

An Effectiveness Trial of a Selected Dissonance-Based Eating Disorder Prevention Program for Female High School Students: Long-Term Effects

Eric Stice, Paul Rohde, Heather Shaw, and Jeff Gau
Oregon Research Institute

Abstract

Objective—Efficacy trials found that a dissonance-based eating disorder prevention program in which female high school and college students with body image concerns critique the thin-ideal reduced eating disorder risk factors, eating disorder symptoms, and future eating disorder onset. The present effectiveness trial tested whether this program produces effects through long-term follow-up when high school clinicians recruit students and deliver the intervention under real-world conditions.

Method—Female high school students with body image concerns ($N = 306$; M age = 15.7 $SD = 1.1$) were randomized to the dissonance intervention or an educational brochure control condition and completed assessments through 3-year follow-up.

Results—Dissonance participants showed significantly greater decreases in body dissatisfaction at 2-year follow-up and eating disorder symptoms at 3-year follow-up than controls; effects on other risk factors, risk for eating disorder onset, and other outcomes (e.g., body mass) were marginal or non-significant.

Conclusions—Although it was encouraging that some key effects persisted over long-term follow-up, effects were on average smaller in this effectiveness trial than previous efficacy trials, which could be due to (a) facilitator selection, training, and supervision, (b) the lower risk status of participants, or (c) the use of a control condition that produces some effects.

Keywords

prevention; body dissatisfaction; eating disorder; effectiveness trial; adolescents

Eating disorders, which afflict 10% of adolescent girls and young women, are marked by functional impairment, morbidity, mental health service utilization, and increased risk for future health and mental health problems (Johnson, Cohen, Kasen, & Brook, 2002; Lewinsohn, Streigel-Moore, & Seeley, 2000; Stice, Marti, Shaw, & Jaconis, 2009a; Wilson, Becker, & Heffernan, 2003). Thus, a public health priority is to develop effective eating disorder prevention programs.

Correspondence should be addressed to Eric Stice, Oregon Research Institute, 1715 Franklin Blvd., Eugene, Oregon, 97403. estice@ori.org.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/ccp

Efficacy trials have produced considerable empirical support for a dissonance-based eating disorder prevention program (Stice, Mazotti, Weibel, & Agras, 2000). In this selective prevention program young women at risk for eating disorders because of body image concerns critique the thin ideal espoused for women in verbal, written, and behavioral exercises. These activities theoretically produce cognitive dissonance that motivates participants to reduce their pursuit of the thin-ideal, producing reductions in body dissatisfaction, unhealthy weight control behaviors, negative affect, and eating disorder symptoms. Efficacy trials show that this prevention program produces greater reductions in eating disorder risk factors (e.g., thin-ideal internalization, body dissatisfaction, self-reported dieting, and negative affect), eating disorder symptoms, functional impairment, and future onset of eating disorders over a 3-year follow-up relative to assessment-only or alternative intervention control conditions (e.g., Stice et al., 2000; Stice, Shaw, Burton, & Wade, 2006; Stice, Marti, Spoor, Presnell, & Shaw, 2008). Efficacy trials evaluate whether preventive interventions produce effects under carefully controlled experimental conditions, in which the facilitators are methodically trained and supervised, the intervention is typically delivered in adequately staffed research clinics, and the participants are often homogenous (Flay, 1986). Efficacy trials conducted by independent labs have also found that dissonance-based eating disorder prevention programs produce greater reductions in risk factors and eating disorder symptoms relative to assessment-only control conditions and alternative interventions (e.g., Becker, Smith, & Ciao, 2005; Mitchell, Mazzeo, Rausch, & Cooke, 2007; Roehrig, Thompson, Brannick, & van den Berg, 2006; Wade, George, & Atkinson, 2009). Impressively, dissonance-based prevention programs have even reduced eating disorder risk factors and symptoms when college students deliver the intervention in dissemination research (Becker, Bull, Schaumberg, Cauble, & Franco, 2008; Becker, Smith, & Ciao, 2006; Perez, Becker, & Ramirez, 2010).

Consistent with the intervention theory for the dissonance-based eating disorder prevention program, there is evidence that reductions in thin-ideal internalization mediate the effects of the intervention on change in the outcomes (Seidel, Presnell, & Rosenfield, 2009; Stice, Presnell, Gau, & Shaw, 2007b; Stice, Marti, Rohde, & Shaw, in press). In support of the notion that dissonance induction contributes to the effects of this intervention, participants assigned to high-dissonance versions of this program show significantly greater reductions in eating disorder symptoms than those assigned to low-dissonance versions of this program (Green, Scott, Divankova, Gasser, & Pederson, 2005; McMillan, Stice, & Rohde, 2011), though these trials clearly imply that intervention content and non-specific factors (e.g., perceived group support) also contribute to intervention effects.

Given the empirical support for this eating disorder prevention program from efficacy trials conducted by independent labs, we initiated a large effectiveness trial of this intervention. Effectiveness trials evaluate whether interventions produce effects when delivered by endogenous providers (e.g., school counselors) who are not closely supervised under real world conditions in natural settings with heterogeneous populations (Flay, 1986). It is vital to conduct effectiveness trials because a prevention program that produces effects in highly controlled efficacy trials may be ineffective when delivered under real world conditions by endogenous clinicians. Effectiveness trials can also provide information concerning the degree of training and supervision necessary to achieve intervention effects and have the potential to reveal problems that must be resolved before the prevention program can be successfully disseminated (e.g., effectiveness trials might reveal that typical high schools do not have clinical staff with experience delivering group-based interventions).

The present trial utilized several design elements consistent with effectiveness research (Roy-Byrne et al., 2003). First, we implemented this prevention program in three entire school districts rather than a subset of schools in a district. Second, high school nurses and

counselors, rather than research staff, recruited participants and delivered the intervention within the school environment. Third, we standardized and streamlined facilitator training; endogenous providers completed a 4-hour training session and received limited supervision to mimic real-world conditions. We also used a new 4-session version of the intervention (versus the original 3-session version) that made it easier for endogenous providers to cover the same intervention exercises. Fourth, we minimized exclusion criteria; only adolescent females who met criteria for an eating disorder were excluded. Fifth, we used a minimal intervention educational brochure control condition because this was the only extant resource for students in local high schools, making this an ecologically valid control condition.

An earlier report from the present effectiveness trial found that female high school students who were randomized to the dissonance intervention showed significantly greater decreases in thin-ideal internalization, body dissatisfaction, self-reported dieting, and eating disorder symptoms from pretest to posttest than educational brochure controls, with the effects for body dissatisfaction, dieting, and eating disorder symptoms persisting through 1-year follow-up (Stice, Rohde, Gau, & Shaw, 2009b). This is noteworthy because almost no eating disorder prevention programs have produced intervention effects for eating disorder symptoms through 1-year follow-up (Stice, Shaw, & Marti, 2007c). The first aim of the current report is to test whether intervention effects from this effectiveness trial persist through 2- and 3-year follow-up. We hypothesized that participants randomized to the dissonance intervention would show greater reductions in risk factors (thin-ideal internalization, body dissatisfaction, self-reported dieting, negative affect), eating disorder symptoms, and risk for onset of threshold and subthreshold anorexia nervosa, bulimia nervosa, and binge eating disorder relative to educational brochure controls. The second aim is to investigate the effects of this prevention program on other ecologically meaningful outcomes, including risk for overweight or obesity onset, psychosocial functioning, and mental health care utilization. The third aim is to benchmark the magnitude of the effects from this effectiveness trial against those observed in our large efficacy trial (Stice et al., 2008).

Methods

Participants and Procedure

Participants were 306 adolescent girls (M age = 15.7, SD = 1.1) with a mean body mass index ($BMI = \text{kg}/\text{m}^2$) of 24.9 (SD = 6.1) recruited from a mid-sized city in the Northwest US. We focused on females between the ages of 14 and 19 because the peak period of risk for eating disorder onset occurs during late adolescence and because females are at much higher risk for eating disorders than males (Lewinsohn et al., 2000; Stice et al., 2009a). The sample was 2% Asian/Pacific Islander, 2% African American, 9% Hispanic, 81% European American, and 6% who specified other or mixed racial heritage, which was representative of the county (2% Asian/Pacific Islander, 2% African American, 82% European American, 6% Hispanic). Parental education, a proxy for socioeconomic status, was 17% high school graduate or less, 24% some college, 38% college graduate, and 21% advanced graduate/professional degree, which was somewhat higher than the education of adults in the county (26% high school graduate or less; 36% some college; 16% college graduate; 10% graduate degree).

From April 2005 to November 2007, participants were recruited from high schools using direct mailings and posters inviting females between the ages of 14 and 19 with body image concerns to participate in a study evaluating interventions designed to improve body acceptance. We provided schools with sample recruitment material, but most schools adapted this material or created their own. School staff were responsible for the mailings,

hanging the posters, or other recruitment efforts (e.g., making announcements on the school PA system). Participants had to verbally affirm that they had body image concerns during a phone call conducted by research staff to schedule the baseline assessment. Assessors collected informed written consent from participants (and their parents if they were minors) before data collection began. The sole exclusion criterion was that participants could not meet criteria for DSM-IV anorexia nervosa, bulimia nervosa, or binge eating disorder at pretest. The 2 individuals who met criteria for these disorders were strongly encouraged to seek treatment, provided with referrals, and told that these interventions were not sufficient for them. Figure 1 provides information on participant flow through this trial.

Participants were randomly assigned to the dissonance intervention or an educational brochure control condition via coin toss within schools. The dissonance intervention consisted of 4 weekly 1-hour group sessions with 6–10 participants. Two nurses or counselors typically delivered the intervention at each school. A scripted manual was used for the dissonance intervention (Stice & Presnell, 2007). Facilitator training involved reading the manual to become familiar with the intervention, and attending a 4-hour workshop to learn the conceptual rationale for the intervention and supporting evidence, discuss and role-play key elements from the sessions, and discuss process issues, including confidentiality, making referrals, and achieving good homework compliance and retention. All sessions were audio-recorded and reviewed by the first author for general guidance in how to improve intervention delivery and fidelity to the scripted manual.

Participants provided interview and survey data at pretest, posttest (intervention termination), and at 6-month, 1-, 2- and 3-year follow-ups. Female assessors, who had a B.A., M.A., or Ph.D. in psychology, were blinded to the condition of participants. Assessors attended 24 hours of training, wherein they received instruction in interview skills, reviewed diagnostic criteria for relevant disorders, observed simulated interviews, and role-played interviews. They also attended annual training workshops. They had to demonstrate high inter-rater agreement ($\kappa [k] > .80$) with supervisors using 12 tape-recorded interviews conducted with individuals with and without eating disorders before collecting data. Weekly consensus meetings were held to resolve ambiguous diagnostic issues. Participants were paid \$15 for completing each assessment, but were not paid for attending the intervention. The local Institutional Review Board approved this project. Greater details regarding intervention content, supervision, fidelity, and competence ratings are reported elsewhere (Stice et al., 2009b).

Dissonance intervention—Participants voluntarily engaged in verbal, written, and behavioral exercises in which they critiqued the thin-ideal ideal during sessions and in homework activities. For example, they wrote a counter-attitudinal essay about the costs associated with pursuit of the thin-ideal and engaged in a counter-attitudinal role-play in which they attempted to dissuade facilitators from pursuing the thin-ideal.

Educational Brochure Control Condition—Participants received a 2-page brochure produced by the National Eating Disorders Association in 2002 which describes negative and positive body image, notes that negative body image is associated with increased risk for onset of eating disorders, and offers 10 steps for achieving a positive body image. Control participants were mailed the educational brochures after randomization, which occurred after the baseline assessment. Participants were referred to treatment if they met criteria for threshold or subthreshold anorexia nervosa, bulimia nervosa, or binge eating disorder at any follow-up assessment (as were participants in the dissonance condition).

Measures

Thin-ideal internalization—The Ideal-Body Stereotype Scale-Revised assessed thin-ideal internalization (Stice et al., 2006). Items used a response format ranging from 1 = *strongly disagree* to 5 = *strongly agree*. Items were averaged for this scale and those described below. This scale has shown internal consistency ($\alpha = .91$), 2-week test-retest reliability ($r = .80$), predictive validity for bulimic symptom onset, and sensitivity to detecting intervention effects (Stice et al., 2006; $\alpha = .78$ at T1).

Body dissatisfaction—Items from the Satisfaction and Dissatisfaction with Body Parts Scale (Berscheid, Walster, & Bohrnstedt, 1973) assessed dissatisfaction with nine body parts using a response scale ranging from 1 = *extremely satisfied* to 6 = *extremely dissatisfied*. This scale has shown internal consistency ($\alpha = .94$), 3-week test-retest reliability ($r = .90$), predictive validity for bulimic symptom onset, and sensitivity to detecting intervention effects (Stice et al., 2006; $\alpha = .91$ at T1).

Dieting—The Dutch Restrained Eating Scale (DRES; van Strien, Frijters, van Staveren, Defares, & Deurenberg, 1986) assesses the frequency of various dieting behaviors using a response scale ranging from 1 = *never* to 5 = *always*. The DRES has shown internal consistency ($\alpha = .95$), 2-week test-retest reliability ($r = .82$), convergent validity with self-reported caloric intake, but not objectively measured caloric intake, predictive validity for bulimic symptom onset, and sensitivity to detecting intervention effects (Stice, Cooper, Schoeller, Tappe, & Lowe, 2007a; Stice et al., 2006; van Strien et al., 1986; $\alpha = .92$ at T1).

Depressive symptoms—The Center for Epidemiologic Studies-Depression Scale (CESD; Radloff, 1977) was used to assess depressive symptoms. For each item, participants select among four response options reflecting increasing levels of symptom severity (0 = *never/less than 1 day in past week* to 3 = *most of the time/5–7 days in the past week*). The CESD has shown internal consistency ($\alpha = .74 - .91$), reliability (test-retest $r = .57 - .59$), and convergent validity with clinician ratings of depressive symptoms ($M r = .88$; Andrews, Lewinsohn, Hops, & Roberts, 1993; Roberts, Lewinsohn, & Seeley, 1991; $\alpha = .94$ at T1).

Eating pathology—The Eating Disorder Diagnostic Interview, a semi-structured interview adapted from the Eating Disorder Examination (Fairburn et al., 1995), assessed DSM-IV eating disorder symptoms. Items assessing symptoms in the past month were summed to form an overall eating disorder symptom composite. A log base₁₀ transformation was used to normalize this composite. We also tested whether the intervention reduced risk for onset of threshold or subthreshold anorexia nervosa, bulimia nervosa, and binge eating disorder among those free of these conditions at pretest following the definitions used previously (Stice et al., 2009a; 2009b; Stice, Marti et al., 2008). For a subthreshold anorexia nervosa diagnosis we required participants to have a BMI of between 85% and 90% of that expected for age and gender, report a definite fear of weight gain, and report that weight and shape was definitely an aspect of self-evaluation. For subthreshold bulimia nervosa we required participants to report at least 6 uncontrollable binge eating episodes and 6 compensatory behavior episode over a 3-month period (an average of twice monthly for each), and to report that weight and shape was definitely an aspect of self-evaluation. For subthreshold binge eating disorder we required that participants report at least 12 uncontrollable binge eating episodes/days over a 6-month period, report fewer than 6 compensatory behavior episodes, report marked distress about binge eating, and report that binge eating was characterized by 3 or more of the following: rapid eating, eating until uncomfortably full, eating large amounts when not hungry, eating alone due to embarrassment, feeling disgusted, depressed, or guilty after overeating.

The symptom composite has shown internal consistency ($\alpha = .92$), 1-week test-retest reliability ($r = .90$), sensitivity to detecting effects from eating disorder prevention and treatment interventions, and predictive validity for future onset of depression in past studies of adolescent girls and young women (Burton & Stice, 2006; Stice et al., 2009b). In the current trial the symptom composite showed internal consistency ($\alpha = .84$ at T1), inter-rater agreement (ICC $r = .93$) for 70 randomly selected participants, and 1-week test-retest reliability (ICC $r = .95$) for 72 randomly selected participants. Threshold and subthreshold eating disorder diagnoses have shown 1-week test retest reliability ($\kappa = .96$) and inter-rater agreement ($\kappa = .86$) in past studies (Stice, Marti et al., 2008). Further, individuals who have received a diagnosis of threshold or subthreshold eating disorders versus those who did not report significantly greater mental health treatment, psychiatric distress, and functional impairment (Stice et al., 2009).

Body mass—Age-adjusted BMI percentile was used to reflect height-adjusted weight (Pietrobelli et al., 1998) to accommodate for the fact that many participants increased in height during the 3-year follow-up. After removal of shoes and coats, height was measured to the nearest millimeter using stadiometers and weight was assessed to the nearest 0.1 kg using digital scales. Two measures of each were obtained and averaged in an effort to ensure reliable measurement. BMI correlates with direct measures of body fat such as dual energy x-ray absorptiometry ($r = .80-.90$) and health measures such as blood pressure, adverse lipoprotein profiles, and diabetes mellitus (Pietrobelli et al., 1998). Because of the normative increases in BMI that occur during adolescence as a function of physical maturation, we used the international standards developed by Cole and colleagues (2000) that represent downward, age-adjusted extensions of the 25.0 and 30.0 kg/m² cut-points recommended by the World Health Organization (2000) for overweight and obesity, respectively. These cut points correspond to BMI values that are associated with particularly increased risk for weight-related medical problems such as diabetes mellitus (Cole et al., 2000).

Functional impairment—Items from the Social Adjustment Scale (Weissman & Bothwell, 1976) assessed psychosocial functioning in the family, peer group, school, and work spheres. Response formats ranged from 1 = *never* to 5 = *always*. The adapted version of this scale has shown internal consistency ($\alpha = .77$), 1-week test-retest reliability ($r = .83$) and sensitivity to detecting intervention effects (Stice et al., 2006).

Health care utilization—Two items were used to assess health care utilization during the past six months: the number of times a person saw a doctor because of illness, injury, long term health problems, or regular check-ups; and the number of times a person saw a psychologist, psychiatrist, or counselor/therapist because of mental health problems. Responses to these two items were averaged.

Results

Preliminary Analyses

Participants in the two groups were compared on demographic characteristics and all baseline measures of the outcomes. Groups did not significantly differ on any demographic characteristics or baseline study outcomes with the exception that controls had higher thin-ideal internalization scores ($t[304] = 2.58, p = .010$; control group $M = 3.50$ [SD=0.46], dissonance control $M = 3.35$ [SD=0.59]), but lower body dissatisfaction scores ($t[304] = 3.41, p = .001$; control group $M = 3.15$ [SD=0.77], dissonance control $M = 3.47$ [SD=0.82]). Subsequent models statistically controlled for baseline differences on these measures

between study groups. There was no evidence that participant age or ethnicity moderated intervention effects.

Attrition rates were 10% and 16% at the 2- and 3-year follow-up, respectively. The participants who did not complete the assessments through 3-year follow-up did not differ from the 84% of participants retained in the trial on any demographic or outcome measures at pretest and attrition did not differ across conditions. Nonetheless, we used full information maximum likelihood (ML) estimation to impute missing data because this intent-to-treat approach produces more accurate and efficient parameter estimates than alternative imputation approaches such as last-observation-carried-forward (Schafer & Graham, 2002).

Omnibus repeated measures ANOVA models tested whether there were differential changes in outcomes across the dissonance and control conditions from baseline to 2- and 3-year follow-up (condition was a 2-level between-subjects factor and time was a 3-level within-subject factor). Time x condition interactions indicated there was significantly differential change over time across body dissatisfaction ($F [2/608] = 3.97, p = .019, \text{Cohen's } d = .23$) and eating disorder symptoms ($F [2/608] = 3.96, p = .020, d = .23$), but not dieting ($F [2/608] = 1.22, p = .296, d = .13$), thin ideal internalization ($F [2/608] = 0.93, p = .397, d = .11$), or depression symptoms ($F [2/608] = 1.04, p = .355, d = .11$). Table 1 provides Ms and SDs for these outcomes across time points and conditions.

We next conducted focused follow-up repeated measures ANOVA models that tested whether intervention effects were statistically significant at 2-year and at 3-year follow-up to determine how long effects persisted. In each model, condition was a 2-level between-subjects factor and time was a 2-level within-subjects factor (pretest to 2-year follow-up and pretest to 3-year follow-up). The time x condition interactions test whether participants in one condition showed significantly greater decreases on the outcome than participants in the other condition at each particular follow-up. Time x condition interactions presented in Table 2 indicate that the dissonance intervention produced significantly greater decreases in body dissatisfaction and marginally greater decreases in eating disorder symptoms at the 2-year follow-up. At the 3-year follow-up differences in body dissatisfaction were no longer significant, however, the dissonance condition produced significantly greater decreases in eating disorder symptoms.

To examine the clinical significance of change in the eating disorder symptoms over the 3-year follow-up, we conducted reliable change score analysis using the reliable change index (Jacobson & Truax, 1991). Rates of reliable change on the eating disorder symptom composite score was only marginally differ across conditions ($p = .077, \text{OR} = 1.51, [95\% \text{CI} = 0.96-2.41]$), with 44% of the dissonance condition participants showed clinically significant change at the 1-year follow-up compared to 35% of the control participants.

We also examined the effects of the program on other ecologically meaningful outcomes at the 3-year follow-up: psychosocial functioning, BMI, and health care utilization (see Table 1 for Ms and SDs across conditions). Omnibus repeated measures ANOVA models (described above) showed marginally greater improvements from pretest to 2- and 3-year follow-ups for psychosocial functioning ($F [2/608] = 2.46, p = .086, d = .18$), but non-significant differential change in BMI ($F [2/608] = 0.34, p = .711, d = .06$) and health care utilization ($F [2/608] = 0.50, p = .610, d = .09$).

We next tested whether the intervention reduced risk for future onset of any threshold or subthreshold eating disorder among participants initially free of these conditions with one-tailed Cox proportional hazard models. We used directional tests because the dichotomous outcomes models are less sensitive than the continuous outcome models. By 3-year follow-up, 9 participants in the educational brochure control condition (5.5%) and 5 participants in

the dissonance condition (4.0%) showed onset of a threshold or subthreshold eating disorder. Although this difference corresponds to an odds ratio of 1.4, which reflects a 27% reduction in risk for onset of eating disorders and a small effect size ($d=.27$), this difference did not reach significance (estimate=-0.33, SE=0.56, Wald $\chi^2[1] = .35, p=.28$).

When the analysis is restricted to future onset of any threshold eating disorder by the 3-year follow-up 5 participants in the educational brochure control condition (3.0%) and 1 participant in the dissonance condition (0.8%) showed onset of a threshold eating disorder. The difference did not reach significance (estimate=-1.38, SE=1.09, Wald $\chi^2[1] = 1.59, p=.10$), but does correspond to an odds ratio of 4.0 and reflects a 73% reduction in risk for onset of eating disorders and a large effect size ($d=1.15$).

The dissonance intervention did not significantly reduce onset of overweight or obesity among participants initially free of these conditions (OR = 1.55, $p = .211$, 95% CI = 0.78-3.07). By 3-year follow-up, 22 participants in the educational brochure control condition (21.2%) and 22 participants in the dissonance condition (29.3%) showed onset of overweight or obesity.

Table 3 compares the effects sizes from the present high school effectiveness trial to parallel effects from the efficacy trial conducted with a combination of high school and college students (51% high school, 49% college; Stice et al., 2006). On average effects were somewhat smaller in magnitude in the current effectiveness trial. The 95% confidence intervals shown in Table 3 indicate that the effect for body dissatisfaction in the present trial was significantly smaller than the effect from the efficacy trial at 3-year follow-up. Interesting, the effect for eating disorder symptoms was significantly larger in the present trial than in the earlier efficacy trial; the remaining effects reported in Table 3 did not differ significantly across the two trials

Discussion

This effectiveness trial indicated that high school participants in the dissonance-based eating disorder prevention program, relative to educational brochure controls, showed significantly greater decreases in body dissatisfaction at 2-year follow-up and significantly greater decreases in eating disorder symptoms at 3-year follow-up when school personnel were responsible for recruitment and intervention delivery. Four other effects were marginal, including ones for reduced eating disorder symptoms at 2-year follow-up, greater clinically significant change in eating disorder symptoms, reduced risk for threshold eating disorder onset, and improved psychosocial functioning over 2- and 3-year follow-up. The significant effects were moderate in magnitude and the marginal effects were small to large in magnitude. These findings converge with those from a previous trial (Matussek, Wendt, & Wiseman 2004) that found that an abbreviated version of the dissonance eating disorder prevention program produced greater reductions in eating disorder symptoms in college females with body image concerns relative to assessment-only controls when health educators at colleges delivered the intervention (though research staff were responsible for recruitment in that trial).

It is encouraging that this brief 4-hour dissonance-based eating disorder prevention program produced intervention effects that persisted through 2- and 3-year follow-up, particularly in light of the fact that most eating disorder prevention programs have not been shown to reduce eating disorder symptoms (Stice et al., 2007c). The present study is the first to find evidence that a prevention intervention significantly reduced eating disorder symptoms through 3-year follow-up. The effect for eating disorder symptoms at 3-year follow-up was significantly larger in the present effectiveness trial than in the previous efficacy trial,

though the body dissatisfaction effect at 3-year follow-up was significantly smaller in the present effectiveness trial relative to the efficacy trial. In addition, this is one of the only effectiveness trials to test whether an empirically supported eating disorder prevention program produces intervention effects when high school clinicians deliver the intervention under real world conditions. These findings extend the evidence-based for the dissonance eating disorder prevention program by verifying that this intervention can produce lasting effects when endogenous providers deliver the intervention under real-world conditions. This is an important contribution to the literature because, to our knowledge, this is the first effectiveness trial of an eating disorder prevention program in which school staff were completely responsible for recruiting participants for a selected program and delivering the intervention.

On the other hand, the effects for the current high school effectiveness trial were smaller on average and did not persist as long as was observed in the efficacy trial involving high school and college students (Stice et al., 2006; 2008). Specifically, the intervention effects for thin-ideal internalization and self-reported dieting observed in the report of the effects through 1-year follow-up (Stice et al., 2009) did not persist over longer-term follow-up. In addition, several effects observed in the efficacy trial (Stice et al., 2006; 2008), including reduced risk for onset of eating disorders and overweight/obesity, and reduced psychosocial impairment and health service utilization, did not reach significance in the present effectiveness trial.

There are a number of potential explanations for why the effects are more limited in this effectiveness trial relative to the past efficacy trial. First, school facilitators may have been less effective than the doctoral level psychologist and advanced graduate students who delivered the intervention in the efficacy trial. Although this is not our clinical impression and ratings of sessions suggested that the dissonance intervention was delivered with high fidelity and competence by school clinicians in this effectiveness trial (Stice et al., 2009), it is possible that the research clinicians who conducted the intervention in the efficacy trial were superior on these dimensions (parallel fidelity and competence ratings were not collected in that efficacy trial). This interpretation would suggest that it is important to improve the selection, training, and supervision of endogenous clinicians in future dissemination efforts. For instance, selecting only facilitators that have experience delivering group-based interventions could improve intervention effects. Moreover, experience suggests that providing detailed supervision based on videotapes of sessions seems to result in improved intervention delivery (email supervision in the present trial was only based on audiotape review). It is important to note that the facilitator from our large efficacy trial delivered the dissonance intervention 18 times, whereas the interventionists in the present effectiveness trial delivered the prevention program an average of only 2 times at each school. Thus, the effectiveness trial may have produced somewhat smaller effects because the facilitators did not have sufficient time to develop expertise. It is also possible that the fact that the clinicians in the high school/college efficacy trial were more similar in age to participants than the school facilitators were in relation to high school students in the present effectiveness trial may have contributed to weaker effects in the present trial, given that therapists who are more similar are more persuasive than those who are less similar (Beutler, Jobe, & Elkins, 1974).

A second potential explanation for the weaker effects in the current trial relative to the efficacy trial is that participants in the high school effectiveness trial had mean body dissatisfaction and thin-ideal internalization scores that were .5 SD lower than participants in the efficacy trial involving high school and college students (Stice et al., 2006). Similarly, the risk for onset of eating disorders was only 6% over the 3-year follow-up in the present effectiveness study, relative to 15% in the efficacy trial. Thus, it appears that this high

school sample from the Northwest US was at lower risk for eating pathology than the sample of high school and college students from the Southwest US, which may have contributed to the smaller effects because eating disorder prevention programs tend to show larger effects with higher-risk participants (Stice et al., 2007c). It is possible that regional differences are responsible, such as the fact that people wear less revealing clothes in colder climates, which may result in a reduced focus on appearance.

A third potential explanation for the weaker effects is that the dissonance intervention was compared to an educational brochure control condition in the present effectiveness trial, but to an assessment-only control condition in the efficacy trial. Receiving an educational brochure has been found to result in significantly greater reductions in risk factors and eating disorder symptoms than assignment to an assessment-only control condition (Mutterperl & Sanderson, 2002; Sanderson & Du Vernay, 2008). Thus, it follows that the effects may be smaller in the present trial because it used an educational brochure control condition whereas the previous efficacy trial used an assessment-only control condition. Consistent with this explanation, pre-post reductions in eating disorder symptoms ($M_1 - M_2 / SD_1$) were 23% larger for participants who received the educational brochures ($d = .27$) in the current effectiveness trial than for assessment-only controls ($d = .22$) in the prior efficacy trial. In contrast, the parallel pre-post reductions in eating disorder symptoms in the dissonance intervention conditions were 22% smaller in the current effectiveness trial ($d = .47$) than in the prior efficacy trial ($d = .60$), implying that one of the two explanations discussed previously contributed to the reduced intervention effects.

It was concerning that intervention effects on thin-ideal internalization did not persist through 2- and 3-year follow-up, whereas effects for body dissatisfaction and eating disorder symptoms did, because according to the intervention model, voluntarily critiquing the thin-ideal theoretically resulted in cognitive dissonance that motivates participants to reduce their subscription to this ideal, putatively resulting in a consequent reduction in body dissatisfaction, dieting, and eating disorder symptoms. An earlier report found evidence that change in thin-ideal internalization did appear to mediate the intervention effects on the other outcomes, but that this process primarily occurred during the pre-to-post period, as this is when most of the change in the mediator and outcomes happened (Stice et al., in press). It is possible the reductions in thin-ideal internalization are important initially and may reduce eating disorder symptoms, but that the lasting reductions in body dissatisfaction play a more prominent role in reducing eating disorder symptoms on a long-term basis. Another possible interpretation is that the measure of thin-ideal internalization was less sensitive than the measures of body dissatisfaction and eating disorder symptoms.

Limitations

This trial improved upon many prior prevention trials by using random assignment, blinded diagnostic interviews, a large sample, a 3-year follow-up, and ratings of intervention fidelity and competence. It is also one of the few effectiveness trials of an eating disorder prevention program. However, it is also important to consider the study limitations when interpreting the findings. We relied on self-report data from surveys and interviews, with the exception of the direct measures of height and weight, introducing the possibility that reporter bias might have distorted our estimates of intervention effects. Second, several of the assessors were B.A. level research assistants, which might have affected the validity of eating disorder diagnoses. Third, the sample was relatively homogeneous with regard to ethnicity and socioeconomic status, suggesting that care should be taken in generalizing the results to more diverse populations. However, the dissonance intervention has produced similar effects for Latino, Asian, and European American participants (Rodriguez, Marchand, Ng, & Stice, 2008). Fourth, there was moderate attrition by the 3-year follow-up, which may have biased the tests of intervention effects.

Implications for Prevention and Future Research

These data suggest that high school clinicians can successfully recruit female students at high risk for eating pathology and deliver this brief evidence-based eating disorder prevention program. Future studies should investigate ways to increase the magnitude of the effects of this intervention in real world settings. We recently developed an enhanced dissonance version of this intervention, which increase activities that are theorized to induce dissonance, including increasing the perception that participation in the sessions is completely voluntary, the degree of effort required for the in-session and between-sessions exercises, and the degree of accountability for anti-thin-ideal statements (e.g., by videotaping the sessions and home practice exercises; McMillan et al., 2011). Based on the fact that we are observing larger intervention effects with the new enhanced dissonance version of this prevention program in an ongoing trial conducted with colleges and universities, we recommend that future trials used this new intervention. It will also be important to conduct effectiveness research with other populations, such as middle school students and college students. Finally, future studies should investigate how best to disseminate this prevention program and implement it on a large-scale basis, which will be necessary to affect a reduction in the incidence of eating disorders in at-risk youth. Utilizing a peer-leader dissemination model may be a particularly effective way to extend the reach of this intervention (Becker et al., 2008).

Acknowledgments

This study was supported by a grant (MH70699) and an administrative supplement, care of Amy Goldstein, from the National Institutes of Health.

We thank project research assistants Cara Bohon, Krista Heim, Erica Marchand, and Janet Ng, our undergraduate volunteers, the Eugene, Springfield, and Bethel School Districts, the school staff who recruited participants and facilitated the groups, and the participants who made this study possible.

References

- Andrews JA, Lewinsohn PM, Hops H. Psychometric properties of scales for the measurement of psychosocial variables associated with depression in adolescence. *Psychological Reports*. 1993;73:1019–1046. [PubMed: 8302975]
- Becker CB, Bull S, Schaumberg K, Cauble A, Franco A. Effectiveness of peer-led eating disorders prevention: A replication trial. *Journal of Consulting and Clinical Psychology*. 2008;76:347–354. [PubMed: 18377130]
- Becker CB, Smith LM, Ciao AC. Peer facilitated eating disorders prevention: A randomized effectiveness trial of cognitive dissonance and media advocacy. *Journal of Counseling Psychology*. 2006;53:550–555.
- Becker CB, Smith L, Ciao AC. Reducing eating disorder risk factors in sorority members: A randomized trial. *Behavior Therapy*. 2005;36:245–254.
- Berscheid E, Walster E, Bohrnstedt G. The happy American body: A survey report. *Psychology Today*. 1973;7:119–131.
- Beutler L, Jobe A, Elkins D. Outcomes in group psychotherapy: Using persuasion theory to increase treatment efficacy. *Journal of Consulting and Clinical Psychology*. 1974;42:547–553. [PubMed: 4847261]
- Burton E, Stice E. Evaluation of a healthy-weight treatment program for bulimia nervosa: A preliminary randomized trial. *Behaviour Research & Therapy*. 2006;44:1727–1738.
- Cole TJ. Establishing a standard definition for child overweight and obesity worldwide: International survey. *British Medical Journal*. 2000;320:1240–1240. [PubMed: 10797032]
- Dietz WH, Robinson TN. Use of body mass index (BMI) as a measure of overweight in children and adolescents. *The Journal of Pediatrics*. 1998;132:191–193. [PubMed: 9506622]

- Fairburn CG, Norman PA, Welch SL, O'Connor ME, Doll HA, Peveler RC. A prospective study of outcome in bulimia nervosa and the long-term effects of three psychological treatments. *Archives of General Psychiatry*. 1995;52:304–312. [PubMed: 7702447]
- Flay B. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. *Preventive Medicine*. 1986;15:451–474. [PubMed: 3534875]
- Gillham JE, Reivich KJ, Freres DR, Lascher M, Litzinger S, Shatte A, Seligman MEP. School-based prevention of depression and anxiety symptoms in early adolescence: A pilot of a parent intervention component. *School Psychology Quarterly*. 2006;21:323–248.
- Green M, Scott N, Diyankova I, Gasser C, Pederson E. Eating disorder prevention: An experimental comparison of high level dissonance, low level dissonance, and no-treatment control. *Eating Disorders*. 2005;13:157–169.
- Jacobson N, Truax P. Clinical significance: A statistical approach to defining change in psychotherapy research. *Journal of Consulting and Clinical Psychology*. 1991;59:12–19. [PubMed: 2002127]
- Johnson JG, Cohen P, Kasen S, Brook JS. Eating disorders during adolescence and the risk for physical and mental disorders during early adulthood. *Archives of General Psychiatry*. 2002;59:545–552. [PubMed: 12044197]
- Lewinsohn PM, Striegel-Moore RH, Seeley JR. Epidemiology and natural course of eating disorders in young women from adolescence to young adulthood. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2000;39:1284–92. [PubMed: 11026183]
- Mann T, Nolen-Hoeksema S, Huang K, Burgard D, Wright A, Hanson K. Are two interventions worse than none? Joint primary and secondary prevention of eating disorders in college females. *Health Psychology*. 1997;16:215–225. [PubMed: 9152699]
- Matusek JA, Wendt SJ, Wiseman CV. Dissonance thin-ideal and didactic healthy behavior eating disorder prevention programs: Results from a controlled trial. *International Journal of Eating Disorders*. 2004;36:376–388. [PubMed: 15558649]
- McMillan W, Stice E, Rohde P. High- and low-level dissonance-based eating disorder prevention programs with young women with body image concerns: An experimental trial. *Journal of Consulting and Clinical Psychology*. 2011;79:129–134. [PubMed: 21261438]
- Mitchell KS, Mazzeo SE, Rausch SM, Cooke KL. Innovative interventions for disordered eating: Evaluating dissonance-based and yoga interventions. *International Journal of Eating Disorders*. 2007;40:120–128. [PubMed: 17089413]
- Mutterperl JA, Sanderson CA. Mind over matter: Internalization of the thinness norm as a moderator of responsiveness to norm misperception education in college women. *Health Psychology*. 2002;21:519–523. [PubMed: 12211520]
- Neumark-Sztainer D, Butler R, Palti H. Eating disturbances among adolescent girls: Evaluation of a school-based primary prevention program. *Journal of Nutritional Education*. 1995;27:24–31.
- Newman DL, Moffitt TE, Caspi A, Magdol L, Silva PA, Stanton WR. Psychiatric disorder in a birth cohort of young adults: Prevalence, comorbidity, clinical significance, and new case incidence from ages 11 to 21. *Journal of Consulting and Clinical Psychology*. 1996;64:552–562. [PubMed: 8698949]
- Pietrobelli A, Faith M, Allison D, Gallagher D, Chiumello G, Heymsfield S. Body mass index as a measure of adiposity among children and adolescents: A validation study. *Journal of Pediatrics*. 1998;132:204–210. [PubMed: 9506629]
- Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*. 1977;1:385–401.
- Roberts RE, Lewinsohn PM, Seeley JR. Screening for adolescent depression: A comparison of depression scales. *Journal of the American Academy of Child and Adolescent Psychiatry*. 1991;30:58–66. [PubMed: 2005065]
- Rodriguez R, Marchand E, Ng J, Stice E. Effects of a cognitive-dissonance-based eating disorder prevention program are similar for Asian American, Hispanic, and White participants. *International Journal of Eating Disorders*. 2008;41:618–625. [PubMed: 18528871]
- Rohde P, Clarke GN, Mace DE, Jorgensen JS, Seeley JR. An efficacy/effectiveness study of cognitive-behavioral treatment for adolescents with comorbid major depression and conduct disorder.

- Journal of the American Academy of Child and Adolescent Psychiatry. 2004;43:660–668. [PubMed: 15167082]
- Roy-Byrne PP, Sherbourne CD, Graskie MG, Stein MB, Katon W, Sullivan G, et al. Moving treatment research from clinical trials to the real world. *Psychiatric Services*. 2003;54:327–332.
- Presnell K, Stice E. An experimental test of the effect of weight-loss dieting on bulimic pathology: Tipping the scales in a different direction. *Journal of Abnormal Psychology*. 2003;112:166–170. [PubMed: 12653425]
- Sanderson, CA.; De Vernay, M. Unpublished manuscript. 2008. Can a brochure-based intervention impact disordered eating? A preliminary examination of individual differences in responsiveness to a universal prevention approach.
- Schafer JL, Graham JW. Missing data: Our view of the state of the art. *Psychological Methods*. 2002;7:147–177. [PubMed: 12090408]
- Stewart DA, Carter JC, Drinkwater J, Hainsworth J, Fairburn CG. Modification of eating attitudes and behavior in adolescent girls: A controlled study. *International Journal of Eating Disorders*. 2001;29:107–118. [PubMed: 11429973]
- Stice E, Cooper JA, Schoeller DA, Tappe K, Lowe MR. Are dietary restraint scales valid measures of moderate- to long-term dietary restriction? Objective biological and behavioral data suggest not. *Psychological Assessment*. 2007a;19:449–458.
- Stice E, Marti N, Rohde P, Shaw H. Testing mediators hypothesized to account for the effects of a dissonance eating disorder prevention program over longer-term follow-up. *Journal of Consulting and Clinical Psychology*. (in press).
- Stice E, Marti CN, Shaw H, Jaconis M. An 8-year longitudinal study of the natural history of threshold, subthreshold, and partial eating disorders from a community sample of adolescents. *Journal of Abnormal Psychology*. 2009a;118:587–597.
- Stice E, Marti N, Spoor S, Presnell K, Shaw H. Dissonance and healthy weight eating disorder prevention programs: Long-term effects from a randomized efficacy trial. *Journal of Consulting and Clinical Psychology*. 2008;76:329–340. [PubMed: 18377128]
- Stice E, Mazotti L, Weibel D, Agras WS. Dissonance prevention program decreases thin-ideal internalization, body dissatisfaction, dieting, negative affect, and bulimic symptoms: A preliminary experiment. *International Journal of Eating Disorders*. 2000;27:206–217. [PubMed: 10657894]
- Stice, E.; Presnell, K. *The Body Project: Promoting Body Acceptance and Preventing Eating Disorders, Facilitators Guide*. New York: Oxford University Press; 2007.
- Stice E, Presnell K, Gau J, Shaw H. Testing mediators of intervention effects in randomized controlled trials: An evaluation of two eating disorder prevention programs. *Journal of Consulting and Clinical Psychology*. 2007b;75:20–32.
- Stice E, Rohde P, Gau J, Shaw H. An effectiveness trial of a dissonance-based eating disorder prevention program for high-risk adolescent girls. *Journal of Consulting and Clinical Psychology*. 2009b;77:825–834.
- Stice E, Shaw H, Burton E, Wade E. Dissonance and healthy weight eating disorder prevention programs: A randomized efficacy trial. *Journal of Consulting and Clinical Psychology*. 2006;74:263–275. [PubMed: 16649871]
- Stice E, Shaw H, Marti CN. A meta-analytic review of eating disorder prevention programs: Encouraging Findings. *Annual Review of Clinical Psychology*. 2007c;3:233–257.
- van Strien T, Frijters JE, Van Staveren WA, Defares PB, Deurenberg P. The predictive validity of the Dutch Restrained Eating Scale. *International Journal of Eating Disorders*. 1986;5:747–755.
- Wade T, George WM, Atkinson M. A randomized controlled trial of brief interventions for body dissatisfaction. *Journal of Consulting and Clinical Psychology*. 2009;77:845–854. [PubMed: 19803565]
- World Health Organization. WHO Technical Report Series No. 894. Geneva: Author; 2000. Obesity: Preventing and managing the global epidemic.
- Wilson, GT.; Becker, CB.; Heffernan, K. Eating Disorders. In: Mash, EJ.; Barkley, RA., editors. *Child Psychopathology*. 2. New York: Guilford; 2003. p. 687-715.

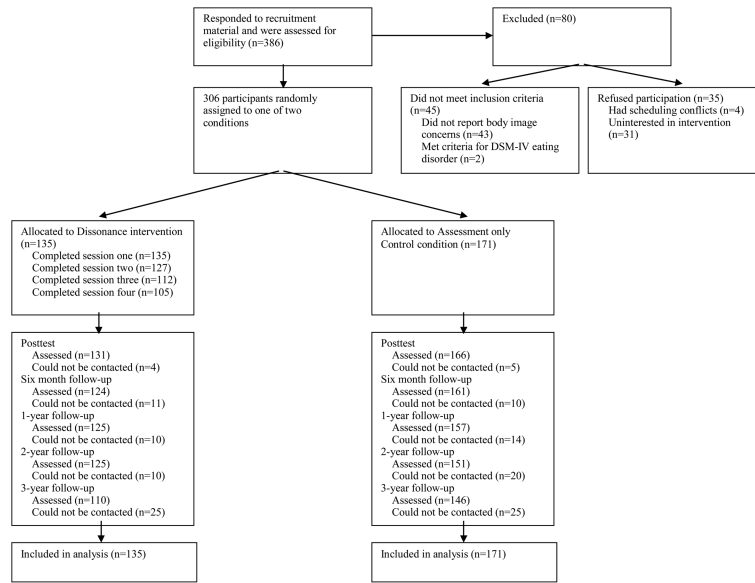


Figure 1.
Participant Flow Throughout the Study

Table 1

Descriptive Statistics

	Pretest		2-year Follow-up		3-year Follow-up	
	M	SD	M	SD	M	SD
Thin ideal						
Dissonance	3.35	0.59	3.21	0.68	3.23	0.62
Control	3.51	0.46	3.30	0.59	3.40	0.52
Body dissatisfaction						
Dissonance	3.47	0.83	3.06	0.85	3.21	0.84
Control	3.15	0.78	3.00	0.84	3.00	0.83
Dieting						
Dissonance	2.35	0.86	1.92	0.82	2.10	0.76
Control	2.30	0.87	2.02	0.82	2.08	0.87
Depression						
Dissonance	18.10	12.61	11.95	11.42	12.06	11.63
Control	16.89	13.08	12.58	12.91	10.88	12.54
<i>I</i> Eating disorder symptoms						
Dissonance	11.57	13.60	5.70	7.54	4.77	5.29
Control	8.94	10.34	5.29	6.06	5.43	6.76
Psychosocial functioning						
Dissonance	2.34	0.44	1.96	0.49	1.88	0.48
Control	2.31	0.45	2.04	0.55	1.84	0.47
Age-adjusted BMI percentile						
Dissonance	79.54	19.86	77.84	22.69	77.22	24.25
Control	75.99	22.54	75.08	22.03	73.01	24.54
Health service utilization						
Dissonance	2.39	2.79	2.10	2.54	1.90	2.17
Control	1.99	2.19	1.92	1.92	1.79	2.16

I Tabled means and standard deviations are non-transformed symptom counts

Table 2
Significance Levels, and Effect Sizes (Cohen's *d*) for the Focused Repeated Measures Time x Condition Interactions

	Pretest to 2-Year Follow-up Time x Condition Interaction		Pretest to 3-Year Follow-up Time x Condition Interaction	
	F-value	p-value	F-value	p-value
Body dissatisfaction	7.48	.001	1.17	.280
Eating disorder symptoms	2.64	.105	6.68	.010
				<i>d</i>
				.15
				.30

Table 3
 Comparison of Effect Sizes (Cohen's d) from the Large Efficacy Trial (Stice et al., 2008) and the Present Effectiveness Trial

	Large Efficacy Trial		Present Effectiveness Trial					
	Pretest to 2-year Follow-up		Pretest to 3-year Follow-up		Pretest to 2-year Follow-up		Pretest to 3-year Follow-up	
	d		d		d	95% CI	d	95% CI
Body dissatisfaction	.28		.43		.39	.23-.54	.15	.01-.30
Eating disorder symptoms	.19		.07		.20	.01-.40	.30	.08-.52