1 to 8 to represent the following age categories: 18–25, 26–30, 31–35, 36–40, 41–45, 46–50, 51–55, and 55 & over. BMI (kg/m^2) was calculated from facilitator self-reported height and weight. Groups were also coded for the presence of male facilitators (i.e., at least one facilitator was male).

Group-level predictors

Group size

Group size was defined as the number of participants assigned to attend each group.

Group attendance rate

Participants were coded as irregular attenders if they attended two or fewer sessions and regular attenders if they attended three or four sessions. The proportion of regular attenders was computed for each group.

Statistical methods

Missing data

Multiple imputation was used to replace missing values following best-practice recommendations (Graham, 2009). Missing data were imputed using the R Amelia package (Honaker, King, & Blackwell, 2011), which uses all available data to impute missing data via a bootstrapping approach. Missing data points were replaced with imputed data in 20 data sets, which were analyzed individually. Model parameters and standard errors across analyses were then combined to generate inferences incorporating within and between model parameter variability (Rubin, 1987). The possibility that effects differed as a function of dropout was assessed by fitting pattern-mixture models in which a dummy variable indicating dropout was included in a model as a main effect and as an interaction with each model parameter (Hedeker & Gibbons, 2006); significant dropout effects indicate a NMAR pattern and, in that event, the dropout parameters would be retained.

Model building

We applied natural log transformations to eating disorder symptoms to reduce the impact of extreme observations. Linear mixed effects models accommodate unevenly spaced longitudinal assessments (Raudenbush & Bryk, 2002) and were fit using the R lme function (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2013). Intercept coefficients exhibited significant variability across sites based on the difference in deviance between a model with and without a random site intercept and the variance term and was therefore included as a level-3 grouping factors in the multilevel structure, in which level-1 units were time points, and level-2 units were participants. Longitudinal outcome change was assessed following recommendations from Singer and Willett (2003) in which an unconditional means model, an unconditional linear growth model, and various unconditional non-linear models were compared using the Akaike Information Criterion (AIC). An elevation change model, in which longitudinal change is dummy coded (pretest = 0, posttest and 1-year follow-up = 1) to contrast follow-up measures with pretest measures, was selected to model change across time. This model was selected based on a prior assessment of longitudinal change in this data set (Stice, Butryn, et al., 2013) in which linear, quadratic, log-linear, elevation change and slope and elevation change models of change were assessed and it was determined that the elevation change model was the best fit to the data following longitudinal model building strategies from Singer and Willett (2003). The elevation change coding scheme is henceforth referred to as time. Each moderator was evaluated in separate models. All models examining participant moderators included pretest measure of eating disorder symptoms, time × condition (*Body Project* condition = 1) × moderator interaction, main effects of the variables in the interaction, and all possible two-way interactions between variables comprising the three-way interaction. Models examining facilitator- and group-level predictors were limited to the *Body Project* condition participants as these variables were not relevant to those in the brochure control condition. All models examining facilitator predictor variables included the pretest measure of eating disorder symptoms, the time × moderator interaction, and the main effects of the variables in the interaction. Effect sizes for model parameters were estimated using a formula for converting *t* to *r* (Lipsey & Wilson, 2001) and effect sizes for contrasts between estimated marginal means that were used to probe interaction effects were computed using an approximation of Cohen's *d*, in which the difference is divided by the baseline standard deviation (Feingold, 2009).

Results

Descriptive information

Participants were 408 female students recruited from 7 universities (*M* age = 21.6, *SD* = 5.6; *M* BMI (kg/m^2) = 24.4, *SD* = 5.0). Additional descriptive statistics for participant variables examined as moderators are shown in Table 1. The mode facilitator education level was a masters level (*M* = 4.24 [*SD* = 0.66]; 4% had BA/BS, 64% had MA/MS, and 32% had PhD). The median maximum facilitator age was 36–40 years-old (*M* = 4.30 [*SD* = 0.1.96]; 4% were 18–25 years of age, 7% were 26–30, 33% were 31–35, 26% were 36–40, 0% were 41–45, 7% were 46–50, 15% were 51–55, and 7% were 55 & over). There were five male facilitators who facilitated eight groups (30% of the 27 groups). Average maximum facilitator self-reported BMI was *M* = 24.08 (*SD* = 4.82). Average group size was *M* = 7.59 (*SD* = 1.22). The average proportion with "regular attendance" (i.e., attended more than 2 of the 4 sessions) was *M* = 0.83 (*SD* = 0.15).

The most frequently reported symptoms were feeling depressed or very guilty after overeating (17%), engaging in exercise to compensate for overconsumption of eating or drinking in the past month (22%), definitely fearing fatness or weight gain for more than half of the past month (26%), feeling fat for more than half of the past month (34%), and definitely using weight or shape as one of the main aspects of self-evaluation (61%). The least frequently reported were taking laxatives or diuretics to control shape or weight in past month (2%), making one's self sick as a means of controlling shape or weigh in past month (3%), binging in the past month (11%), eating alone due to embarrassment (12%), fasting as a means of controlling your shape or weight in past month (13%), eating much more rapidly than normal (15%), eating large amounts of food when not physically hungry (15%), eating until uncomfortably full (16%), feeling upset about not controlling eating (16%).

Table 1
Means and standard deviations of participant baseline variables examined as participant moderators.

Variable	Brochure controls		*Body Project*	
	M	*SD*	*M*	*SD*
Eating disorder symptoms	10.99	9.03	12.70	10.53
Thin-ideal internalization	3.85	0.55	3.90	0.58
Body dissatisfaction	3.31	0.70	3.40	0.73
Negative affect	11.59	8.59	12.53	9.07

Missing data

All participants completed a baseline interview and survey. Baseline data was imputed for individual items, however, all moderators had <1% missing data. The eating disorder symptom composite had 1% missing data at post-test and 4% missing at the 12-month follow-up. There were no significant effects indicating that missingness impacted model parameters, indicating that reported pattern of results was robust to dropout effects and that it was thus not necessary to fit pattern-mixture models.

Participant-level risk factor moderators

Results for Aim 1 are presented in Table 2. There was a significant effect for the condition × time × baseline eating disorder symptoms ($t[391] = -2.08$, $p = .039$, $r = -.10$). To interpret the moderation effect of baseline eating disorder symptoms, we examined the predicted values for change in eating disorder symptoms for *Body Project* and brochure control participants at low and high levels of baseline eating disorder symptoms (defined as 1 *SD* below and above the mean, respectively) at post-intervention time points; results are shown in Fig. 1. A single effect was sufficient for both post-test time points due to the fact that our time model was an elevation change model that tested for a mean difference between pre-test and both post-intervention measures. The post-intervention difference between control and *Body Project* participants was significant among low symptom participants ($t[372] = 3.04$, $p = .003$, $d = 0.24$) as well as high symptom participants ($t[394] = 7.40$, $p < .001$, $d = 0.58$). No other significant moderator effects were observed for participant-level eating disorder risk factors (M *d* for remaining three variables = 0.11). There was no significant effects participant BMI, age, or ethnicity and the average effect for these demographic variables was small (M $d = 0.07$).

Facilitator- and group-level predictors

Results for Aims 2 and 3 are presented in Table 3. None of the facilitator- or group-level effects were statistically significant and the average effect sizes was small (M $d = 0.10$).

Table 2
Model coefficients for participant moderators (Aim 1).

Model	Parameter	*B*	*SE*	*df*	*t*	*p*	*r*
Eating disorder symptoms	*Body Project* × time × Eating disorder symptoms	−0.17	0.08	391	−2.08	.039	−.10
Thin-ideal internalization	*Body Project* × time × thin-ideal internalization	−0.13	0.10	394	−1.19	.234	−.06
Body dissatisfaction	*Body Project* × time × body dissatisfaction	−0.06	0.08	388	−0.71	.476	−.04
Negative affect	*Body Project* × time × negative affect	−0.09	0.08	392	−1.25	.213	−.06
BMI	*Body Project* × time × BMI	−0.01	0.01	396	−0.80	.426	−.04
Age	*Body Project* × time × age	−0.00	0.01	400	−0.36	.716	−.02
White vs. non-white	*Body Project* × time × white	0.10	0.12	392	0.78	.436	.04

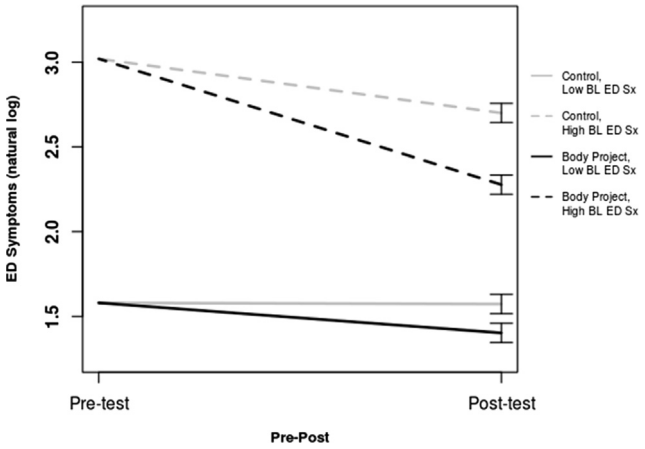

Fig. 1. Fitted values for the condition × time × eating disorder symptoms interaction for control and *Body Project* participants at pre- and post-intervention at high (1 *SD* above the mean) and low (1 *SD* below the mean) baseline eating disorder symptoms (BL ED Sx) with standard error of the difference for post-test comparisons.

Exploratory analysis: Impact of eating disorder symptoms in other group members

To more fully understand the nature of the one significant moderator (which reflected an individual-level variable) related to any group-level effects, we explored whether average eating disorder symptom level or the range of eating disorder symptoms in a group predicted outcome for individual participants within that group. That is to say, do participants who are in a group that has a higher overall level of eating disorder pathology (i.e., higher group mean level of eating disorder symptoms) or in a group that is more homogenous in terms of symptoms (i.e., lower group standard deviation of eating disorder symptoms) tend to show greater individual change compared to participants who are in either low symptomatology or more heterogeneous groups. Investigating whether the individual-level moderator translated into group effects would be important to understand, as the results would have implications for maximizing the yield of prevention efforts given limited resources (e.g., whether participants should be stratified by eating disorder symptom levels to create more homogeneous groups to produce larger effects). The average group log-transformed eating disorder symptoms was $M = 2.36$ ($SD = 0.29$) and the average group log-transformed eating disorder symptoms standard deviation was $M = 0.70$ ($SD = 0.18$). These analyses are included at the end of Table 3. Both effects were significant: time × group mean level of symptoms $t(193) = -2.62$, $p = .010$, $r = -.19$ and time × group standard deviation of symptoms $t(195) = 2.04$, $p = .043$, $r = .14$.

To interpret these effects, we examined the predicted values of symptom change for *Body Project* participants who attended high symptom groups versus low symptom groups (defined as groups that were 1 *SD* above or below the mean of eating disorder symptoms, respectively), and who attended more heterogeneous versus more homogeneous groups (defined as groups where the *SD* of baseline eating disorder symptoms was above or below the mean standard deviation, respectively). We found that, consistent with the finding that high participant baseline eating disorder symptoms predicted greater change in the individual, participants who were from groups where the overall mean symptoms among group members was elevated showed greater individual reductions in eating disorder symptoms over time ($t[193] = 9.34$, $p < .001$, $d = 0.85$) compared to participants from the groups with lower mean symptoms ($t[194] = 5.78$, $p < .001$, $d = 0.51$). Regarding the impact of symptom variability within the group, *Body Project* participants from more homogeneous groups tended to show greater reductions in their symptom levels ($t[194] = 8.98$, $p < .001$, $d = 0.81$) compared to those from more heterogeneous groups ($t[195] = 5.92$, $p < .001$, $d = 0.54$).

Discussion

The purpose of this study was to investigate an array of factors hypothesized to moderate the effects of the *Body Project* at 1-year follow-up. The study sought to determine if previous findings regarding participant moderators would replicate in the real-world context of an effectiveness trial. This study also had the novel aims of examining facilitator characteristics and group-level variables as outcome predictors, which have not been examined previously. These aims have critical relevance for dissemination, because they inform decisions about which individuals are most likely to benefit

Table 3
Model coefficients for facilitator and group-level moderators (Aims 2 and 3).

Model	Parameter	*B*	*SE*	*df*	*t*	*p*	*r*
Facilitator education	Time × facilitator	0.10	0.07	190	1.46	.145	.11
Facilitator age	Time × facilitator age	0.00	0.02	193	0.04	.967	.00
Facilitator BMI	Time × facilitator BMI	0.00	0.01	193	0.48	.635	.03
Male facilitator	Time × male facilitator	0.08	0.11	193	0.71	.476	.05
Group size	Time × group size	−0.02	0.05	192	−0.49	.628	−.04
Proportion high attendance	Time × proportion high attendance	0.15	0.33	194	0.47	.639	.03
Group mean of ED Sx	Time × Group ED Sx mean	−0.41	0.16	193	−2.62	.010	−.19
Group SD of ED Sx	Time × ED Sx SD	0.52	0.26	195	2.04	.043	.14

from this selective prevention program, whether facilitator characteristics may influence the effectiveness of the intervention, and whether characteristics of the group are related to amount of symptom change for individual participants.

Baseline eating disorder symptoms significantly moderated the effect of condition on reductions in eating disorder symptoms. *Body Project* intervention effects were significant for both participants who had low or high baseline symptom levels, but the effect size was approximately twice as large for participants with high baseline symptom levels, reflecting a small magnitude difference. Overall, symptom reductions from baseline to 1-year were greatest among participants who had the highest level of symptoms at baseline. Although it is possible that this single significant moderator was a chance finding (and would not have been significant had we applied a correction for experiment-wide error), as only one out of the 16 moderators was significant, this seems unlikely because this finding replicates results from past independent trials for both the *Body Project* (Müller & Stice, 2013; Stice, Marti, Shaw, et al., 2008; Stice, Marti, Spoor, et al., 2008) and an alternative eating disorder prevention program (*Healthy Weight*; Stice, Rohde, Shaw, et al., 2012; 2013); thus three trials have replicated the same moderator effect, increasing confidence in this finding. It is possible that participants with greater baseline symptoms were more distressed and thus more likely to work for change, or experience greater cognitive dissonance when completing exercises in group or at home (and thus experience more eventual symptom change). It is tempting to argue that the greater symptom reduction was simply a function of higher baseline elevations providing greater potential for change, but this would not logically explain why *Body Project* participants showed greater reductions than control participants if they have higher baseline eating pathology, as the control participants with initially elevated eating pathology would be expected to exhibit the same regression to the mean. The consistently across studies in identifying baseline eating disorder symptoms as a moderator suggests that it is the critical moderator among participant baseline characteristics.

Given the significance of individual baseline eating disorder symptoms as a moderator of change in the individual, we conducted two post-hoc analyses examining whether the treatment effects for the participant were predicted by either the overall level or the variance of eating disorder symptomatology present in the particular *Body Project* group each participant attended. Average symptom level of the group was a significant predictor, such that treatment effects for individual participants were greatest in groups in which average baseline symptom level was higher. The group-level predictor measure of average symptom level for the group is related to but distinct from the participant-level moderator measure of baseline symptoms. An individual participant's symptom level obviously contributes to the group's mean but is not synonymous with the group average, and we thought these two measures were worth examining separately. One possible explanation for this group-level finding is that several participants in those groups were especially motivated to actively participate because of their initial elevated levels of symptomatology and that created a group culture of active engagement that benefited all participants. It is encouraging that groups with the most symptomatic participants showed no signs of iatrogenic effects; conversely, effects were strongest in those groups. Variability in symptom level of the group also was a significant moderator, such that the greatest improvements were seen for individuals who attended groups that had relatively homogenous symptom levels. It is possible that this effect had a clinical origin, in that greater group cohesion and more cognitive dissonance occurred in more homogenous groups. Alternatively, it might have been easier for the facilitator to connect with group participants when they are more similar regarding symptom levels.

Unlike previous research with blended college/high school samples (Müller & Stice, 2013; Stice, Marti, Shaw, et al., 2008; Stice, Marti, Spoor, et al., 2008), the present study, which involved solely college women, found that neither baseline level of body dissatisfaction nor thin-ideal internalization moderated treatment effects, which suggests that a broader range of college women may benefit from the *Body Project*, whereas effects for high school women are potentiated by two risk factors in addition to eating disorder symptoms. Lack of statistical power is not a likely explanation for the null effects, as this study was adequately powered to find clinically meaningful effects.

Intervention effects were not significantly related to facilitator's education level, age, BMI, or sex. Effect sizes for these four possible moderators averaged $r = .05$, which is quite small. Among these non-significant variables, facilitator education level had the largest effect size ($p = .145$, $r = .11$), with a pattern suggesting that participants who were in groups where one or both facilitators had higher education levels (i.e., Ph.D. versus masters) tended to experience greater eating disorder symptom change. However, that pattern of results must be interpreted with great caution given that this predictor was not significant. One explanation for the lack of significant facilitator effects is that this study implemented a strong and fairly intensive training and supervision protocol for all facilitators, which helps to produce robust intervention effects but may have reduced variance in the effectiveness of intervention delivery, making it difficult to detect facilitator-level moderators. The lack of significant facilitator-level predictors has encouraging implications for broad dissemination.

Group size did not significantly predict treatment effects. In efficacy research on the *Body Project*, group size was tightly controlled and typically averaged 6 to 8 participants. In the present effectiveness study, variability was slightly larger but still somewhat limited, with groups ranging from 5 to 9 participants. As the *Body Project* program is disseminated, greater variability in group size may be more likely, and thus this variable may be more likely to be detected as a significant moderator. It is unknown if groups that are smaller or larger than average may be more or less effective in producing symptoms reductions, and hypotheses could support effects in either direction (smaller groups provide more opportunities for individual effort which increases dissonance induction; larger groups create more accountability which increases dissonance induction). At present, the data do not suggest that group size, within the range we examined, influences the impact of the treatment on symptom change.

Group attendance rate (i.e., the absence of drop-out) also did not significantly predict symptom reduction. In efficacy research on the *Body Project*, participant attendance was excellent because researchers conducted careful screening and selection of participants. Intervention attendance rates in this effectiveness study remained very high: 62% attended all 4 sessions and only 5% of participants assigned to the *Body Project* attended none or only one session. As the *Body Project* is disseminated in real-world settings, it is possible that attendance rates for participants may be lower, creating more variability in this group factor. As with group size, hypotheses can offer suggesting stronger effects for either high average attendance (e.g., greater group cohesion; perception of shared value for the intervention) or low average attendance (e.g., greater sense of personal accountability and commitment), though at present, no evidence supports either hypothesis.

This study has several limitations. First, there were some naturally occurring restrictions in values of some of the examined moderators. For example, although there was a meaningful amount of variability in the universities at which participants were

enrolled, including an Ivy League institution, a public urban university, and a small liberal arts college, and the sample of participants was ethnically diverse (42% ethnic minority), parent education level was high (84% of participants had parents with at least some college education). Variability in participants likely could have been greater by, for instance, including a two-year community college or a college with a high enrollment of students from socioeconomically disadvantaged backgrounds. Although the range of scores for some of the moderators (e.g., group size, participant or facilitator age) was not broad, we had adequate variability for the context in which this prevention program was designed and we did not find evidence that the *Body Project* was less effective for subgroups within this sampling frame. Another study limitation is that the research in this area could benefit from longer follow-up periods, as this study reported moderators of symptom change through only 1-year follow-up.

Sufficient evidence of clinically meaningful effects has been previously gathered to warrant dissemination of the *Body Project*. A facilitator guide that provides detailed instruction in intervention delivery has been published (Stice, Rohde, & Shaw, 2013) and a website is available that provides free access to intervention scripts and videos illustrating exemplars of intervention delivery (http://www.bodyprojectsupport.org). The *Body Project* has been implemented in over 100 universities in the US, as well as in 10 other countries. Because this intervention has a large body of research supporting its efficacy for reducing risk of eating disorder onset (Levine & Smolak, 2006), it is encouraging that stakeholders are choosing this as the vehicle for eating disorder prevention efforts. The present study provides valuable information to guide dissemination efforts. Examining moderators of intervention effects is critical for understanding under what circumstances the intervention is most effective, and for which individuals (Kraemer et al., 2002). This study demonstrates that participants with either low or high eating disorder symptoms at baseline will benefit from the intervention but if resources are limited, targeting young women with elevated eating disorder symptoms may be sensible in dissemination efforts and/or in future research. These effects could be clinically driven (e.g., greater motivation to change, sense of group cohesion, or cognitive dissonance induction). Targeting enrollment also may be sensible given that participants in this study benefited most when they were in enrolled in groups in which the average level of baseline symptoms were high. Future research should include measures of moderators such as motivation to change, as well as mediators such as group cohesion and the induction of cognitive dissonance to further understand the potential impact of these effects. There was no evidence that effectiveness of intervention delivery was moderated by facilitator education level, age, BMI, or sex. While the null hypothesis cannot be accepted, the statistical power to detect effects in this study was adequate, lending support to the possibility that with high-quality training and supervision, a wide range of facilitators may be able to effectively deliver the intervention to benefit a wide range of at-risk college women.

Acknowledgment

This study was supported by Grant MH086582 from the National Institute of Mental Health.

References

Arcelus, J., Mitchell, A., Wales, J., & Nielsen, S. (2011). Mortality rates in patients with anorexia nervosa and other eating disorders: a meta-analysis of 36 studies. *Archives of General Psychiatry, 68*, 724–731.

Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77–100.

Becker, C. B., Smith, L. M., & Ciao, A. C. (2005). Reducing eating disorder risk factors in sorority members: a randomized trial. *Behavior Therapy, 36*, 245–253.

Beintner, I., Jacobi, C., & Taylor, C. B. (2012). Effects of an Internet-based prevention programme for eating disorders in the USA and Germany – a meta-analytic review. *European Eating Disorders Review, 20*, 1–8.

Ben-Tovim, D. I., Walker, K., Gilchrist, P., Freeman, R., Kalucy, R., & Esterman, A. (2001). Outcome in patients with eating disorders: a 5-year study. *The Lancet, 357*, 1254–1257.

Berscheid, E., Walster, E., & Bohrnstedt, G. (1973). The happy American body: a survey report. *Psychology Today, 7*, 119–131.

Brookes, S. T., Whitely, E., Egger, M., Smith, G. D., Mulheran, P. A., & Peters, T. J. (2004). Subgroup analyses in randomized trials: risks of subgroup-specific analyses: power and sample size for the interaction test. *Journal of Clinical Epidemiology, 57*, 229–236.

Bulik, C. M. (2013). Are we really paddling as fast as we can? Reflections on why eating disorders treatment and research always seem to be one step behind. *International Journal of Eating Disorders, 46*, 489–491.

Burton, E., & Stice, E. (2006). Evaluation of a healthy-weight treatment program for bulimia nervosa: a preliminary randomized trial. *Behaviour Research and Therapy, 44*, 1727–1738.

Cialdini, R. B. (2008). *Influence: Science and practice* (5th ed.). Boston: Allyn & Bacon.

Fairburn, C. G. (2008). *Cognitive behavior therapy and eating disorders*. Guilford Press.

Feingold, A. (2009). Effect sizes for growth-modeling analysis for controlled clinical trials in the same metric as for classical analysis. *Psychological Methods, 14*, 43–53.

Graham, J. W. (2009). Missing data analysis: making it work in the real world. *Annual Review of Psychology, 60*, 549–576.

Halliwell, E., & Diedrichs, P. C. (2014). Testing a dissonance body image intervention among young girls. *Health Psychology, 33*, 201–204.

Hedeker, D., & Gibbons, R. (2006). *Longitudinal data analysis*. New York, NY: Wiley.

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: a program for missing data. *Journal of Statistical Software, 45*, 1–47. Retrieved from http://www.jstatsoft.org/v45/i07/.

Kraemer, H. C., Wilson, G. T., Fairburn, C. G., & Agras, W. S. (2002). Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry, 59*, 877–883.

Levine, M. P., & Smolak, L. (2006). *The prevention of eating problems and eating disorders: Theory, research, and practice*. Mahway, NJ: Lawrence Erlbaum Associates.

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Matusek, J. A., Wendt, S. J., & Wiseman, C. V. (2004). Dissonance thin-ideal and didactic healthy behavior eating disorder prevention programs: results from a controlled trial. *International Journal of Eating Disorders, 36*, 376–388.

Mitchell, K. S., Mazzeo, S. E., Rausch, S. M., & Cooke, K. L. (2007). Innovative interventions for disordered eating: evaluating dissonance-based and yoga interventions. *International Journal of Eating Disorders, 40*, 120–128.

Müller, S., & Stice, E. (2013). Moderators of the intervention effects for a dissonance-based eating disorder prevention program; results from an amalgam of three randomized trials. *Behaviour Research and Therapy, 51*, 128–133.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., &, R Core Team. (2013). *nlme: Linear and nonlinear mixed effects models (version 3.1-109)*. Retrieved from http://cran.r-project.org/web/packages/nlme/nlme.pdf.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Rodriguez, R., Marchand, E., Ng, J., & Stice, E. (2008). Effects of a cognitive-dissonance-based eating disorder prevention program are similar for Asian American, Hispanic, and White participants. *International Journal of Eating Disorders, 41*, 618–625.

Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.

Stice, E., Butryn, M. L., Rohde, P., Shaw, H., & Marti, C. N. (2013). An effectiveness trial of a new enhanced dissonance eating disorder prevention program among female college students. *Behaviour Research and Therapy, 51*, 862–871.

Stice, E., Marti, C. N., & Cheng, Z. H. (2014). Effectiveness of a dissonance-based eating disorder prevention program for ethnic groups in two randomized controlled trials. *Behaviour Research and Therapy, 55*, 54–64.

Stice, E., Marti, N., & Rohde, P. (2013). Prevalence, incidence, impairment, and course of the proposed DSM-5 eating disorder diagnoses in an 8-year prospective community study of young women. *Journal of Abnormal Psychology, 122*, 445–457.

Stice, E., Marti, N., Shaw, H., & O'Neil, K. (2008). General and program-specific moderators of two eating disorder prevention programs. *International Journal of Eating Disorders, 41*, 611–617.

Stice, E., Marti, C. N., Spoor, S., Presnell, K., & Shaw, H. (2008). Dissonance and healthy weight eating disorder prevention programs: long-term effects from a randomized efficacy trial. *Journal of Consulting and Clinical Psychology, 76*, 329–340.

Stice, E., Rohde, P., Durant, S., & Shaw, H. (2012). A preliminary trial of a prototype internet dissonance-based eating disorder prevention program for young women with body image concerns. *Journal of Consulting and Clinical Psychology, 80*, 907–916.

Stice, E., Rohde, P., Gau, J., & Shaw, H. (2009). An effectiveness trial of a dissonance-based eating disorder prevention program for high-risk adolescent girls. *Journal of Consulting and Clinical Psychology, 77*, 825–834.

Stice, E., Rohde, P., & Shaw, H. (2013). *The body project: A dissonance-based eating disorder prevention intervention*. Oxford University Press.

Stice, E., Rohde, P., Shaw, H., & Gau, J. (2011). An effectiveness trial of a selected dissonance-based eating disorder prevention program for female high school students: long-term effects. *Journal of Consulting and Clinical Psychology, 79*, 500–508.

Stice, E., Rohde, P., Shaw, H., & Marti, C. N. (2012). Efficacy trial of a selective prevention program targeting both eating disorder symptoms and unhealthy weight gain among female college students. *Journal of Consulting and Clinical Psychology, 80*, 164–170.

Stice, E., Rohde, P., Shaw, H., & Marti, N. (2013). Efficacy trial of a selective prevention program targeting both eating disorders and obesity among female college students: 1- and 2-year follow-up effects. *Journal of Consulting and Clinical Psychology, 81*, 183–189.

Stice, E., Shaw, H., Burton, E., & Wade, E. (2006). Dissonance and healthy weight eating disorder prevention programs: a randomized efficacy trial. *Journal of Consulting and Clinical Psychology, 74*, 263–275.

Swanson, S., Crow, S., Le Grange, D., Swendsen, J., & Merikangas, K. (2011). Prevalence and correlates of eating disorders in adolescents: results from the national comorbidity survey replication adolescent supplement. *Archives of General Psychiatry, 68*, 714–723.

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