Psychological Bulletin
2004, Vol. 130, No. 2, 206–227

0033-2909/04/$12.00 DOI: 10.1037/0033-2909.130.2.206

Eating Disorder Prevention Programs: A Meta-Analytic Review

Eric Stice and Heather Shaw
University of Texas at Austin

This meta-analysis of eating disorder prevention programs found that intervention effects ranged from an absence of any effects to reductions in current and future eating pathology. Certain effects persisted as long as 2 years and were superior to minimal-intervention control conditions. Larger effects occurred for selected (vs. universal), interactive (vs. didactic), and multisession (vs. single session) programs; for programs offered solely to females and to participants over age 15; for programs without psychoeducational content; and for trials that used validated measures. The results identify promising prevention programs and delineate sample, format, and design features that are associated with larger effects, but they suggest the need for improved methodological rigor and statistical modeling of trials and enhanced theoretical rationale for interventions.

Eating pathology, including threshold and subthreshold anorexia nervosa, bulimia nervosa, and binge eating disorder, is one of the most prevalent psychiatric problems for women and girls and is marked by chronicity and relapse (Fairburn, Cooper, Doll, Norman, & O'Connor, 2000; Lewinsohn, Striegel-Moore, & Seeley, 2000). Eating pathology results in subjective distress and functional impairment, as well as inpatient hospitalization, suicide attempts, and mortality (Newman et al., 1996; Patton, Selzer, Coffey, Carlin, & Wolfe, 1999). Eating pathology also increases the risk for future onset of obesity, depressive disorders, suicide attempts, anxiety disorders, substance abuse, and health problems (Johnson, Cohen, Kasen, & Brook, 2002; Stice, Cameron, Killen, Hayward, & Taylor, 1999; Stice, Hayward, Cameron, Killen, & Taylor, 2000). Because less than a third of individuals with eating disorders receive treatment (Fairburn et al., 2000; Johnson et al., 2002), and treatment produces symptom remission for only 40% to 60% of patients (Agras, Walsh, Fairburn, Wilson, & Kraemer, 2000; Fairburn, Jones, Peveler, Hope, & O'Connor, 1993; Telch, Agras, & Linehan, 2001; Wilfley et al., 2002), much effort has been devoted to developing prevention programs.

The first generation of eating disorder prevention programs delivered didactic psychoeducational material about eating disorders in universal interventions involving all available adolescents. This type of intervention was implicitly based on the assumption that information about the adverse effects of eating disorders would deter individuals from initiating these maladaptive behaviors. The second generation of prevention programs was also universal in focus and didactic in format but included components focusing on resistance of sociocultural pressures for thinness and healthy weight-control behaviors. These interventions were based on the assumption that sociocultural pressures played a key role in the etiology of eating pathology and that adolescents turned to radical dieting and compensatory behaviors primarily for weight-control purposes. The third generation of interventions has included selective programs that target high-risk individuals with interactive exercises that focus on risk factors that have been shown to predict onset of eating pathology (e.g., body dissatisfaction).

Despite the profusion of prevention trials, the results of these evaluations have not been reviewed and analyzed with objective meta-analytic procedures. Thus, the primary aims of this review are to provide a comprehensive summary of these prevention programs and their effects and to examine sample, intervention, and design features that are associated with larger intervention effects. Given the heterogeneity in the effects from these interventions, a systematic consideration of the moderators that are associated with larger intervention effects is important. The secondary aims of this article are to discuss theoretical, methodological, and statistical limitations of the literature and to explore promising directions for future research.

Putative Moderators of Intervention Effects

A unique feature of meta-analysis is that it permits an examination of factors associated with variation in effect sizes. Elucidating the factors that moderate prevention programs effects is informative because it illuminates aspects of the samples, interventions, and designs that are associated with the strongest intervention effects. This information should increase the yield of future prevention efforts by identifying the conditions under which optimal prevention effects occur. Analyses of moderators of intervention effect sizes should also advance general theories regarding effective routes to alter maladaptive health behaviors and attitudes. Accordingly, we investigated several potential moderators of intervention effects that were selected on the basis of theory, prior findings, and clinical experience.

Eric Stice and Heather Shaw, Department of Psychology, University of Texas at Austin.

Preparation of this article was supported by a career award (MH01708) and research grants (MH/DK61957 and MH64560) from the National Institutes of Health.

We thank Michael Levine for sharing the results of his literature search on eating disorder prevention trials and Amy Greenwold, Erika Hinds, and Emily Hoffman for their assistance with the literature search. Thanks also go to the numerous authors who were kind enough to run additional analyses when requested.

Correspondence concerning this article should be addressed to Eric Stice, Department of Psychology, University of Texas, Austin, TX 78712. E-mail: stice@psy.utexas.edu

Risk Status of Participants

It has been suggested that selected programs that are offered to high-risk participants are more effective than universal programs that are offered to all available participants because effect sizes from universal interventions are often larger for high-risk subgroups versus the full sample (Killen et al., 1993). Researchers have also noted that the effects of eating disorder prevention programs (Stice, Mazotti, Weibel, & Agras, 2000) and violence prevention programs (Fields & McNamara, 2003) appear to be larger for selective programs than for universal programs. Theoretically, high-risk individuals that are targeted in selected programs are more motivated to engage in the prevention program content, which may result in greater benefits. It is also likely that low-risk individuals have less room for improvement on the outcomes (a floor effect). Thus, we hypothesized that intervention effects would be larger for selected programs versus universal programs. However, the distinction between universal and selected programs is often blurred. For example, most universal programs focus solely on females—a subpopulation at high risk for eating pathology. We considered interventions delivered to all participants in intact classrooms and trials that did not mention the intervention objective during recruitment (e.g., body acceptance) to be universal programs. We considered interventions that screened participants for a risk factor or that used recruitment strategies that implicitly screened participants, such as advertisements for a body acceptance intervention, to be selected programs. Because the key distinction between selected and universal programs is that the former are offered to high-risk individuals, we use the term *risk status of participants* to refer to this variable in this report.

Participant Sex

Research indicates that girls and women are at much higher risk for eating pathology than boys and men (Lewinsohn, Hops, Roberts, Seeley, & Andrews, 1993; Newman et al., 1996). Thus, girls and women might be expected to be more likely to engage in eating disorder prevention programs than boys and men. In addition, the low base rate of eating pathology in boys and men might produce a floor effect that makes it difficult to observe intervention effects. Therefore, we hypothesized that intervention effects would be stronger for female samples versus male or mixed-sex samples.

Participant Age

Prospective studies suggest that eating pathology is most likely to emerge between the ages of 15 and 19 in adolescent girls (Lewinsohn et al., 2000; Stice, Killen, Hayward, & Taylor, 1998) and that the rates of eating pathology are very low during early adolescence (Keel, Fulkerson, & Leon, 1997; Stice, Presnell, & Bearman, 2001). It has been theorized that prevention programs are most effective when delivered during the developmental period in which the pathological condition emerges (Maggs, Schulenberg, & Hurrelmann, 1996). Consistent with this assertion, our clinical experience suggests that adolescents younger than 15 often have not experienced sufficient subjective distress to motivate them to engage in a program designed to ameliorate and prevent body image and eating disturbances.[1] Younger adolescents may also possess limited insight because their abstract reasoning skills are still developing, which might constrain their ability to benefit from interventions. In addition, there may be a floor effect because the rates of eating pathology are low during early adolescence. Thus, we hypothesized that adolescents over age 15, relative to younger adolescents and children, would be more likely to show a positive response to interventions.

Program Format

A meta-analysis of substance abuse prevention programs found that interactive programs were more effective than didactic programs (Tobler et al., 2000). Theoretically, participants in interactive programs show greater intervention effects because this presentation format helps them become engaged in the program content, which facilitates skill acquisition and attitudinal change. Interactive programs are also more likely to involve exercises that allow participants to apply the skills taught in the intervention, which should facilitate skill acquisition. Accordingly, we predicted that interactive programs would be more effective than didactic programs.

Number of Sessions

Researchers have concluded that brief single-session eating disorder prevention interventions, which are typically 1 hr in length, are insufficient to produce lasting attitudinal and behavioral change because these types of programs rarely produce effects (Martz & Bazzini, 1999). A meta-analysis of substance abuse prevention programs similarly indicated that multisession interventions that were spread over a longer period of time produced superior effects than brief interventions (Rooney & Murray, 1996). Theoretically, multisession interventions, which typically involve at least three hour-long sessions that occur on a weekly basis, allow participants to reflect on intervention material between repeated sessions, thereby maximizing internalization of program tenants. Furthermore, multisession interventions also give participants a chance to try new skills and then return to the group for troubleshooting advice. Thus, we hypothesized that intervention effects would be stronger for multiple-session interventions relative to brief single-session interventions.

Program Content

Another important factor that should determine whether a program produces intervention effects is the content of the intervention. Theoretically, interventions that target established risk factors for eating pathology will be more effective than those that focus on

[1] Note that the predictive power for other cutpoints on participant age, such as younger than 18 years old (which would correspond roughly to entry into college), showed a weaker relation to effect sizes than the younger than age 15 cutpoint suggested by clinical experience. For instance, the effect for the age-15 cutpoint was larger ($r = .47$, $p = .001$) than the effect for the age-18 cutpoint ($r = .34$, $p = .044$) in the prediction of effect sizes for eating pathology at posttest. Post hoc analyses also verified that the continuous age variable did not account for any more variance in effect sizes than did the dichotomous age variable, which suggests that the effects were not attenuated because of our decision to dichotomize this variable.

nonestablished risk factors (Stice, Mazotti, et al., 2000). More generally, research has indicated that psychoeducational content is ineffective in producing behavioral change (Clarke, Hawkins, Murphy, & Sheeber, 1993; Larimer & Cronce, 2002). Accordingly, we hypothesized that programs that focused on increasing resistance to sociocultural pressures for thinness, body satisfaction, self-esteem, and healthy weight-management skills would produce larger intervention effects than programs that did not focus on these factors. Conversely, we expected that programs focusing on stress and coping skills would be associated with smaller intervention effects because these variables are not established eating disorder risk factors. Finally, we hypothesized that psychoeducational programs would be associated with weaker intervention effects. Because of the considerable content overlap of various interventions, it was necessary to create nonorthogonal content variables that reflected whether each particular program focused on each of these six content areas (i.e., sociocultural pressures, healthy weight control, body satisfaction, self-esteem, psychoeducational information on eating disorders, and stress and coping).

Use of Validated Measures

Prevention researchers have noted that use of unreliable measures may result in an underestimation of intervention effects in prevention trials (Kalichman, Carey, & Johnson, 1996). Presumably, trials that use reliable measures should be better positioned to detect intervention effects that do occur because these measures are more sensitive. Thus, we hypothesized that trials that used validated outcome measures would observe larger intervention effects than trials that used measures for which reliability and validity have not been established.

Trials that used measures with established reliability and validity (minimum requirement was evidence of internal consistency greater than .70 and test–retest coefficients greater than .60 or evidence of predictive validity) for at least 50% of the outcomes were coded as having used validated measures and the remaining were coded as having used nonvalidated measures.

Method

Sample of Studies

Several procedures were used to retrieve published and unpublished articles. First, a computer search was performed on PsycINFO, MEDLINE, *Dissertation Abstracts International,* and Cumulative Index to Nursing and Allied Health Literature for the years 1980–2003 (through April 2003) using the following key words: *eating disorder, eating pathology, anorexia, anorexic, bulimia, bulimic, binge eating, prevention, preventive,* and *intervention.* Two research assistants and a professional librarian performed independent searches to increase the odds that all relevant articles would be retrieved. Eric Stice reviewed the products of all three searchers to identify pertinent articles. Second, the tables of contents for journals that commonly publish articles in this area were reviewed for this same period (e.g., *International Journal of Eating Disorders, Journal of Consulting and Clinical Psychology*). Third, the reference sections of all identified articles, review chapters, and books in this area were examined. Finally, established eating disorder prevention researchers were contacted and asked for copies of unpublished articles (under review or in press) describing prevention trials.

Inclusion and Exclusion Criteria

The defining feature of a successful eating disorder prevention program is that it reduces current or future eating disorder symptoms or rates of clinically significant eating pathology relative to the changes in these outcomes observed in a control group. Reductions in initial symptoms are clinically important because there are nontrivial rates of subdiagnostic eating pathology in unselected samples that are associated with subjective distress and functional impairment (Lewinsohn et al., 2000; Patton et al., 1999). The rationale for interventions that focus solely on risk factors is that a reduction in these factors should produce decreases in eating pathology (i.e., intervention effects on eating pathology are mediated by reductions in the risk factors). Thus, to be considered optimally successful, an eating disorder prevention program should reduce both the risk factor and eating pathology—even if the intervention focuses solely on reducing the risk factor. Nonetheless, we included prevention trials that tested for intervention effects on eating pathology, as well as those that solely tested for intervention effects on risk factors that have been found to predict onset of eating pathology.[2]

We focused exclusively on prevention programs that were evaluated in controlled trials. We included trials in which participants were randomly assigned to an intervention or to a minimal-intervention, placebo, waitlist, or assessment-only control condition, as well as trials in which some relevant comparison group was used (e.g., matched controls) in a quasi-experimental design. Random assignment to condition is optimal because it is the best approach to generating comparison groups that are equated on any potential confounding variables. Because many potential confounds are unknown, random assignment is preferable to the use of control groups that are matched to the intervention group on certain dimensions. Nonetheless, carefully selected comparison groups can permit useful inferences regarding intervention effects if analyses test for significant differences in change over time across conditions. Seventeen studies without minimal-intervention, placebo, waitlist, or assessment-only control groups were omitted because there was no way to determine whether any observed changes were due to the intervention or to the passage of time, regression to the mean, or measurement artifacts (i.e., pretest sensitization). Admittedly, the use of waitlist or assessment-only control groups does not permit one to rule out the possibility that any observed effects are due to expectancies or demand characteristics. However, since only five trials used minimal-intervention or placebo control conditions, we did not require these types of control conditions.

We also focused exclusively on studies that tested whether the change in the outcomes over time was significantly greater in the intervention group versus the control group. This could take the form of a Time $\times$ Condition interaction in a repeated measures analysis of variance (ANOVA) model or an analysis of covariance (ANCOVA) model that controls for initial levels of the outcome variable. Growth curve and survival models could also be used if they controlled for initial levels of the outcome. Studies that only tested for significant changes within condition were not included because this type of analysis does not test whether the reductions in the intervention condition are significantly greater than the reductions in the control condition. With this approach there is no way to separate the effects of the intervention versus those from alternative sources, such as regression to the mean. Authors of prevention trials that did not test for significant differences in change in the outcomes across conditions were asked if they could

[2] We included three trials of cognitive–behavioral interventions that were designed to reduce body dissatisfaction (Butters & Cash, 1987; Dworkin & Kerr, 1987; Rosen, Saltzberg, & Srebnik, 1989) even though these trials were not originally conceived of as prevention programs because several subsequent eating disorder prevention programs were based on these programs (Bearman et al., 2003; Celio et al., 2000; Winzelberg et al., 1998, 2000) and because body dissatisfaction is one of the most robust risk factors for onset of eating pathology (Stice, 2002).

provide the results of such analyses. Of 10 authors contacted, 3 provided the requested results so that their findings could be included in this review[3] (the other seven trials were excluded). Four studies that included only posttest measures were excluded because it is not possible to model change in the outcomes with this type of design, and thus there is no way of ruling out the possibility that any observed findings resulted because random assignment failed to create equivalent groups at pretest.[4] Finally, one study that did not collect information to allow the authors to pair pretest and posttest data was excluded because it is not possible to model change with unmatched data.

Effect Size Estimation Procedures

We calculated effect sizes for outcomes assessing eating pathology and risk factors for eating pathology. We included outcomes that were examined in at least 10 trials. We required a minimum of 10 trials because Stice (2002) found it difficult to detect moderators of effect sizes with fewer data points. We included thin-ideal internalization, body dissatisfaction, dieting, negative affect, and body mass because each of these factors has been found to predict subsequent onset of eating pathology in prospective risk factor studies (Stice, 2002). Although deficient knowledge about eating disorders is not an established risk factor for eating pathology, this outcome was included because it was examined in a large portion of the trials.

We calculated separate effect sizes for intervention effects at termination and at follow-up (averaging across all available follow-ups) to provide discrete tests of acute and long-term effects. Reporting one overall index of intervention effects that averaged across termination and follow-up data would have unfairly underestimated effects for programs that were evaluated in trials with longer follow-ups because effects tend to become smaller over time. When multiple measures of the same outcome were included in a trial, effect sizes were calculated for the scale that had the most evidence of reliability and validity. If both measures possessed equal evidence of reliability and validity, an average effect size and probability value (p value) was calculated using the fixed-effects formulas provided in Shadish and Haddock (1994, pp. 265–266). When multiple measures of conceptually similar outcomes were included in a study, the one with the strongest construct validity was selected (e.g., a measure that solely reflected body dissatisfaction was selected over one that reflected body dissatisfaction, thin-ideal internalization, and dieting). Effect sizes were from analyses that used all available data (completer analyses) rather than intent-to-treat analyses (wherein the last available data point for participants who drop out is carried forward) because most trials only reported the former.[5]

If effect sizes were not reported, they were directly calculated by reconstituting the data (e.g., using weighted p values to conduct a chi-square test) or were estimated from the p values (exact p values were directly calculated from test statistics when possible) using the formula provided in Rosenthal (1991). If these options were not possible because insufficient information was provided, effect sizes were requested from the authors. Of the 23 authors contacted, 17 provided the requested effect sizes.[6] Despite these efforts, we were unable to generate effect sizes estimates for the nonsignificant intervention effects from five trials because the authors could not be located or because the authors could not locate their data (we were able to generate effect sizes for the significant effects). As the optimal unbiased estimate of these missing nonsignificant effect sizes is zero (Rosenthal, 1991), we assumed an r of .00 and a p value of .50 for these missing effects. Fortunately, less than 5% of the effect sizes examined in this meta-analysis were set to zero because they were missing.

The correlation coefficient (r) was selected as the index of effect size because of its similar interpretation across different combinations of interval, ordinal, and nominal variables (Pearson's r, Spearman's rho, and point-biserial correlations; Rosenthal, 1991). Cohen's (1988) criteria for small ($r = .10$), medium ($r = .30$) and large ($r = .50$) effects were used.

Results

The literature search identified 51 published and unpublished studies that met the inclusion criteria, in which 38 eating disorder prevention programs were evaluated in 53 separate controlled trials (11 programs were evaluated in more than one trial, five trials evaluated two interventions simultaneously, and two reports described the results from two separate trials). In total, this resulted in 60 separate effect sizes estimates for eating disorder prevention programs for this meta-analysis.

A brief description of the sample, program content, and intervention effects is provided in Table 1 for universal prevention programs and in Table 2 for selected prevention programs. Tables 3 and 4 report the magnitude of effect sizes (r) and provide the sample, design, and content features that were investigated as potential moderators of intervention effects for universal and selected prevention programs, respectively. Only the effect sizes for outcomes examined in 10 or more trials are summarized in Tables 3 and 4, but relevant outcomes that were examined in fewer trials

[3] Three authors who did not report the results of analyses testing whether there was significantly differential change across conditions in the original article were generous enough to provide this information when contacted (Franko, 1998; Kater, Rohwer, & Londre, 2002; Martz & Bazzini, 1999).

[4] For similar reasons, we did not use effect size estimates, such as d, that focus on posttest mean differences across conditions without correcting for pretest mean differences. Such effect sizes estimates are not able to rule out the possibility that differences at baseline between the conditions artificially amplified or attenuated effect size estimates. This approach also introduces greater error variance in effect sizes estimates and therefore decreases power in analyses testing heterogeneity of treatment effects and moderators of treatment effects.

[5] It should be noted that intent-to-treat analyses may provide an overly liberal or overly conservative indication of intervention effects depending on when most participants dropped and whether attrition favored the intervention or control group. If most participants dropped before providing data at intervention termination, intent-to-treat analyses may be conservative because the analyses assume that participants remain as symptomatic as they were at pretest. However, if most participants dropped after providing termination data, intent-to-treat analyses may be overly liberal because the analyses assume that any intervention gains are preserved over follow-up (which is typically not the case). To complicate matters further, intent-to-treat analyses provide an overestimate of intervention effects if more participants drop from the control group before providing termination data or if more participants drop from the intervention condition after providing termination data. Intent-to-treat analyses provide an underestimate of intervention effects if more participants drop from the intervention group before providing termination data or if more participants drop from the control condition after providing termination data. Furthermore, intent-to-treat analyses produce greater biases on effect sizes for less efficient (longer) prevention programs than for more efficient (shorter) programs because (all else being equal) there is greater attrition from longer trials.

[6] The authors of the following studies were gracious enough to provide effect sizes estimates when none were reported in the original article and none could be calculated with reported data: Baranowski & Hetherington, 2001; Dalle Grave, De Luca, & Campello, 2001; Mann et al., 1997; McVey & Davis, 2002; McVey, Davis, Tweed, & Shaw, in press; McVey, Lieberman, Voorberg, Wardrope, & Blackmore, 2003; Neumark-Sztainer, Butler, & Palti, 1995; O'Dea & Abraham, 2000; Paxton, 1993; Santonastaso et al., 1999; Stewart et al., 2001; Varnado-Sullivan et al., 2001; Wade, Davidson, & O'Dea, 2003; Winzelberg et al., 1998, 2000; Withers, Twigg, Wertheim, & Paxton, 2002; Zabinski, Calfas, Gehrman, Wilfley, & Sallis, 2001.

Table 1
Descriptions of the Sample, Intervention Content, and Findings From Universal Eating Disorder Prevention Trials

Study	Sample	Intervention	Findings
Buddeberg-Fischer et al. (1998)	314 middle school boys and girls mean age = 16.1	Didactic. Provided information on normative physical development, nutrition, healthy weight-control behaviors, eating disorders, and risk factors for eating disorders.	No significant intervention effects on eating disorder symptoms, general psychiatric symptoms, or physical symptoms at termination relative to an assessment-only control group.
Dalle Grave et al. (2001)	106 middle school girls and boys mean age = 11.6	Interactive. Provided information about eating disorders and risk factors for eating disorders. Attempted to reduce overevaluation of appearance and promote self-acceptance and healthy weight-control behaviors.	Effects for knowledge at posttest and 6-month follow-up relative to an assessment-only control group, but no effects for body dissatisfaction, dieting, negative affect, or eating pathology.
Jerome (1987, 1991)	135 and 109 high school girls and boys mean age = 15.1; 15.7	Psychoeducational video. Described bulimia nervosa and detailed the processes that putatively cause development of this eating disorder.	Effect for knowledge relative to assessment-only controls at posttest in first trial, but no effects for perceived pressure to be thin, body dissatisfaction, dieting, or eating pathology. Effects for knowledge at posttest in second trial, but no effects for body dissatisfaction, dieting, negative affect, or eating pathology.
Kater et al. (2002)	415 grade and middle school girls and boys mean age = 10.0	Psychoeducational. Provided information on determinants of body shape and healthy weight-control behaviors and promoted body acceptance, coping skills, and critical thinking about mass media.	No effects for knowledge, healthy weight-control behaviors, thin-ideal internalization, body satisfaction, or negative affect at posttest for girls or boys relative to assessment-only controls.
Killen et al. (1993)	838 middle school girls mean age = 12.4	Didactic psychoeducational. Provided information on harmful effects of unhealthy weight-control behaviors, promoted healthy weight-control practices, and taught coping skills to resist sociocultural pressures for thinness.	Effect for knowledge relative to assessment-only controls at posttest, but no effect at 24-month follow-up. No effects for healthy weight-control behaviors, perfectionism, body dissatisfaction, dieting, negative affect, eating pathology, or body mass at posttest or follow-up.
Kusel (1999)	172 middle school girls mean age = 10.1	Media literacy program promoted resistance to sociocultural pressure for thinness by enhancing critical viewing of mass media.	Effects for thin-ideal internalization, body dissatisfaction, dieting, and negative affect relative to assessment-only controls at posttest, but no effects for eating pathology; no effects at 3-month follow-up.
Mann et al. (1997)	113 college women mean age = 17.9	Didactic psychoeducational. Provided information on eating disorders and consequences of eating disorders; included presentations from people in recovery from eating disorders.	No effects for body dissatisfaction, negative affect, or eating pathology at 1-month or 3-month follow-up assessments relative to the assessment-only control group.
Martz & Bazzini (1999, Studies 1 & 2)	114 and 77 college women mean age = 19.0; 19.0	Didactic psychoeducational. Distributed treatment referrals and provided information on eating disorders, putative causes of eating disorders, and healthy weight-control behaviors. Intervention delivered by peer educator in Study 1 and by a professional in Study 2.	No effects for thin-ideal internalization, body dissatisfaction, dieting, or body mass in either evaluation at 1-month follow-up relative to assessment-only controls.
McVey & Davis (2002), McVey et al. (in press)	263 and 258 middle school girls mean age = 10.9; 11.2	Interactive program promoted critical use of the media, body acceptance, healthy weight-control behaviors, and stress management skills. Provided information regarding determinants of body mass.	No effects for body dissatisfaction or eating pathology at posttest or follow-up relative to assessment-only controls in first trial. Effects for body dissatisfaction, dieting, and negative affect at posttest and follow-up, and for bulimic symptoms at posttest but not at follow-up, in second trial, but no effects for perfectionism.
Moreno & Thelen (1993)	219 middle school girls mean age = 13.75	Didactic psychoeducational presentation. Provided information on eating disorders, consequences and putative causes of eating pathology, healthy weight-control behaviors, and peer pressure resistance skills. Delivered by a research assistant (Study 1) and a home economics teacher (Study 2).	Effects for knowledge and behavioral intentions to diet at posttest and 1-month follow-up relative to assessment-only controls in both trials.
Mutterperl & Sanderson (2002)	107 college women mean age = 18.1	Brochure to correct misperceptions about normative eating, dieting, and exercise behaviors of college women.	No effects for thin-ideal internalization, eating disorder symptoms, or body mass at posttest or follow-up relative to a control brochure discussing healthy weight-control behaviors.

Table 1 (*continued*)

Study	Sample	Intervention	Findings
Nebel (1995)	203 college women mean age not available	Interactive psychoeducational. Provided information on eating disorders, healthy weight-control behaviors, and stress management; promoted body satisfaction and self-esteem.	Effects for bulimic symptoms at posttest relative to an assessment-only control condition, but no effects for knowledge, body satisfaction, or negative affect.
Neumark-Sztainer et al. (1995)	341 high school girls mean age = 15.3	Didactic psychoeducational. Presented information on healthy weight-control behaviors, body image, eating disorders, putative causes of eating disorders, and social pressure resistance skills.	Effects for knowledge and eating pathology at 1-month follow-up; knowledge, healthy weight-control behaviors, dieting, and binge eating at 6-month follow-up; and binge eating at 24-month follow-up.
Neumark-Sztainer et al. (2000)	208 girl scouts mean age = 10.6	Provided psychoeducational information on normative physical development. Included self-esteem enhancement exercises and interactive activities focused on helping adolescents become critical consumers of thin-ideal media.	Effects for knowledge at posttest but not at 3-month follow-up, and effects for thin-ideal internalization at 3-month follow-up but not at posttest, relative to waitlist controls. No effects for body dissatisfaction, dieting, or eating pathology.
Nicolino et al. (2001)	85 college women mean age = 18.9	Cognitive–behavioral intervention focusing on emotional and behavioral precipitants of body dissatisfaction, behavioral rituals regarding appearance, challenging negative cognitions regarding body dissatisfaction, and relapse prevention.	No effects for thin-ideal internalization, body dissatisfaction, dieting, and negative affect at the 1-month follow-up relative to minimal-intervention controls.
O'Dea & Abraham (2000)	470 middle school girls and boys mean age = 12.9	Interactive intervention that promoted positive self-esteem, coping skills, and social skills.	Effects for thin-ideal internalization and dieting at 12-month follow-up but not posttest; effects for body dissatisfaction at posttest but not 12-month follow-up relative to assessment-only controls. No effects for perfectionism, dieting, negative affect, or eating pathology.
Outwater (1991)	50 middle school girls and boys mean age = 11.5	Didactic psychoeducational. Focused on enhancing body satisfaction and self-esteem.	No effects for body satisfaction or negative affect at posttest or 1-month follow-up relative to an assessment-only control group.
Paxton (1993)	136 high school girls mean age = 14.1	Didactic psychoeducational. Provided information about sociocultural pressures; determinants of body size, nutrition, weight-control methods, and emotional eating. Interactive discussions about these topics in small groups.	No effects for thin-ideal internalization, body dissatisfaction, dieting, negative affect, eating pathology, body mass, or healthy behaviors at 11-month follow-up relative to assessment-only controls.
Richman (1993, 1998)	180 and 463 primary school girls and boys mean age = 10.4; 10.9	Psychoeducational. Presented information on eating disorders and healthy weight-control behaviors, attempted to enhance self-esteem, and encouraged participants to resist the thin ideal.	Effects for knowledge and body satisfaction compared with assessment-only controls at posttest in first trial, but no effects for dieting and eating pathology. Effects for knowledge, body satisfaction, and dieting relative to assessment-only control group at posttest, but no effects for bulimic pathology.
Santonastaso et al. (1999)	265 vocational school girls mean age = 16.1	Provided psychoeducational information about normative physical development and eating disorders. Girls engaged in unstructured discussions about topics such as body image concerns, sociocultural pressures, and coping with stress.	Effects for body dissatisfaction at 12-month follow-up relative to assessment-only controls, but no effects for body mass, negative affect, perfectionism, or eating pathology.
Shepard (2001)	153 high school girls mean age = 14.4	Psychoeducational interactive program provided information on sociocultural pressures, body image disturbances, dieting and healthy weight-control behaviors, and eating disorders.	Effects for knowledge at posttest and follow-up relative to assessment-only controls, but no effects for thin-ideal internalization, body dissatisfaction, or eating pathology.
Smolak et al. (1998a, 1998b)	222 and 266 grade school girls and boys mean age = 10.0; 9.0	Didactic psychoeducational. Provided information on nutrition, healthy weight-control techniques, and diversity of body shapes and promoted critical evaluation of thin-ideal media.	Effects for thin-ideal internalization at posttest relative to assessment-only controls in first trial, but no effects for knowledge, healthy weight-control behaviors, body dissatisfaction, and dieting. No effects for knowledge, healthy weight-control behaviors, thin-ideal internalization, body dissatisfaction, or dieting in second trial.
Stewart et al. (2001)	459 grade school girls mean age = 13.4	Interactive program focused on resisting cultural pressures for thinness, determinants of body weight, body acceptance, effects of cognitions on emotions, nature and consequences of eating disorders, self-esteem enhancement, stress management, and healthy weight-control behaviors.	Effects for knowledge, dieting, and eating pathology at both posttest and 6-month follow-up, and effects for body dissatisfaction at posttest, but not at 6-month follow-up, relative to an assessment-only control group. No effects for negative affect.

(*table continues*)

Table 1 (*continued*)

Study	Sample	Intervention	Findings
Varnado-Sullivan et al. (2001)	287 grade school girls and boys mean age = 12.0	Interactive psychoeducational intervention focused on the causes and consequences of body dissatisfaction, particularly cultural influences, and healthy weight-control behaviors.	Effects for girls at posttest for thin-ideal internalization and eating pathology, but not for dieting and negative affect relative to waitlist controls. Effects for boys at posttest for thin-ideal internalization and negative affect, but not for dieting and eating pathology.
Wade et al. (2003)	86 middle school girls and boys mean age = 13.42	Media literacy program promoted critical evaluation of thin-ideal images. Self-esteem program promoted positive self-esteem, coping skills, and social skills.	No effects on body dissatisfaction, dieting, or negative affect at posttest or at 3-month follow-up for either intervention relative to assessment-only controls.
Weiss (2000)	173 high school girls mean age = 14.3	Interactive program that provided information on eating pathology and risk factors for eating pathology and promoted healthy weight-control behaviors, social pressures resistance, and body satisfaction and self-esteem.	Effects for knowledge at posttest and follow-up, and for body dissatisfaction at follow-up relative to assessment-only controls, but no effects for dieting, negative affect, or bulimic pathology.
Withers et al. (2002)	242 middle school girls mean age = 12.5	Psychoeducational video provided information on causes and consequences of body image and eating problems, body shape determinants, social pressures for thinness, and healthy eating.	Effects for knowledge at posttest and 1-month follow-up relative to an assessment-only control condition, but no effects for body dissatisfaction or dieting.
Wolf-Bloom (1999)	70 high school girls mean age = 12.3	Media literacy intervention. Promoted critical viewing of media images of the thin ideal and the effects on girls and women.	No effects for thin-ideal internalization, body dissatisfaction, negative affect, or eating pathology at posttest relative to assessment-only controls.

are described in Tables 1 and 2. The effect sizes reported in Tables 3 and 4 are for the intervention versus a waitlist or assessment-only control group when available. It would have unfairly underestimated the effects of the interventions that were compared with more rigorous minimal-intervention or placebo control groups, relative to those interventions that were compared with waitlist or measurement-only control groups. Nonetheless, the effects of the intervention relative to a minimal-intervention or placebo control condition are described in Tables 1 and 2 when available. The effect sizes reflect analyses performed on the entire samples used in these studies. The effects for separate analyses with high-risk subgroups are not reported here because few studies conducted such analyses and because trials that did conduct these analyses focused on qualitatively different high-risk subgroups. The findings from these separate high-risk subgroups are discussed below.

Summary of Intervention Effects

Overall, this review suggests that a number of promising eating disorder prevention programs have been evaluated. Thirty-two (53%) of the interventions resulted in significant reductions in at least one established risk factor for eating pathology, such as body dissatisfaction (which represented 26 distinct programs). Even more exciting was the fact that 15 (25%) of the interventions resulted in significant reductions in eating pathology (which represented 10 distinct programs), and we found evidence that certain interventions both reduce extant eating pathology and prevent increases in eating pathology that were observed in control groups. These promising findings represent a relatively new development, given that most of the effective prevention programs have been evaluated in the past few years. Despite the encouraging findings, there was a wide variety of intervention effects, ranging from those that produced no effects on any outcomes to those that produced significant effects for all outcomes. Thus, it is important to examine factors that moderated the effect sizes observed across interventions.

Average Intervention Effect Sizes and Moderators of Effect Sizes

In this section we report the weighted average random effect size for each of the outcomes that were examined in at least 10 trials—the minimum deemed necessary to permit adequate power to detect moderator effects. The random-effects SPSS macro developed by Lipsey and Wilson (2001) was used to estimate the weighted average effect sizes. We then tested whether there was significant heterogeneity in the effect sizes for each outcome with the random-effects Q test using the SPSS macro. In the event of significant heterogeneity, univariate models tested whether the 12 putative moderators were related to observed effect sizes using the random effects SPSS macro for inverse variance weighted regression with maximum likelihood estimation. We first estimated univariate models that examined each moderator individually to gain insight into these relations that were not obscured by multicolinearity. Next, we estimated multivariate models that included the moderators that were significant in the univariate models. We also tested whether there was significant residual heterogeneity in effect sizes after the moderators were entered into the multivariate

models. This process was conducted separately for acute (pretest to termination) and long-term effects (pretest to follow-up).

Preliminary analyses tested whether publication status (published = 1, nonpublished = 0) moderated the magnitude of effect sizes, which would have suggested that this variable should be included in the analyses. Publication status did not show a significant relation to effect sizes for any of the seven outcomes for either posttest or follow-up data. Similar analyses indicated that use of random assignment to condition (randomized = 1, not randomized = 0) did not show a significant relation to any of the effect sizes examined in this report. Accordingly, neither publication status nor use of random assignment was included as a moderator.

The moderators were coded as follows. Risk factor status was coded such that 1 = selective and 0 = universal. Participant sex was coded such that 1 = female sample and 0 = male or mixed-sex sample. Participant age was coded such that 1 = mean age greater than 15 years and 0 = mean age less than 15 years. Program format was coded such that 1 = interactive program and 0 = didactic program. Number of sessions was coded such that 1 = multiple-session intervention and 0 = single-session intervention. Sociocultural resistance skills, body satisfaction enhancement, self-esteem enhancement, healthy weight-control skills, stress and coping skills, and psychoeducational content was coded such that 1 = present and 0 = absent. Validated measures was coded such that 1 = validated and 0 = nonvalidated.

Knowledge of eating disorders. The weighted average effect size for knowledge of eating disorders at termination across the 16 prevention trials was .38 ($p < .001$). There was statistically significant heterogeneity of effect sizes, $\chi^2(15, N = 16) = 190.52$, $p < .001$, indicating that it was appropriate to investigate potential moderators of intervention effects. We were unable to investigate the effects of risk factor status and psychoeducational content, as there was insufficient representation of the levels of these moderators for this particular outcome. We required that there be at least two studies with each of the two levels for each moderator to be included in the analyses. None of the moderators examined were significant predictors of effect size for knowledge of eating disorders at termination.

The weighted average effect size for knowledge of eating disorders at follow-up across the 13 trials was .29 ($p < .001$). There was statistically significant heterogeneity of effect sizes, $\chi^2(12, N = 13) = 97.76$, $p < .001$, indicating that it was appropriate to investigate moderators of intervention effects. Preliminary analysis indicated that length of follow-up period was not significantly related to the effect sizes for knowledge, therefore this variable was not used as a covariate in subsequent models. There was insufficient variability in the following moderators for them to be included in these analyses: risk factor status, sociocultural content, stress/coping content, and psychoeducational content. None of the remaining moderators were significant predictors of effect size for knowledge of eating disorders at follow-up.

Thin-ideal internalization. The weighted average effect size for thin-ideal internalization at termination across the 20 trials was .19 ($p < .001$). There was significant heterogeneity of effect sizes, $\chi^2(19, N = 20) = 60.60$, $p < .001$. There was insufficient variability in self-esteem enhancement content to examine this factor in these moderator analyses. A univariate model indicated that selected programs produced significantly ($z = 2.93$, $p < .005$) larger decreases in thin-ideal internalization (mean $r = .27$, $p < .001$) than universal programs (mean $r = .12$, $p < .01$). Interventions produced significantly ($z = 2.50$, $p < .05$) larger effect sizes for thin-ideal internalization in trials that used validated measures (mean $r = .24$, $p < .001$) than in trials that used nonvalidated measures (mean $r = .06$, *ns*). Studies that focused on stress and coping skills (mean $r = .05$, *ns*) showed significantly smaller effect sizes ($z = -2.20$, $p < .05$) than those that did not (mean $r = .23$, $p < .001$). Intervention effects were significantly larger ($z = 2.04$, $p < .05$) when delivered to female-only samples (mean $r = .23$, $p < .001$) versus samples with male participants (mean $r = .07$, *ns*). Programs focusing on participants over the age of 15 produced significantly ($z = 2.16$, $p < .05$) larger effect sizes (mean $r = .26$, $p < .001$) than those focusing on younger participants (mean $r = .13$, $p < .01$). In the multivariate model with the five variables that showed significant univariate effects, none remained significant. There was no significant residual heterogeneity in effect sizes, $\chi^2(15, N = 16) = 19.64$, *ns*, when these five predictors were entered into the model.

The weighted average effect size for thin-ideal internalization at follow-up across the 17 trials was .15 ($p < .001$). However, because there was no significant heterogeneity of effect sizes for thin-ideal internalization at follow-up, $\chi^2(16, N = 17) = 25.88$, *ns*, we did not test for moderators of effect size.

Body dissatisfaction. The weighted average effect size for body dissatisfaction at termination across 46 trials was .13 ($p < .001$). There was significant heterogeneity of effect sizes, $\chi^2(45, N = 46) = 85.02$, $p < .001$. A univariate model indicated that selected programs produced significantly ($z = 4.71$, $p < .001$) larger decreases in body dissatisfaction (mean $r = .23$, $p < .001$) than did universal programs (mean $r = .08$, $p < .001$). Interactive programs produced significantly ($z = 1.97$, $p < .05$) larger decreases in body dissatisfaction (mean $r = .15$, $p < .001$) than did studies that used didactic programs (mean $r = .08$, $p < .005$). Significantly ($z = 2.03$, $p < .05$) larger effects were observed for trials with a mean participant age of 15 or older (mean $r = .18$, $p < .001$) than were observed for trials with younger participants (mean $r = .10$, $p < .001$). Multisession programs produced significantly ($z = 2.15$, $p < .05$) larger effects (mean $r = .14$, $p < .001$) than did single-session programs (mean $r = -.03$, *ns*). Programs with psychoeducational content (mean $r = .09$, $p < .001$) had significantly smaller effects ($z = -3.52$, $p < .001$) than did programs without this content (mean $r = .21$, $p < .001$). In the multivariate model with the five moderators that showed significant univariate effects, only the effect for selected versus universal programs ($z = 3.23$, $p < .005$) remained significant. There was no significant residual heterogeneity in effect sizes, $\chi^2(40, N = 41) = 41.23$, *ns*, when these five predictors were entered into the model.

The weighted average effect size for body dissatisfaction at follow-up across the 37 trials was .12 ($p < .001$). There was significant heterogeneity of effect sizes, $\chi^2(36, N = 37) = 76.11$, $p < .001$. Because preliminary analysis indicated that length of follow-up period was not significantly related to the effect sizes for body dissatisfaction, the former variable was not used as a covariate. A univariate model indicated that selected programs produced significantly ($z = 6.72$, $p < .001$) larger decreases in body dissatisfaction (mean $r = .25$, $p < .001$) than did universal programs (mean $r = .06$, $p < .001$). Significantly ($z = 2.52$, $p < .05$) larger decreases in body dissatisfaction were observed for interventions

Table 2
Descriptions of the Sample, Intervention Content, and Findings From Selected Eating Disorder Prevention Trials

Study	Sample	Intervention	Findings
Baranowski & Hetherington (2001)	29 grade school girls mean age = 11.5	Provided information on eating disorders, causes and consequences of eating disorders, body dissatisfaction, thin-ideal internalization, and negative affect.	Significant intervention effects for dieting at termination and follow-up relative to minimal-intervention control group, but not for body dissatisfaction, negative affect, or eating pathology.
Bearman et al. (2003)	74 college women mean age = 18.9	Brief four-session version of the cognitive–behavioral intervention for body dissatisfaction developed by Cash and Rosen (see, e.g., Butters & Cash, 1987).	Effects for body dissatisfaction and negative affect at posttest and 3-month follow-up relative to waitlist controls, but no effects for dieting or bulimic symptoms.
Butters & Cash (1987)	31 college women mean age = 21.3	Cognitive–behavioral intervention promoted body satisfaction by challenging negative cognitions regarding appearance and using systematic desensitization to reduce anxiety about body dissatisfaction.	Effects for body dissatisfaction and negative affect at posttest.
Chase (2001)	91 college women mean age = 18	Didactic intervention that provided education about healthy weight-control behaviors, precursors and consequences of eating disorders, the influence of cognitions on feelings and behaviors, and social pressure resistance skills.	Effects for knowledge, body dissatisfaction, and dieting at posttest and 1-month follow-up relative to minimal-intervention controls, but no effects for thin-ideal internalization, negative affect, eating pathology, or body mass.
Dworkin & Kerr (1987)	79 college women mean age not available	Cognitive intervention promoted body satisfaction through cognitive restructuring. A cognitive–behavioral intervention added self-reinforcement for cognitive restructuring and body acceptance role-plays.	Effects for body dissatisfaction and negative affect at posttest for both interventions relative to minimal-intervention controls.
Franko (1998)	19 college women mean age not available	Intervention promoted critical evaluation of cultural pressure for thinness, presented healthy weight-control skills, challenged dysfunctional cognitions about body image, and introduced affect-regulation skills.	No effects for thin-ideal internalization, body dissatisfaction, or bulimic symptoms relative to assessment-only controls at posttest.
Kaminski & McNamara (1996)	29 college women mean age = 18.3	Interactive program provided information about eating disorders, putative risk factors for eating disorders, and healthy and unhealthy weight control; presented cognitive interventions for body image and eating disturbances; and taught communication and affect regulation skills.	Effects for perfectionism, thin-ideal internalization, body dissatisfaction, dieting, and negative affect relative to assessment-only control group at posttest. All effects except perfectionism persisted through 1-month follow-up.
McVey et al. (2003)	214 grade school girls mean age = 12.5	Interactive program promoted critical use of the media, body acceptance, healthy weight-control behaviors, and stress management skills. Provided information regarding determinants of body mass.	Effects for body dissatisfaction, dieting, and bulimic symptoms relative to assessment-only controls at posttest and 3-month follow-up. Prevent natural increases in bulimic symptoms observed in controls.
Presnell & Stice (2003)	81 college women mean age = 20.0	Intervention promoted a low-calorie diet, adapted from Brownell's (1997) obesity treatment program, for nonobese individuals.	Effects for bulimic pathology and body mass, but not in negative affect or dieting, relative to the waitlist control group at posttest.
Rosen et al. (1989)	23 college women mean age = 19.0	Cognitive–behavioral intervention was similar in content to that evaluated by Butters and Cash (1987) but was delivered in a small group format, provided corrective feedback regarding body size misperception, and sought to reduce avoidance of behaviors that elicit body image concerns.	Effects for body size estimation, body dissatisfaction, and behavioral avoidance relative to minimal-intervention control group at posttest and 2-month follow-up.
Stice, Mazotti, et al. (2000)	30 college women mean age = 18.0	A dissonance-based intervention wherein girls with elevated levels of thin-ideal internalization voluntarily critiqued the thin ideal in a series of verbal, written, and behavioral exercises.	Effects for thin-ideal internalization, body dissatisfaction, negative affect, and bulimic symptoms but not dieting at posttest relative to waitlist controls. All effects except negative affect persisted at 1-month follow-up. Prevented increases in bulimic symptoms observed in the waitlist controls over 2-month study.

Table 2 (*continued*)

Study	Sample	Intervention	Findings
Stice & Ragan (2002), Orjada & Stice (2003)	66 and 60 college women mean age = 21.0; 21.3	Psychoeducational eating disorder class presented information on eating disorders and obesity, putative causes of these disorders, and prevention and treatment programs for these disorders.	Effects for thin-ideal internalization, body dissatisfaction, dieting, eating pathology, and body mass at posttest, but not for negative affect or healthy weight-control behaviors relative to assessment-only controls in first trial. Prevented natural increases in weight observed in controls. Effects for body mass, thin-ideal internalization, and eating pathology at posttest, but not for body dissatisfaction, dieting, negative affect, or healthy weight-control behaviors relative to controls in second trial. Prevented weight gain observed in controls.
Stice, Trost, & Chase (2003)	148 and 193 high school girls and college women	In dissonance intervention, girls with elevated levels of thin-ideal internalization voluntarily critiqued the thin ideal in verbal, written, and behavioral exercises.	In first trial, dissonance program produced effects for thin-ideal internalization, body dissatisfaction, negative affect, and bulimic symptoms at posttest, and for negative affect and bulimic symptoms at 6-month follow-up. No effects for dieting. Healthy weight program produced effects for thin-ideal internalization, negative affect, and bulimic symptoms at posttest and 6-month follow-up. No effects for body dissatisfaction or dieting. In second trial, dissonance program produced effects for thin-ideal internalization, body dissatisfaction, dieting, negative affect, and eating pathology, but not healthy behaviors or body mass, relative to waitlist and minimal-intervention conditions at posttest. Effects relative to waitlist condition persisted through 6-month follow-up, with the exception of the negative affect effect, but no effects for dissonance program relative to minimal-intervention controls at 6-month follow-up. Healthy weight program produced effects for healthy behaviors and eating pathology relative to waitlist controls at posttest, and for healthy behaviors, body dissatisfaction, and eating pathology relative to minimal-intervention controls at posttest. Effects for thin-ideal internalization, body dissatisfaction, dieting, and eating pathology at 6-month follow-up relative to waitlist controls. Only effect for body dissatisfaction relative to minimal-intervention controls at 6-month follow-up.
Stice, Presnell, et al. (2003)	mean age = 17.4; 17.0	Healthy weight-management intervention promoted permanent decreases in high-fat food consumption and increases in exercise.	
Winzelberg et al. (1998, 2000), Celio et al. (2000)	57, 60, and 76 college women mean age = 19.7; 20.0; 19.6	Computer-administered program based on cognitive–behavioral interventions for body dissatisfaction (Butters & Cash, 1987). Intervention provided information on eating disorders, healthy weight-control behaviors, and nutrition and included unstructured e-mail-support interchange that allowed participants to talk about emotionally important material and reactions to the program. First and second versions were similar, but third version included interactive cognitive–behavioral exercises, in-person meetings with facilitators, weekly reading assignments, and critical reflection papers on readings. Third trial compared intervention with a university-taught eating disorder class.	Effects for body dissatisfaction at posttest and 3-month follow-up relative to waitlist controls in first trial, but no effects for knowledge, eating pathology, or body mass. Effects for body dissatisfaction at follow-up, but not at posttest, relative to waitlist controls in second trial, but no effects for eating pathology. Effects for body dissatisfaction at posttest, but not 6-month follow-up, and dieting at posttest and 6-month follow-up relative to waitlist controls in third trial, but no effects for bulimic symptoms. No effects on any outcomes for the eating disorder class relative to computer-administered intervention or waitlist controls.
Zabinski et al. (2001)	338 college women and men mean age = 24	Psychoeducational program promoted lifestyle change that incorporates regular physical activity.	No effects on body dissatisfaction or body mass at posttest relative to assessment-only controls.

Table 3
Moderator Values and Effect Sizes for Universal Programs

Study	Moderators										
	Sex	Age	Format	No. of sessions	Psycho-educational	Sociocultural	Healthy weight	Stress/ coping	Self-esteem	Body enhancement	Validated measures
Buddeberg-Fischer et al. (1998)	0	1	0	1	1	1	0	0	0	0	1
Dalle Grave et al. (2001)	0	0	1	1	1	1	1	0	0	0	1
Jerome (1987)	0	1	0	0	1	0	0	0	0	0	1
Jerome (1991)	0	1	0	0	1	0	0	0	0	0	1
Kater et al. (2002, girls)	1	0	0	1	1	1	1	1	0	0	0
Kater et al. (2002, boys)	0	0	0	1	1	1	1	1	0	0	0
Killen et al. (1993)	1	0	0	1	1	1	1	0	0	0	1
Kusel (1999)	1	0	1	1	0	1	0	0	0	0	1
Mann et al. (1997)	1	1	0	0	1	0	0	0	0	0	1
Martz & Bazzini (1999)	1	1	0	0	1	0	0	0	0	0	1
Martz & Bazzini (1999)	1	1	0	0	1	0	1	0	0	0	1
McVey & Davis (2002)	1	0	1	1	1	1	1	1	1	1	0
McVey et al. (in press)	1	0	1	1	1	1	1	1	1	1	0
Moreno & Thelen (1993, Study 1)	1	0	0	0	1	1	1	0	0	0	0
Moreno & Thelen (1993, Study 2)	1	0	0	0	1	1	1	0	0	0	0
Mutterperl & Sanderson (2002)	1	1	—	0	1	0	0	0	0	0	0
Nebel (1995)	1	1	1	1	1	0	1	1	1	1	1
Neumark-Sztainer et al. (1995)	1	1	1	1	1	1	1	0	0	0	0
Neumark-Sztainer et al. (2000)	1	0	1	1	1	1	0	0	0	0	0
Nicolino et al. (2001)	1	1	1	0	1	1	0	0	0	0	1
O'Dea & Abraham (2000)	0	0	1	1	0	0	0	1	1	0	1
Outwater (1991)	0	0	0	1	1	0	0	0	1	1	1
Paxton (1993)	1	0	1	1	1	1	1	0	0	0	1
Richman (1993)	0	0	0	1	1	1	1	0	1	0	0
Richman (1998)	0	0	0	1	1	1	1	0	1	0	1
Santonastaso et al. (1999)	1	1	1	1	1	1	0	0	0	0	1
Shepard (2001)	0	0	1	1	1	1	1	0	0	0	1
Smolak et al. (1998a)	0	0	0	1	1	1	1	0	0	1	0
Smolak et al. (1998b)	0	0	0	1	1	1	1	0	0	1	0
Stewart et al. (2001)	1	0	1	1	1	1	1	1	1	1	1
Varnado-Sullivan et al. (2001, girls)	1	0	1	1	1	1	1	0	0	1	1
Varnado-Sullivan et al. (2001, boys)	0	0	1	1	1	1	1	0	0	1	1
Wade et al. (2003, media)	0	0	1	1	0	1	0	0	0	0	1
Wade et al. (2003, esteem)	0	0	1	1	0	0	0	1	1	0	1
Weiss (2000)	1	0	1	1	1	1	1	0	1	1	1
Withers et al. (2002)	1	0	0	0	1	1	1	0	1	1	0
Wolf-Bloom (1999)	1	0	1	0	0	1	0	0	0	0	1

Note. Effect sizes (r) at posttest are reported first, and effect sizes for follow-up are reported second in parentheses; significant effects ($p < .05$) are marked by an asterisk. Effect sizes are for the intervention versus a waitlist or assessment-only control group when available. Dashes represent dependent variables that were not examined in particular studies. Moderators are coded such that 1 = yes, 0 = no.

offered to female-only samples (mean $r = .15$, $p < .001$) than were observed for interventions offered to samples containing male participants (mean $r = .05$, $p < .05$). Significantly ($z = 3.81$, $p < .001$) larger effects were observed for trials focusing on participants over age 15 (mean $r = .20$, $p < .001$) than were observed for trials focusing on younger participants (mean $r = .06$, $p < .001$). Interventions produced significantly ($z = 2.69$, $p < .01$) larger effects in trials that used validated measures (mean $r = .15$, $p < .001$) than in trials that used nonvalidated measures (mean $r = .03$, *ns*). In the multivariate model with the four variables that showed significant univariate effects, only the effect for selected versus universal focus ($z = 4.42$, $p < .001$) remained significant. There was no significant residual heterogeneity in effect sizes, $\chi^2(32, N = 33) = 25.89$, *ns*, when these four predictors were entered into the model.

Dieting. The weighted average effect size for dieting at termination across the 33 trials was .11 ($p < .001$). There was significant heterogeneity of effect sizes, $\chi^2(32, N = 33) = 59.76$, $p < .005$. There was insufficient variability in the number of sessions for this variable to be included in these moderator analyses. A univariate model indicated that selected programs (mean $r = .18$, $p < .001$) produced significantly ($z = 2.88$, $p < .005$) larger decreases in dieting than did universal programs (mean $r = .07$, $p < .005$). Studies that used interactive programs produced significantly ($z = 2.54$, $p < .05$) larger effects (mean $r = .13$, $p < .001$) than did studies that used didactic programs ($r = .02$, *ns*). Multisession programs produced significantly ($z = 3.27$, $p < .01$) larger effects (mean $r = .12$, $p < .001$) than did single-session programs (mean $r = -.05$, *ns*). In the multivariate model, the only significant predictor of effect size was whether a program had a selected versus a universal focus ($z = 2.13$, $p < .05$). There was no significant residual heterogeneity in effect sizes, $\chi^2(29, N = 30) = 29.91$, *ns*, when these three predictors were entered into the model.

Effect sizes for outcomes						
Knowledge	Body mass	Thin-ideal internalization	Body dissatisfaction	Dieting	Negative affect	Eating pathology
—	—	—	—	—	—	.05
.38* (.38*)	—	—	.10 (.08)	.01 (.00)	.02 (.01)	.21 (.10)
.29*	—	—	.00	.00	—	.00
.46*	—	—	.00	.00	.00	.00
—	—	.06	.10	—	.05	—
—	—	.08	.06	—	.11	—
.57*	.00	—	.00	(.00)	.00	.00 (.00)
—	—	.20 (.00)	.21* (.00)	.25* (.00)	.22* (.00)	.00 (.00)
—	.15 (.11)	—	.03 (.21)	−.22 (.03)	.03 (.08)	−.09 (.07)
—	(.01)	(.08)	(.00)	(.08)	—	—
—	(.00)	(.08)	(.06)	(.13)	—	—
—	—	—	−.07	—	—	−.02
—	—	—	.13* (.15*)	.15* (.14*)	.15* (.13*)	.17* (.10)
.45* (.42*)	—	—	—	—	—	—
.57* (.57*)	—	—	—	—	—	—
—	.03 (.03)	.01	—	—	—	.06 (.06)
.03	—	—	.05	—	.07	.21*
.12* (.12)	.09 (.13*)	—	.03 (.03)	.09 (.07)	.00 (.05)	.15* (.15*)
.14* (.04)	—	.05 (.14*)	.10 (.06)	.07 (.04)	—	.03 (.05)
—	—	(.15)	(.02)	(.20)	(.08)	—
—	(−.11*)	.02 (.13*)	.14* (.05)	.03 (.05)	.05 (.05)	.05 (.05)
—	—	—	.08 (.08)	—	.01 (.01)	—
—	(.03)	(.10)	(.05)	(.13)	(.04)	(.17)
.39*	—	—	.24*	.00	—	.05
.37* (.33*)	—	—	.15* (.07)	.11* (.05)	—	.03 (.08)
—	(.01)	—	(.14)	—	(.11)	(.10)
.57* (.40*)	—	.04 (.04)	.03 (.02)	—	—	.13 (.11)
—	—	(.18*)	(−.05)	(.00)	—	—
(.09)	—	(.00)	(−.05)	(.00)	—	—
.71* (.46*)	—	—	.11* (.06)	.19* (.08*)	.06 (.04)	.12* (.10*)
—	—	.43*	—	.05	.06	.28*
—	—	.21*	—	.10	.31*	.12
—	—	—	.16 (.06)	.18 (.03)	.03 (.17)	—
—	—	—	.05 (.05)	.03 (.00)	.11 (.07)	—
.20* (.20*)	—	—	.03 (.13*)	.00 (.00)	.00 (.00)	.00 (.00)
.35* (.21*)	—	—	.04 (.03)	.00 (.00)	—	—
—	—	.04	.07	—	.01	.01

The weighted average effect size for dieting scores at follow-up across the 32 trials was .11 ($p < .001$). There was significant heterogeneity of effect sizes, $\chi^2(31, N = 32) = 66.58$, $p < .001$. Length of follow-up period was not significantly related to the effect sizes for dieting, so this variable was not used as a covariate in subsequent models. A univariate model indicated that selected programs produced significantly ($z = 6.40$, $p < .001$) greater decreases in dieting (mean $r = .24$, $p < .001$) than did universal programs (mean $r = .05$, $p < .005$). Significantly ($z = 3.77$, $p < .001$) larger decreases in dieting were observed for trials that focused on participants older than 15 (mean $r = .19$, $p < .001$) than were observed for trials that focused on younger participants (mean $r = .06$, $p < .005$). Interactive programs produced significantly ($z = 2.69$, $p < .01$) larger effects (mean $r = .14$, $p < .001$) than did didactic programs (mean $r = .03$, *ns*). In the multivariate model with the three variables that showed significant univariate effects, only the effect for selected versus universal focus ($z = 2.13$, $p < .05$) remained significant. There was no significant residual heterogeneity in effect sizes, $\chi^2(29, N = 30) = 29.91$, *ns*, when these four predictors were entered into the model.

Negative affect. The weighted average effect size for negative affect at termination across the 34 trials was .14 ($p < .001$). There was significant heterogeneity of effect sizes, $\chi^2(33, N = 32) = 77.41$, $p < .001$. There was insufficient variability in the number of sessions for this variable to be included in these moderator analyses. A univariate model indicated that selected programs produced significantly ($z = 6.20$, $p < .001$) larger decreases in negative affect (mean $r = .25$, $p < .001$) than did universal programs (mean $r = .06$, $p < .001$). Significantly ($z = 2.90$, $p < .005$) larger decreases in negative affect were observed for trials focusing on participants over age 15 (mean $r = .20$, $p < .001$) than were observed for trials with younger participants (mean $r = .07$,

Table 4
Moderator Values and Effect Sizes for Selected Prevention Programs

Study	Moderators										
	Sex	Age	Format	No. of sessions	Psycho-educational	Sociocultural	Healthy weight	Stress/ coping	Self-esteem	Body enhancement	Validated measures
Baranowski & Hetherington (2001)	1	0	1	1	1	0	1	0	1	1	1
Bearman et al. (2003)	1	1	1	1	0	0	0	0	0	1	1
Butters & Cash (1987)	1	1	1	1	0	0	0	1	0	1	1
Celio et al. (2000, computer)	0	1	1	1	1	0	1	0	0	1	1
Celio et al. (2000, class)	0	1	1	1	1	0	1	0	0	1	1
Chase (2001)	1	1	0	1	1	1	1	0	0	1	1
Dworkin & Kerr (1987, CT)	1	1	1	1	0	0	0	0	1	1	1
Dworkin & Kerr (1987, CBT)	1	1	1	1	0	0	0	0	1	1	1
Franko (1998)	1	1	1	1	0	1	1	1	0	1	1
Kaminski & McNamara (1996)	1	1	1	1	1	1	1	0	0	1	1
McVey et al. (2003)	1	0	1	1	1	1	1	1	1	1	1
Orjada & Stice (2003)	1	1	1	1	1	0	0	0	0	0	1
Presnell & Stice (2003)	1	1	1	1	0	0	1	0	0	0	1
Rosen et al. (1989)	1	1	1	1	0	0	0	0	0	1	1
Stice, Mazotti, et al. (2000)	1	1	1	1	0	1	0	0	0	0	1
Stice & Ragan (2002)	1	1	1	1	1	0	0	0	0	0	1
Stice, Presnell, et al. (2003, dissonance)	1	1	1	1	0	1	1	0	0	0	1
Stice, Presnell, et al. (2003, healthy weight)	1	1	1	1	0	1	1	0	0	0	1
Stice, Trost, & Chase (2003, dissonance)	1	1	1	1	0	1	1	0	0	0	1
Stice, Trost, & Chase (2003, healthy weight)	1	1	1	1	0	1	1	0	0	0	1
Winzelberg et al. (1998)	1	1	0	1	1	0	1	0	0	0	1
Winzelberg et al. (2000)	1	1	0	1	1	1	1	0	0	1	1
Zabinski et al. (2001)	0	1	0	1	1	1	1	0	0	1	1

Note. Effect sizes (r) at posttest are reported first, and effect sizes for follow-up are reported second in parentheses; significant effects ($p < .05$) are marked by an asterisk. Effect sizes are for the intervention versus a waitlist or assessment-only control group when available. Dashes represent dependent variables that were not examined in particular studies. Moderators are coded such that 1 = yes, 0 = no. CT = cognitive therapy; CBT = cognitive–behavioral therapy.

$p < .001$). Interactive programs produced significantly ($z = 2.28$, $p < .05$) larger effects (mean $r = .16$, $p < .001$) than did didactic programs (mean $r = .03$, *ns*). Programs with psychoeducational content (mean $r = .07$, $p < .001$) produced significantly smaller effects ($z = -3.40$, $p < .001$) than did programs without this content (mean $r = .22$, $p < .001$). In the multivariate model with the four variables that showed significant univariate effects, only the effect for selected versus universal focus ($z = 3.91$, $p < .001$) remained significant. There was no significant residual heterogeneity in effect sizes, $\chi^2(29, N = 30) = 32.89$, *ns*, when all four predictors were entered into the model.

The weighted average effect size for negative affect at follow-up across the 23 trials was .09 ($p < .001$). However, because there was no significant heterogeneity of effect sizes for negative affect at follow-up, $\chi^2(22, N = 23) = 27.04$, *ns*, we did not test for effect size moderators.

Eating pathology. The weighted average effect size for eating pathology at termination across the 39 trials was .12 ($p < .001$). There was significant heterogeneity of effect sizes, $\chi^2(38, N = 39) = 83.11$, $p < .001$. A univariate model indicated that selected programs produced significantly ($z = 5.34$, $p < .001$) larger decreases in eating pathology (mean $r = .24$, $p < .001$) than did universal programs (mean $r = .07$, $p < .001$). Significantly ($z = 2.55$, $p < .05$) larger decreases in eating pathology were observed for trials focusing on participants that were older than 15 (mean $r = .17$, $p < .001$) than were observed for trials focusing on younger participants (mean $r = .08$, $p < .001$). Interactive programs produced significantly ($z = 3.93$ $p < .001$) larger effects (mean $r = .17$, $p < .001$) than did didactic programs (mean $r = .02$, *ns*). Multisession programs also produced significantly ($z = 2.48$, $p < .05$) larger effects (mean $r = .14$, $p < .001$) than did single-session programs (mean $r = -.01$, *ns*). Significantly smaller effects ($z = -2.60$, $p < .01$) occurred for programs with psychoeducational content (mean $r = .09$, $p < .001$) than occurred for those without this content (mean $r = .21$, $p < .001$). In the multivariate model with the five variables that showed significant univariate effects, the effects for selected versus universal focus ($z = 2.47$, $p < .05$), program format ($z = 3.58$, $p < .001$), and age ($z = 2.08$, $p < .05$) remained significant. There was no significant residual heterogeneity in effect sizes, $\chi^2(32, N = 33) = 25.59$, *ns*, when these five predictors were entered into the model.

The weighted average effect size for eating pathology at follow-up across the 28 trials was .12 ($p < .001$). There was significant heterogeneity of effect sizes, $\chi^2(27, N = 28) = 42.81$, $p < .05$. Length of follow-up period was not significantly related to the effect sizes for eating pathology, and therefore it was not included as a covariate in these models. A univariate model indicated that selected programs produced significantly ($z = 4.80$, $p < .001$) larger decreases in eating pathology (mean $r = .22$, $p < .001$) than did universal programs (mean $r = .07$, $p < .001$). Significantly ($z = 2.93$, $p < .005$) larger decreases in eating pathology occurred for trials focusing on participants older than 15

Effect sizes for outcomes						
Knowledge	Body mass	Thin-ideal internalization	Body dissatisfaction	Dieting	Negative affect	Eating pathology
—	—	—	.29 (.11)	.40* (.38*)	.19 (.14)	.19 (.14)
—	—	—	.39* (.27*)	.13 (.11)	.38* (.28*)	.22* (.20*)
—	—	—	.50*	—	.59*	—
—	—	—	.29* (.17)	.30* (.47*)	—	.26 (.18)
—	—	—	.13 (.27)	.11 (.23)	—	.12 (.05)
.45* (.42*)	—	.10 (.08)	.14* (.11)	.21* (.20*)	.14 (.02)	.00 (.00)
—	—	—	.21*	—	.18*	—
—	—	—	.34*	—	.31*	—
—	—	.31	.20	—	—	.15
—	—	.55* (.44*)	.55* (.56*)	.43* (.44*)	.47* (.55*)	—
—		—	.23* (.23*)	.17* (.27*)	—	.18* (.27*)
—	.38*	.32*	.11	.05	.19	.30*
—	.40*	—	—	.12	.20	.32*
—	—	—	.50* (.32*)	—	—	—
—	—	.40* (.43*)	.46* (.42*)	.36 (.27)	.42* (.32)	.37* (.37*)
—	.32*	.38*	.25*	.31*	.14	.34*
—	.05 (.15*)	.31* (.17*)	.28* (.30*)	.18* (.18*)	.23* (.11)	.17* (.16*)
—	.12 (.12)	.20* (.21*)	.16* (.24*)	.05 (.13)	.16* (.04)	.20* (.17*)
—	—	.24 (.15)	.23 (.13)	.17 (.09)	.27* (.26*)	.23* (.21*)
—	—	.25 (.20)	.12 (.15)	.06 (.13)	.34* (.23*)	.31* (.22*)
—	.11 (.22)	—	.32* (.31*)	—	—	.10 (.20)
—	—	—	.17 (.35*)	—	—	.16 (.23)
—	.07	—	.01	—	—	—

(mean $r = .17$, $p < .001$) than for trials focusing on younger participants (mean $r = .08$, $p < .001$). Interactive programs produced significantly ($z = 1.99$, $p < .05$) larger effects (mean $r = .14$, $p < .001$) than did didactic programs (mean $r = .05$, *ns*). Programs that contained psychoeducational content (mean $r = .09$, $p < .001$) had significantly smaller effects ($z = -2.04$, $p < .001$) than did those without this content (mean $r = .18$, $p < .001$). In the multivariate model with the four variables that showed significant univariate effects, the effects for selected versus universal focus ($z = 3.02$, $p < .005$) and program format ($z = 2.05$, $p < .05$) remained significant. There was no significant residual heterogeneity in effect sizes, $\chi^2(22, N = 23) = 14.47$, *ns*, when these four predictors were entered into the model.

Body mass. The weighted average effect size for body mass at termination across the 11 prevention trials was .12 ($p < .005$). There was statistically significant heterogeneity of effect sizes, $\chi^2(10, N = 11) = 25.95$, $p < .005$. There was insufficient variability in participant sex, age, body satisfaction content, self-esteem content, and stress and coping content for these moderators to be examined. Interactive programs produced significantly ($z = 2.15$, $p < .05$) larger effects (mean $r = .20$, $p < .005$) than did studies that used didactic programs ($r = .02$, *ns*). Programs emphasizing sociocultural resistance skills (mean $r = .03$, *ns*) had significantly smaller effects ($z = -3.42$, $p < .001$) than did programs without this content (mean $r = .23$, $p < .001$). In the multivariate model, both program format ($z = 2.54$, $p < .05$) and sociocultural resistance skill content ($z = -3.75$, $p < .001$) remained significant. There was no significant heterogeneity of effect size in this model, $\chi^2(6, N = 7) = 2.86$, *ns*.

The weighted average effect size for body mass at follow-up across the 11 trials was .05 (*ns*). Moreover, because there was no significant heterogeneity of effect sizes for body mass at follow-up, $\chi^2(10, N = 11) = 18.26$, *ns*, it was not appropriate to test for moderators of effect size.

Discussion

Summary of Average Effect Sizes

The average effect sizes for the outcomes ranged from .11 to .38 at termination and from .05 to .29 at follow-up. Only one of these average effects was not significantly larger than zero. These average effects were small to medium according to Cohen's (1988) criteria. On the one hand, it is encouraging that the average effect sizes were at least small in magnitude, given that the first generation of eating disorder prevention programs was unsuccessful at reducing risk factors for eating disorders or eating pathology. Although recent reviews have commented on the discouraging findings from prevention efforts (e.g., Levine & Smolak, 2001; Striegel-Moore & Steiner-Adair, 1998), the findings from trials published since 2000 suggest there is more cause for optimism. For example, one intervention reduced the rate of threshold and subthreshold eating pathology from 15% in the control condition to 6% in the intervention condition (Stice, Fisher, & Martinez, in

press). Assuming the same recruitment rate for high-risk participants (7%) and the same dropout rate (3%), if this selected prevention program were offered to all 9,844,000 adolescent girls in the United States between the ages of 15 and 19 (U.S. Census Bureau, 2002), it could reduce the rate of threshold and subthreshold eating pathology by a projected 60,157 cases. On the other hand, it was disappointing that the average effects found in this meta-analysis were relatively small. Fortunately, the individual effect sizes from the prevention trials ranged from nonexistent to large, and the moderation analyses identified a number of factors that were associated with larger effects.

Summary of Effect Size Moderators

Risk status of participants. Results indicated that selected programs that were provided to high-risk individuals produced significantly larger intervention effects than did universal programs that were provided to unselected samples for five out of the seven outcomes. It was noteworthy that 10 of the 15 interventions that produced effects for eating pathology were selected. It was also striking that only selected interventions prevented the future increases in eating pathology that were observed in control groups. This finding is important because it suggests that the intervention effects are not merely resulting because the programs decrease initial elevations in eating disturbances. Mirroring the pattern of findings across studies, there was also evidence from within samples that supported this conclusion. Several universal prevention programs were more effective for subgroups of high-risk participants than for the full sample (Buddeberg-Fischer, Klaghofer, Gnam, & Buddeberg, 1998; Killen et al., 1993; Stewart, Carter, Drinkwater, Hainsworth, & Fairburn, 2001; Weiss, 2000). Prevention programs for depression (Clarke et al., 1995), anxiety (Lowry-Webster, Barrett, & Dadds, 2001), behavior problems (Stoolmiller, Eddy, & Reid, 2000), and substance abuse (Murphy et al., 2001) also have been found to produce stronger effects for people with initially elevated symptoms. Theoretically, the subjective distress that characterizes high-risk individuals motivates these participants to engage more effectively in the prevention program, which should produce greater intervention benefits. Further, the relatively lower levels of eating pathology and risk factors for eating pathology in unselected samples relative to high-risk samples may serve to attenuate intervention effects.

Program format. Findings provided support for the assertion that intervention effects would be stronger for programs that were interactive rather than didactic. Intervention effects were significantly stronger for interactive versus didactic programs for five of the seven outcomes. It was noteworthy that all of the interventions that produced effects for eating pathology were interactive. Prevention researchers in other areas have similarly concluded that psychoeducational interventions are less effective than interventions that actively engage participants and teach new skills (Clarke et al., 1993; Larimer & Cronce, 2002). We suspect that an interactive format might be necessary to ensure that participants engage in the program material, which likely facilitates acquisition of concepts and skills and promotes attitudinal and behavioral change.

Participant age. Results also supported the hypothesis that intervention effects would be stronger for samples in which the average age of participants was over 15 versus those that targeted younger adolescents. Intervention effects were significantly larger for older adolescents for five of the seven outcomes. Eleven of the 15 programs that produced effects on eating pathology focused on samples with a mean age of 15 or greater. This pattern of findings provides little support for the suggestion that past prevention programs have shown limited success because participants were too old when the program was delivered (Smolak et al., 1998a). The interventions may be more effective for older adolescents because they were delivered during the period of greatest risk for emergence of the pathological condition, when these individuals are more motivated to engage in the intervention (because of elevated subjective distress). It is also possible that younger adolescents possess more limited insight, based on the fact that their abstract reasoning skills are still developing, which may constrain their ability to benefit from interventions. A final possible explanation is that younger girls have such low levels of eating disturbances and risk factors for eating pathology that it attenuates statistical power to detect intervention effects.

Participant sex. There was also support for the hypothesis that intervention effects would be larger for samples composed solely of females. Intervention effects were significantly larger for programs focusing on females versus those that included males for two of the seven outcomes. It was surprising that this effect only emerged for a subset of the outcomes examined in this review. Our power to detect effects of this potential moderator might have been attenuated by the fact that 70% of the trials targeted females. We theorized that intervention effects would be more pronounced for female samples because the elevated levels of body image and eating disturbances that occur for this sex might motivate this subpopulation to engage more effectively in the intervention and because there might be floor effects for samples containing males.

Number of sessions. Results also provided some support for the hypothesis that brief single-session programs would be less effective than longer multisession programs. Single-session programs produced significantly smaller intervention effects than did multisession programs for three of the seven outcomes examined. It was noteworthy that not 1 of the 11 single-session interventions produced significant effects for eating pathology or risk factors for eating pathology. Multisession interventions theoretically produce stronger effects because it is useful for participants to reflect on intervention material between sessions to maximize internalization of program tenets and because intersession periods give participants a chance to try new skills and then return to the group for troubleshooting advice. It is also possible that repeated meetings with a group of participants struggling with the same issue facilitate group cohesion and perceived support and acceptance and that these factors contribute to the positive intervention effects.

Within this context, we should acknowledge that it is desirable to have prevention programs that are relatively brief. Such programs are more easily disseminated because they are less expensive, more appealing to adolescents, and more acceptable to school administrators with numerous demands for class time. Fortunately, this review identified several effective eating disorder prevention programs that were brief (three to four sessions long).

Program content. One of the factors that would be expected to contribute to effective eating disorder prevention programs is the content of the intervention. However, our results suggest that content is less important in predicting intervention effects than are features of the participants studied and research design. Indeed,

three of the content codes did not show a relation to any of the outcomes and two only showed a relation to one of the seven outcomes. These results may imply that aspects of intervention delivery (e.g., use of an interactive format vs. a didactic format) or of the population targeted (e.g., high-risk vs. all adolescents) may be more important than content. This pattern of findings may also suggest that nonspecific factors are more important than program content in producing intervention effects (although this interpretation does not seem to accord with the fact that only a portion of the programs produced effects). Another alternative is that it may be necessary to combine the proper content with the optimal delivery format and audience type to produce the strongest eating disorder prevention effects. Unfortunately, we did not have sufficient power to test whether content interacted with delivery format and audience type in the prediction of effect sizes.

Nonetheless, there was support for the hypothesis that programs with psychoeducational content are less effective than those without this content for four out of the seven outcomes. These findings echo the conclusions drawn from other prevention areas that psychoeducational interventions are ineffective in producing behavioral change (e.g., Larimer & Cronce, 2002). There was also evidence that programs that focused on sociocultural factors and on stress and coping produced smaller effects for one of the outcomes. However, because these findings did not replicate across multiple outcomes, these results should be interpreted cautiously. It was interesting that the one outcome that was related inversely to a focus on sociocultural pressures was body mass. This might suggest that programs that reduce the impact of cultural pressures for thinness have the unintended effect of condoning obesity—a thesis that has been explored by others (Heinberg, Thompson, & Matzon, 2001).

Within this context it is important to note, however, that the content of the 15 programs that produced intervention effects for eating pathology varied dramatically, including programs that focused on promoting self-esteem, stress management skills, healthy weight-control behaviors, and critical analysis of the thin ideal. This pattern of findings suggests that there may be multiple methods to successfully prevent eating pathology. Nevertheless, the programs that produced the most promising effects appeared to include cognitive interventions that alter maladaptive attitudes, such as thin-ideal internalization or body dissatisfaction, and behavioral interventions that alter maladaptive behaviors, such as fasting and overeating. Accordingly, research should continue to investigate these various promising interventions.

Use of validated measures. There was also support for the hypothesis that intervention effects would be larger for trials that used outcome measures with established reliability and validity. Specifically, intervention effects were significantly larger for trials that used validated scales for two of the seven outcomes examined. We suspect that prevention trials that used validated outcome measures simply have more sensitivity to detect intervention effects. One implication of this finding is that prevention researchers should devote more attention to measurement, as many investigators create their own scales for their trials without conducting psychometric analyses to verify that the scales possess reliability and validity within the population to be studied. An alternative explanation is that better outcomes are associated with use of validated measures solely because use of such measures is a proxy for more skilled or experienced research groups, which also tend to generate more promising interventions.

Qualifications of Moderator Analyses

First, many of the conclusions regarding moderators of intervention effects are based on theoretical considerations rather than on direct empirical evidence. For example, although we conjecture that selected prevention programs produce large effects because the subjective distress that ostensively characterizes high-risk samples leads them to engage more effectively in the program content, to our knowledge no trials have confirmed this with direct measures of distress or engagement. Thus, an important direction for future trials will be to directly test these hypothesized explanations for the moderators of intervention effects.

Second, our power to detect the effects of certain moderators was limited by the modest sample sizes involved in some analyses (e.g., some outcomes were only examined in a small subset of trials) and by the fact that certain values of moderators were rare (e.g., trials focusing on males). The implication of this limitation is that the moderation analyses reported here might represent a conservative estimate of the effects of these factors.

Third, it should be acknowledged that there is insufficient information at this time to draw conclusions regarding certain factors that might moderate intervention effects. For example, it might be fruitful for future prevention trials to investigate other potential moderators of intervention effects, such as the presence of other psychiatric disturbances (mood disorders or substance abuse), ethnic minority status, and intellectual capacity of participants.

Theoretical Limitations of Literature

The most noteworthy theoretical limitation of the eating disorder prevention literature is that etiologic models of eating pathology were not used to guide the design of many of the prevention programs. Most of the early programs focused on providing information about eating pathology, despite the fact that etiologic models do not posit that a lack of information concerning the ill effects of eating pathology is a risk factor for development of these conditions. The evidence that psychoeducational content was inversely related to intervention effects from the current study seems to verify that the provision of knowledge is not an effective approach to preventing eating pathology. It seems preferable to design interventions that reduce established risk factors for eating pathology and increase protective factors. In support of this conjecture, most of the prevention programs that produced intervention effects for eating pathology appear to have focused on reducing empirically established risk factors for eating pathology. One implication of the suggestion to focus on empirically established risk factors is that it might be fruitful for future prevention trials to attempt to modify societal factors that putatively increase the risk for eating disorders (e.g., the glorification of thinness in the mass media), in addition to building resilience to these pressures at the individual level.

Second, little theoretical attention has focused on the specific factors that putatively mediate (i.e., account for) the effects of the various interventions. A clear articulation of the mechanisms by which an intervention is hypothesized to produce reductions in eating pathology is crucial because it yields falsifiable hypotheses

that may lead to refinement of prevention theory. Each intervention has a unique set of mediators that relates to the content covered in the intervention. For example, media literacy interventions would be expected to result in decreased body dissatisfaction and eating disturbances because this type of intervention should result in decreased internalization of the thin ideal and social comparisons with ultrathin media images. Each intervention also has a unique profile of nonspecific factors that depends on the delivery format used for the intervention. For instance, perceived group support may be an important mediator for group-administered interventions but may not be germane to didactic universal programs or interventions delivered over the computer. The fact that several interventions did not produce significantly stronger effects than minimal-intervention control conditions (e.g., Celio et al., 2000; Mutterperl & Sanderson, 2002; Nicolino, Martz, & Curtin, 2001; but see also Dworkin & Kerr, 1987; Rosen et al., 1989; Stice, Presnell, Shaw, & Hoffman, 2003) suggests that nonspecific factors should be examined as potential mediators. Refinement of the theory regarding factors that mediate intervention effects may facilitate the design of more effective interventions.

Third, randomized prevention trials offer a unique opportunity to provide experimental tests of etiologic theory. Prevention programs that focus on reducing a sole putative risk factor (e.g., body dissatisfaction) provide a strong test of the effect of that factor on eating pathology. Such experimental tests of etiologic theory are vital because there is always a possibility that some confounding variable explains the relation between a putative risk factor and future eating pathology from prospective studies (Kraemer, Stice, Kazdin, & Kupfer, 2001). For example, it was noteworthy that two of the successful eating disorder prevention programs increased a variable that is widely accepted to be a risk factor for eating pathology: dieting. Interventions that entailed either a low-calorie diet for weight loss purposes (Presnell & Stice, 2003) or lifestyle changes in caloric intake and exercise for weight maintenance purposes (Stice, Presnell, et al., 2003) resulted in significant reductions in bulimic pathology. These findings converge with the evidence from two randomized obesity treatment trials that low-calorie diets result in decreased binge eating (Goodrick, Poston, Kimball, Reeves, & Foreyt, 1998; Reeves et al., 2001). Uncontrolled trials have also found that placing individuals on low-calorie diets results in decreased binge eating (Epstein, Paluch, Saelens, Ernst, & Wilfley, 2001; Telch & Agras, 1993; Wadden, Foster, & Letizia, 1994).[7] These results suggest the need for theoretical refinement of one of the most widely accepted etiologic risk factors for eating pathology.

Fourth, it is our impression that it would be beneficial if greater theoretical attention were focused on defensiveness and insight. This was suggested by an interesting trend that emerged in this review, wherein many of the most successful interventions were not presented as eating disorder prevention programs to the participants. Specifically, it appeared that 10 of the 15 prevention programs that produced significant effects for eating pathology were not presented as eating disorder prevention programs (many were simply described as body acceptance interventions).[8] This pattern of findings suggests the intriguing possibility that prevention researchers should focus greater effort at developing what has been referred to as *covert prevention programs* (Stice & Ragan, 2002). Individuals may be less defensive about their body image and eating disturbances and thus more willing to entertain alternative perspectives when they are not aware that they are in an intervention focusing on these outcomes. This conjecture accords with the evidence that selected interventions that sought to recruit high-risk participants with eating disorder screening measures were ineffective (e.g., Varnado-Sullivan et al., 2001). An additional benefit of a covert prevention approach is that it does not require insight regarding eating disturbances on the part of the participants to entice them into enrolling in the programs.

Methodological Limitations of Literature

The most important methodological limitation of the eating disorder prevention literature is that over 25% of eating disorder prevention trials located in our literature search did not include a control group (and were therefore excluded from this review). Without a waitlist or assessment-only control group, it is not possible to disentangle true intervention effects from the effects of passage of time, regression to the mean, or measurement artifacts. However, there appears to be little evidence that repeated administration of assessment measures produces artificial decreases in scores. Trials that randomly assigned participants to an assessment-only control group that completed pretest and posttest measures or to posttest-assessment-only conditions have provided no evidence of measurement artifact effects (Kusel, 1999; Mann et al., 1997; Martz & Bazzini, 1999).

Second, only a handful of prevention trials used minimal-intervention or placebo control conditions. Without such control groups, there is no way to rule out the possibility that any observed decreases in the outcomes are due to demand characteristics or expectancy effects. The fact that many of the eating disorder prevention programs did not produce significant reductions in eating pathology suggests that demand characteristics and expectancy effects do not produce artificial findings. However, the one study that included both a minimal-intervention and an assessment-only control condition found that there were significantly greater decreases in risk factors and eating pathology in the minimal-intervention condition relative to the assessment-only control condition from pretest to posttest (Stice, Presnell, et al., 2003).[9] It must be noted, however, that it can be difficult to identify a placebo intervention that is truly ineffective. Unlike

[7] Telch and Agras (1993) concluded that their results suggested that assignment to a low-calorie diet resulted in increased binge eating because initially non-binge-eating participants showed an increase in binge frequency after the low-calorie diet was terminated. However, this result appears to be an artifact of regression to the mean, in that initial non-binge-eating participants showed a general increase in binge eating after the termination of the diet and initial binge eaters showed a corresponding general decrease in this outcome.

[8] We did not include this factor as a moderator of intervention effect sizes because we were only moderately confident in our ability to reliably code which interventions were presented as eating disorder prevention programs and which were not.

[9] Four trials compared an intervention to a minimal-intervention or placebo control group in the absence of a waitlist or assessment-only control group. However, this design provides little information regarding whether there is a placebo response because nonsignificant differences between the intervention and placebo condition on the outcomes could result because there is a placebo response or because the intervention was not effective.

medication trials in which a pill that is known to have no therapeutic effect can be used as the placebo intervention, it is difficult to determine whether a putative therapeutic placebo intervention produces an intervention effect because it is a placebo response or because it represents an alternative treatment. There are examples from both the psychotherapy treatment literature (Fairburn et al., 1993) and the prevention literature (Stice, Trost, & Chase, 2003) of effective interventions that were originally intended to serve as placebo control conditions and were ultimately found to outperform alternative interventions. Thus, it might be best for future prevention trials to use minimal-intervention control conditions. It is also important to note that care should be taken in selecting a placebo or minimal-intervention control condition because it determines what inferences one can draw regarding potential confounds. For example, if the investigators think that nonspecific group factors are clinically important, they should select a placebo or minimal-intervention condition that does not contain these nonspecific factors.

A third methodological limitation is that many prevention trials did not randomly assign participants to condition. Greater confidence can be placed in inferences from randomized prevention trials because this type of study is better positioned to rule out any potential third-variable confounds (because all variables should be roughly equally distributed across conditions with large enough samples).

Fourth, numerous trials apparently did not assess eating pathology (over 30% of those included in this review). This is troublesome because the defining feature of a successful eating disorder prevention program is that it reduces current or future eating disorder symptoms or syndromes—otherwise it should not logically be considered an eating disorder prevention program. The assumption of interventions that target empirically established risk factors for eating pathology is that by reducing the risk factors, consequent reductions in eating pathology should be observed. Such trials should include a measure of eating pathology so that this assumption can be directly tested.

Fifth, few studies have assessed clinically significant eating pathology. Many of the trials that included an eating disorder measure used scales that do not directly assess the diagnostic symptoms of eating disorders. This makes it difficult to establish whether the intervention effects are clinically meaningful. For example, 5 of the 15 studies that reported significant intervention effects for eating pathology used scales that do not assess *DSM–IV* (*Diagnostic and Statistical Manual of Mental Disorders*, 4th ed., American Psychiatric Association, 1994) eating disorder symptoms. Future trials should be sure to include an outcome measure that assesses the key diagnostic features of the currently recognized eating disorders. It is preferable to use clinical interviews to ensure that only clinically significant eating pathology is recorded. This also helps minimize the false-positive problem that occurs with self-report measures of eating pathology (Fairburn & Beglin, 1994). In addition, it would be useful to include clinically meaningful measures of social functioning and mental health service utilization to further establish the clinical significance of eating disorder prevention trials.

A final troublesome feature of past trials is that they have sometimes used scales that do not tap a conceptually clear unidimensional construct (e.g., the Drive for Thinness Scale, Weight Concerns Scale, and the Restraint Scale). This state of affairs renders findings for these outcome measures difficult to interpret unambiguously.

Statistical Limitations of Literature

The most serious statistical limitation of extant prevention trials is that numerous evaluations did not conduct appropriate inferential tests of intervention effects. Similar to randomized clinical trials, it is necessary to test whether there is significantly greater change (decreases or increases) on the outcome variable in the intervention versus control condition. This can be accomplished with repeated measures ANOVA models or some sort of analysis of covariance models that control for initial levels of the outcome variables (i.e., ANCOVA models, survival models, or growth curve models). Some investigators tested whether there was significant change in both intervention and control conditions without directly testing whether there was significantly greater change in the intervention condition. If there was not significantly greater change in the intervention versus control condition, the effects may be due to some alternative explanation (e.g., regression to the mean or passage of time). Other investigators compared the mean levels of the outcome variables across the two groups at follow-up assessments without controlling for initial levels. The problem with this approach is that any differences between the groups at baseline (even if nonsignificant) can artificially inflate or attenuate the observed intervention effects.

Within this context, it is important to note that there is currently no statistically reliable evidence of iatrogenic effects from eating disorder prevention programs. Mann et al. (1997) found no evidence of any intervention effects in either a positive or negative direction when they tested for differential change in the outcomes across conditions in the same participants over time. It was only when they compared follow-up data from different groups of participants who completed surveys at the assessments without controlling for initial levels of the outcome that they found what they interpreted as iatrogenic effect. Other researchers have interpreted the results from Carter, Stewart, Dunn, and Fairburn (1997) as suggesting that prevention programs produce iatrogenic effects. However, because this study did not include a control group, there is no way of ruling out the possibility that control participants would have shown similar, or even greater, increases in dietary restraint over the same 6-month period. Finally, some reviewers have interpreted any evidence of symptom worsening among any participants in an intervention condition as indicative of an iatrogenic effect. However, care should be taken to distinguish statistically reliable evidence of greater symptom worsening in the intervention versus control group from chance fluctuation in symptoms. Only the former should be considered reliable evidence of iatrogenic effects—the latter reflects random noise.

A second statistical limitation of the eating disorder prevention literature is that almost no trials reported effect sizes. This information is vital because it provides an indication of the clinical importance of prevention effects. For example, a prevention effect could be significant but might explain less than 1% of the variance in change in the outcome. Indices such as correlations or variance explained seem to be the most useful effect size measures because of their ease of interpretation and because they are scale invariant.

Third, few studies have conducted analyses that directly test whether the rates of clinically significant change were greater in

intervention versus control conditions. Because it has been argued that intervention effects can be statistically significant but not clinically significant (Jacobson & Truax, 1991), prevention trials should include analyses that directly test whether interventions produce higher rates of clinically significant change.

Fourth, few prevention trials have included tests of the specific factors that putatively mediate the effects of the interventions. This is an important omission because these analyses provide an opportunity to test the prevention theory that underlies interventions. For example, researchers should confirm that healthy weight interventions produce reductions in eating pathology because they promote reduced intake of high-fat and high-sugar foods and increased exercise. If intervention effects on eating pathology were observed in the absence of change in the putative mediators, this would suggest that the intervention theory is incorrect (e.g., nonspecific factors, such as perceived group support, mediate the intervention effects) or that the measures are unreliable. Such analyses may also reveal that interventions did not affect eating pathology because they failed to change the putative mediators (i.e., did not produce actual changes in healthy weight-control behaviors), which could facilitate interpretation of null findings. More generally, an improved understanding of the factors that mediate program effects may help researchers enhance critical program content that focuses on key mediators and prune unnecessary program content that plays no role in mediating intervention effects. Testing mediation is straightforward within the context of randomized trials (Kraemer, Wilson, Fairburn, & Agras, 2002), and there are examples from both the treatment (Wilson, Fairburn, Agras, Walsh, & Kraemer, 2002) and prevention literature (Bearman, Stice, & Chase, 2003; Stice, Mazotti, et al., 2000). Mediation analyses need to verify that the intervention produces effects for both eating pathology and the putative mediator (e.g., thin-ideal internalization), that change in the mediator predicts change in eating pathology, and that the effect of the intervention on eating pathology is reduced when change in the mediator is statistically controlled.

Fifth, prevention trials have rarely included analyses of factors that moderated intervention effects. Several studies have found that intervention effects were more pronounced for high-risk participants with initial elevations in body dissatisfaction (Celio et al., 2000; Killen et al., 1993; Stewart et al., 2001) and eating pathology (Buddeberg-Fischer et al., 1998; Stice et al., in press). These findings echo the trend observed in the current meta-analysis for selected eating disorder prevention programs to be more effective than universal programs. As noted above, these findings may occur because participants with extant body image and eating disturbances are more likely to engage in the prevention program or because there are floor effects for low-risk individuals. Future trials should attempt to identify subgroups of participants who show a particularly favorable response to the intervention (e.g., participants who are most distressed about their body image and eating disturbances) and a particularly unfavorable response (e.g., participants who deny the costs of pursuing the thin-ideal or have unrealistic perceptions of weight norms). This information may facilitate prevention-matching approaches wherein various subgroups are provided with interventions that should be maximally effective for that particular population. These analyses will need to test whether there is evidence of a significant interaction between the moderator and intervention condition in predicting change in the outcomes over time. Analyses that test for factors that are correlated with change in outcome solely among intervention participants are not appropriate because such analyses cannot separate true intervention effects from artifacts that may result in decreases in the outcomes (e.g., regression to the mean). Future trials should also explore the possibility of high-order interactions among moderators of intervention effects because this type of analyses may identify subpopulations that show exceptionally potent or weak intervention effects. Regression tree and classification tree analyses are particularly sensitive analytic approaches for detecting such higher order interactive effects.

Conclusions and Directions for Future Research

This review reveals that a number of promising eating disorder prevention programs have been developed and evaluated. Several prevention programs were found to decrease current eating pathology and the risk for future increases in eating pathology. A larger number of programs were found to decrease risk factors for eating pathology (e.g., body dissatisfaction). Certain intervention effects persisted as long as 2 years and were superior to minimal-intervention control conditions. The heterogeneity in the content of the successful programs implies that there may be several approaches to effectively preventing eating disturbances, but it appeared that successful programs often decreased attitudinal risk factors and promoted healthy weight-control behaviors. This meta-analytic review also suggested that larger intervention effects tended to occur in selected (vs. universal) programs, interactive (vs. didactic) programs, multisession (vs. single-session) programs, and for evaluations that used validated outcome measures. Larger intervention effects also tended to occur for programs offered solely to females and to participants older than 15 years of age. Results also indicated that smaller effects occurred for interventions with psychoeducational content. There was suggestive evidence that programs that were not explicitly presented as an eating disorder prevention program produced more positive effects.

It is hoped that the next generation of eating disorder prevention trials will build upon the promising emerging results and will attempt to address the conceptual, methodological, and statistical limitations enumerated above. It is particularly important for future studies to directly compare the more promising prevention programs in randomized trials, in an effort to isolate the most potent interventions and to provide a forum for exploring the possibility of treatment-matching approaches that might improve the overall success rate of prevention efforts. It would also be useful if greater attention focused on conducting independent replications of the most promising prevention trials. Moreover, greater research attention should be devoted to developing general prevention techniques that are independent of the specific content of the intervention, such as strategic self-presentation, motivational interviewing, and other persuasion techniques from social and clinical psychology. In addition, it is crucial to begin conducting effectiveness trials to determine whether interventions remain effective when delivered by natural clinicians (e.g., school counselors) in ecologically valid settings (e.g., schools). With continued methodologically rigorous and programmatic research, the ultimate goal of decreasing the overall prevalence of eating disorders in the population may yet be realized.

References

Agras, W. S., Walsh, B. T., Fairburn, C. G., Wilson, G. T., & Kraemer, H. C. (2000). A multicenter comparison of cognitive-behavioral therapy and interpersonal psychotherapy for bulimia nervosa. *Archives of General Psychiatry, 57,* 459–466.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Baranowski, M. J., & Hetherington, M. M. (2001). Testing the efficacy of an eating disorder prevention program. *International Journal of Eating Disorders, 29,* 119–124.

Bearman, S. K., Stice, E., & Chase, A. (2003). Effects of body dissatisfaction on depressive and bulimic symptoms: A longitudinal experiment. *Behavior Therapy, 34,* 277–293.

Brownell, K. D. (1997). *The LEARN program for weight control.* Dallas, TX: American Health.

Buddeberg-Fischer, B., Klaghofer, R., Gnam, G., & Buddeberg, C. (1998). Prevention of disturbed eating behaviour: A prospective intervention study in 14- to 19-year-old Swiss students. *Acta Psychiatrica Scandinavica, 98,* 146–155.

Butters, J. W., & Cash, T. F. (1987). Cognitive–behavioral treatment of women's body-image dissatisfaction. *Journal of Consulting and Clinical Psychology, 55,* 889–897.

Carter, J. C., Stewart, D. A., Dunn, V. J., & Fairburn, C. G. (1997). Primary prevention of eating disorders: Might it do more harm than good? *International Journal of Eating Disorders, 22,* 167–172.

Celio, A. A., Winzelberg, A. J., Wilfley, D. E., Eppstein-Herald, D., Springer, E. A., Dev, P., & Taylor, C. B. (2000). Reducing risk factors for eating disorders: Comparison of an Internet- and classroom-delivered psychoeducational program. *Journal of Consulting and Clinical Psychology, 68,* 650–657.

Chase, A. K. (2001). Eating disorder prevention: An intervention for at risk college women. *Dissertation Abstracts International, 62*(3), 1568B.

Clarke, G. N., Hawkins, W., Murphy, M., & Sheeber, L. (1993). School-based primary prevention of depressive symptomatology in adolescents: Findings from two studies. *Journal of Adolescent Research, 8,* 183–204.

Clarke, G., Hawkins, W., Murphy, M., Sheeber, L., Lewinsohn, P. M., & Seeley, J. R. (1995). Targeted prevention of unipolar depressive disorder in an at-risk sample of high school adolescents: A randomized trial of group cognitive intervention. *Journal of the American Academy of Child & Adolescent Psychiatry, 34,* 312–321.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dalle Grave, R. D., De Luca, L., & Campello, G. (2001). Middle school primary prevention program for eating disorders: A controlled study with a twelve-month follow-up. *Eating Disorders, 9,* 327–337.

Dworkin, S. H., & Kerr, B. A. (1987). Comparison of interventions for women experiencing body image problems. *Journal of Counseling Psychology, 34,* 136–140.

Epstein, L. H., Paluch, R. A., Saelens, B. E., Ernst, M. M., & Wilfley, D. E. (2001). Changes in eating disorder symptoms with pediatric obesity treatment. *Journal of Pediatrics, 139,* 58–65.

Fairburn, C. G., & Beglin, S. J. (1994). Assessment of eating disorders: Interview or self-report questionnaire? *International Journal of Eating Disorders, 16,* 363–370.

Fairburn, C. G., Cooper, Z., Doll, H. A., Norman, P. A., & O'Connor, M. E. (2000). The natural course of bulimia nervosa and binge eating disorder in young women. *Archives of General Psychiatry, 57,* 659–665.

Fairburn, C. G., Jones, R., Peveler, R. C., Hope, R. A., & O'Connor, M. (1993). Psychotherapy and bulimia nervosa: Longer-term effects of interpersonal psychotherapy, behavior therapy, and cognitive behavior therapy. *Archives of General Psychiatry, 50,* 419–428.

Fields, S. A., & McNamara, J. R. (2003). The prevention of child and adolescent violence: A review. *Aggression and Violent Behavior, 8,* 61–91.

Franko, D. L. (1998). Secondary prevention of eating disorders in college women at risk. *Eating Disorders, 6,* 29–40.

Goodrick, G. K., Poston, W. S., Kimball, K. T., Reeves, R. S., & Foreyt, J. P. (1998). Nondieting versus dieting treatments for overweight binge-eating women. *Journal of Consulting and Clinical Psychology, 66,* 363–368.

Heinberg, L. J., Thompson, J. K., & Matzon, J. L. (2001). Body image dissatisfaction as a motivator for healthy lifestyle change: Is some distress beneficial? In R. H. Striegel-Moore & L. Smolak (Eds.), *Eating disorders: Innovative directions in research and practice* (pp. 215–232). Washington, DC: American Psychological Association.

Jacobson, N., & Truax, P. (1991). Clinical significance: A statistical approach to defining change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59,* 12–19.

Jerome, L. W. (1987). *Primary intervention for bulimia: The evaluation of a media presentation for an adolescent population.* Unpublished master's thesis, Kent State University, Kent, OH.

Jerome, L. W. (1991). Primary intervention for bulimia: The evaluation of a media presentation for an adolescent population. *Dissertation Abstracts International, 52*(6), 3296B.

Johnson, J. G., Cohen, P., Kasen, S., & Brook, J. S. (2002). Eating disorders during adolescence and the risk for physical and mental disorders during early adulthood. *Archives of General Psychiatry, 59,* 545–552.

Kalichman, S. C., Carey, M. P., & Johnson, B. T. (1996). Prevention of sexually transmitted HIV infection: A meta-analytic review of the behavioral outcome literature. *Annals of Behavioral Medicine, 18,* 6–15.

Kaminski, P. L., & McNamara, K. (1996). A treatment for college women at risk for bulimia: A controlled evaluation. *Journal of Counseling and Development, 74,* 288–294.

Kater, K. J., Rohwer, J., & Londre, K. (2002). Evaluation of an upper elementary school program to prevent body image, eating, and weight concerns. *Journal of School Psychology, 72,* 199–204.

Keel, P. K., Fulkerson, J. A., & Leon, G. R. (1997). Disordered eating precursors in pre- and early adolescent girls and boys. *Journal of Youth and Adolescence, 26,* 203–216.

Killen, J. D., Taylor, C. B., Hammer, L., Litt, I., Wilson, D. M., Rich, T., et al. (1993). An attempt to modify unhealthful eating attitudes and weight regulation practices of young adolescent girls. *International Journal of Eating Disorders, 13,* 369–384.

Kraemer, H. C., Stice, E., Kazdin, A., & Kupfer, D. (2001). How do risk factors work? Mediators, moderators, independent, overlapping, and proxy risk factors. *American Journal of Psychiatry, 158,* 848–856.

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.

Kusel, A. B. (1999). Primary prevention of eating disorders through media literacy training of girls. *Dissertation Abstracts International, 60*(4), 1859B.

Larimer, M. E., & Cronce, J. M. (2002). Identification, prevention, and treatment: A review of individual-focused strategies to reduce problematic alcohol consumption by college students. *Journal of Studies on Alcohol, 14*(Suppl.), 148–163.

Levine, M. P., & Smolak, L. (2001). Primary prevention of body image disturbances and disordered eating in childhood and early adolescence. In J. K. Thompson & L. Smolak (Eds.), *Body image, eating disorders, and obesity in youth: Assessment, prevention and treatment* (pp. 237–260). Washington, DC: American Psychological Association.

Lewinsohn, P. M., Hops, H., Roberts, R. E., Seeley, J. R., & Andrews, J. A. (1993). Adolescent psychopathology: I. Prevalence and incidence of depression and other *DSM–II–R* disorders in high school students. *Journal of Abnormal Psychology, 102,* 133–144.

Lewinsohn, P. M., Striegel-Moore, R. H., & Seeley, J. R. (2000). Epidemiology and natural course of eating disorders in young women from adolescence to young adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 1284–1292.

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

Lowry-Webster, H. M., Barrett, P. M., & Dadds, M. R. (2001). A universal prevention trial of anxiety and depressive symptomatology in childhood: Preliminary data from an Australian study. *Behaviour Change, 18,* 36–50.

Maggs, J. L., Schulenberg, J., & Hurrelmann, K. (1996). Developmental transitions during adolescence: Health promotion implications. In J. Schulenberg, J. Maggs, & K. Hurrelmann (Eds.), *Health risks and developmental transitions during adolescence* (pp. 522–543). New York: Cambridge University Press.

Mann, T., Nolen-Hoeksema, S., Huang, K., Burgard, D., Wright, A., & Hanson, K. (1997). Are two interventions worse than none? Joint primary and secondary prevention of eating disorders in college females. *Health Psychology, 16,* 215–225.

Martz, D., & Bazzini, D. (1999). Eating disorders prevention programming may be failing: Evaluation of two one-shot programs. *Journal of College Student Development, 40,* 32–42.

McVey, G. L., & Davis, R. (2002). A program to promote positive body image: A 1-year follow-up assessment. *Journal of Early Adolescence, 22,* 96–108.

McVey, G. L., Davis, R., Tweed, S., & Shaw, B. F. (in press). Evaluation of a school-based program designed to improve body image satisfaction, global self-esteem, and eating attitudes and behaviours: A replication study. *International Journal of Eating Disorders.*

McVey, G. L., Lieberman, M., Voorberg, N., Wardrope, D., & Blackmore, E. (2003). School-based peer support groups: A new approach to the prevention of disordered eating. *Eating Disorders: Journal of Treatment and Prevention, 11,* 169–186.

Moreno, A. B., & Thelen, M. H. (1993). A preliminary prevention program for eating disorders in a junior high school population. *Journal of Youth and Adolescence, 22,* 109–124.

Murphy, J. G., Duchnick, J. J., Vuchinich, R. E., Davison, J. W., Karg, R. S., Olson, A. M., et al. (2001). Relative efficacy of a brief motivational intervention for college student drinkers. *Psychology of Addictive Behaviors, 15,* 373–379.

Mutterperl, J. A., & Sanderson, C. A. (2002). Mind over matter: Internalization of the thinness norm as a moderator of responsiveness to norm misperception education in college women. *Health Psychology, 21,* 519–523.

Nebel, M. A. (1995). Prevention of disordered eating among college women: A clinical intervention. *Dissertation Abstracts International, 56*(5), 2878B.

Neumark-Sztainer, D., Butler, R., & Palti, H. (1995). Eating disturbances among adolescent girls: Evaluation of a school-based primary prevention program. *Journal of Nutritional Education, 27,* 24–31.

Neumark-Sztainer, D., Sherwood, N. E., Coller, T., & Hannan, P. J. (2000). Primary prevention of disordered eating among preadolescent girls: Feasibility and short-term effect of a community-based intervention. *Journal of the American Dietetic Association, 100,* 1466–1473.

Newman, D. L., Moffitt, T. E., Caspi, A., Magdol, L., Silva, P. A., & Stanton, W. R. (1996). 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, 64,* 552–562.

Nicolino, J. C., Martz, D. M., & Curtin, L. (2001). Evaluation of a cognitive-behavioral therapy intervention to improve body image and decrease dieting in college women. *Eating Behaviors, 2,* 353–362.

O'Dea, J. A., & Abraham, S. (2000). Improving the body image, eating attitudes, and behaviors of young male and female adolescents: A new educational approach that focuses on self-esteem. *International Journal of Eating Disorders, 28,* 43–57.

Orjada, K., & Stice, E. (2003). *A controlled evaluation of an eating disturbance psychoeducational intervention for college students: A replication.* Manuscript in preparation.

Outwater, A. D. (1991). An intervention project to improve body image and self-esteem in 6th grade boys and girls as a potential prevention against eating disorders. *Dissertation Abstracts International, 51*(12), 4029A.

Patton, G. C., Selzer, R., Coffey, C., Carlin, J. B., & Wolfe, R. (1999). Onset of adolescent eating disorders: Population based cohort study over 3 years. *British Medical Journal, 318,* 765–768.

Paxton, S. J. (1993). A prevention program for disturbed eating and body dissatisfaction in adolescent girls: A 1 year follow-up. *Health Education Research, 8,* 43–51.

Presnell, K., & Stice, E. (2003). An experimental test of the effect of weight-loss dieting on bulimic pathology: Tipping the scales in a different direction. *Journal of Abnormal Psychology, 112,* 166–170.

Reeves, R. S., McPherson, R. S., Nichaman, M. Z., Harrist, R. B., Foreyt, J. P., & Goodrick, G. K. (2001). Nutrient intake of obese female binge eaters. *Journal of the American Dietetic Association, 101,* 209–215.

Richman, R. D. (1993). *Primary prevention of eating disorders: A pilot program.* Unpublished master's thesis, Simon Fraser University, Burnaby, British Columbia, Canada.

Richman, R. D. (1998). Preventing disordered eating: Promoting healthy attitudes and behaviors: A school-based program. *Dissertation Abstracts International, 52*(6), 0425B.

Rooney, B. L., & Murray, D. M. (1996). A meta-analysis of smoking prevention programs after adjustment for errors in the unit of analysis. *Health Education Quarterly, 23,* 48–64.

Rosen, J. C., Saltzberg, E., & Srebnik, D. (1989). Cognitive behavior therapy for negative body image. *Behavior Therapy, 20,* 393–404.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Thousand Oaks, CA: Sage.

Santonastaso, P., Zanetti, T., Ferrara, S., Olivotto, M. C., Magnavita, N., & Savaro, A. (1999). A preventive intervention program in adolescent school girls: A longitudinal study. *Psychotherapy and Psychosomatics, 68,* 46–50.

Shadish, W. R., & Haddock, C. K. (1994). Combining estimates of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 261–281). New York: Russell Sage Foundation.

Shepard, R. E. (2001). The body and soul program: Evaluation of a peer educator-led eating disorder education and prevention program. (Doctoral dissertation, University of Oregon, 2000). *Dissertation Abstracts International, 61*(7), 3862B.

Smolak, L., Levine, M., & Schermer, F. (1998a). A controlled evaluation of an elementary school primary prevention program for eating problems. *Journal of Psychosomatic Research, 44,* 339–353.

Smolak, L., Levine, M., & Schermer, F. (1998b). Lessons from lessons: An evaluation of an elementary school prevention program. In W. Vandereyken & G. Noordenbos (Eds.), *The prevention of eating disorders* (pp. 137–172). London: Athlone.

Stewart, D. A., Carter, J. C., Drinkwater, J., Hainsworth, J., & Fairburn, C. G. (2001). Modification of eating attitudes and behavior in adolescent girls: A controlled study. *International Journal of Eating Disorders, 29,* 107–118.

Stice, E. (2002). Risk and maintenance factors for eating pathology: A meta-analytic review. *Psychological Bulletin, 128,* 825–848.

Stice, E., Cameron, R., Killen, J. D., Hayward, C., & Taylor, C. B. (1999). Naturalistic weight reduction efforts prospectively predict growth in relative weight and onset of obesity among female adolescents. *Journal of Consulting and Clinical Psychology, 67,* 967–974.

Stice, E., Fisher, M., & Martinez, E. (in press). Eating Disorder Diagnostic

Scale: Additional evidence of reliability and validity. *Psychological Assessment.*

Stice, E., Hayward, C., Cameron, R., Killen, J. D., & Taylor, C. B. (2000). Body image and eating related factors predict onset of depression in female adolescents: A longitudinal study. *Journal of Abnormal Psychology, 109,* 438–444.

Stice, E., Killen, J. D., Hayward, C., & Taylor, C. B. (1998). Age of onset for binge eating and purging during adolescence: A 4-year survival analysis. *Journal of Abnormal Psychology, 107,* 671–675.

Stice, E., Mazotti, L., Weibel, D., & Agras, W. S. (2000). Dissonance prevention program decreases thin-ideal internalization, body dissatisfaction, dieting, negative affect, and bulimic symptoms: A preliminary experiment. *International Journal of Eating Disorders, 27,* 206–217.

Stice, E., Presnell, K., & Bearman, S. K. (2001). Relation of early menarche to depression, eating disorders, substance abuse, and comorbid psychopathology among adolescent girls. *Developmental Psychology, 37,* 608–619.

Stice, E., Presnell, K., Shaw, H., & Hoffman, E. (2003). *Dissonance and healthy weight control eating disorder prevention programs: Results from a randomized trial.* Manuscript in preparation.

Stice, E., & Ragan, J. (2002). A controlled evaluation of an eating disturbance psychoeducational intervention. *International Journal of Eating Disorders, 31,* 159–171.

Stice, E., Trost, A., & Chase, A. (2003). Healthy weight control and dissonance-based eating disorder prevention programs: Results from a controlled trial. *International Journal of Eating Disorders, 33,* 10–21.

Striegel-Moore, R. H., & Steiner-Adair, C. (1998). Primary prevention of eating disorders: Further considerations from a feminist perspective. In W. Vandereyken & G. Noordenbos (Eds.), *The prevention of eating disorders* (pp. 1–22). London: Athlone.

Stoolmiller, M., Eddy, J. M., & Reid, J. B. (2000). Detecting and describing preventive intervention effects in a universal school-based randomized trial targeting delinquent and violent behavior. *Journal of Consulting and Clinical Psychology, 68,* 296–306.

Telch, C. F., & Agras, W. S. (1993). The effects of a very low calorie diet on binge eating. *Behavior Therapy, 24,* 177–193.

Telch, C. F., Agras, W. S., & Linehan, M. M. (2001). Dialectical behavior therapy for binge eating disorder. *Journal of Consulting and Clinical Psychology, 69,* 1061–1065.

Tobler, N. S., Roona, M. R., Ochshorn, P., Marshall, D. G., Streke, A. V., & Stackpole, K. M. (2000). School-based adolescent drug prevention programs: 1998 meta-analysis. *Journal of Primary Prevention, 20,* 275–336.

U.S. Census Bureau. (2002). *Statistical abstracts of the United States: 2002* (122nd ed.). Washington, DC: Author.

Varnado-Sullivan, P. J., Zucker, N., Williamson, D. A., Reas, D., Thaw, J., & Netemeyer, S. B. (2001). Development and implementation of the Body Logic Program for adolescents: A two-stage prevention program for eating disorders. *Cognitive and Behavioral Practice, 8,* 248–259.

Wadden, T. A., Foster, G. D., & Letizia, K. A. (1994). One-year behavioral treatment of obesity: Comparison of moderate and severe caloric restriction and the effects of weight maintenance therapy. *Journal of Consulting and Clinical Psychology, 62,* 165–171.

Wade, T. D., Davidson, S., & O'Dea, J. A. (2003). A preliminary controlled evaluation of a school-based media literacy program and self-esteem program for reducing eating disorder risk factors. *International Journal of Eating Disorders, 33,* 371–383.

Weiss, K. R. (2000). *An evaluation of a primary prevention program for disordered eating in adolescent girls: Examining responses of high- and low-risk girls.* Unpublished doctoral dissertation, La Trobe University, Bundoora, Victoria, Australia.

Wilfley, D. E., Welch, R. R., Stein, R. I., Spurrell, E. B., Cohen, L. R., Saelens, B. E., et al. (2002). A randomized comparison of group cognitive-behavioral therapy and group interpersonal psychotherapy for the treatment of overweight individuals with binge eating disorder. *Archives of General Psychiatry, 59,* 713–721.

Wilson, G. T., Fairburn, C. G., Agras, W. S., Walsh, B. T., & Kraemer, H. K. (2002). Cognitive–behavioral therapy for bulimia nervosa: Time course and mechanisms of change. *Journal of Consulting and Clinical Psychology, 70,* 267–274.

Winzelberg, A. J., Eppstein, D., Eldredge, K. L., Wilfley, D., Dasmahapatra, R., Dev, P., & Taylor, C. B. (2000). Effectiveness of an Internet-based program for reducing risk factors for eating disorders. *Journal of Consulting and Clinical Psychology, 68,* 346–350.

Winzelberg, A. J., Taylor, C. B., Sharpe, T., Eldredge, K. L., Dev, P., & Constantinou, P. S. (1998). Evaluation of a computer-mediated eating disorder intervention program. *International Journal of Eating Disorders, 24,* 339–349.

Withers, G. F., Twigg, K., Wertheim, E. H., & Paxton, S. J. (2002). A controlled evaluation of an eating disorders primary prevention videotape using the elaboration likelihood model of persuasion. *Journal of Psychosomatic Research, 53,* 1021–1027.

Wolf-Bloom, M. S. (1999). Using media literacy training to prevent body dissatisfaction and subsequent eating problems in early adolescent girls. *Dissertation Abstracts International, 59*(8), 4515B.

Zabinski, K. J., Calfas, K. J., Gehrman, C. A., Wilfley, D. E., & Sallis, J. F. (2001). Effects of a physical activity intervention on body image in university seniors: Project GRAD. *Annals of Behavioral Medicine, 23,* 247–252.

Received January 6, 2003
Revision received September 3, 2003
Accepted September 9, 2003 ■