A Meta-Analytic Review of Obesity Prevention Programs for Children and Adolescents: The Skinny on Interventions That Work

Eric Stice, Heather Shaw, and C. Nathan Marti
University of Texas at Austin

This meta-analytic review summarizes obesity prevention programs and their effects and investigates participant, intervention, delivery, and design features associated with larger effects. A literature search identified 64 prevention programs seeking to produce weight gain prevention effects, of which 21% produced significant prevention effects that were typically pre- to post effects. Larger effects emerged for programs that targeted children and adolescents (vs. preadolescents) and females, programs that were relatively brief, programs that solely targeted weight control versus other health behaviors (e.g., smoking), programs evaluated in pilot trials, and programs wherein participants must have self-selected into the intervention. Other factors, including mandated improvements in diet and exercise, sedentary behavior reduction, delivery by trained interventionists, and parental involvement, were not associated with significantly larger effects.

Keywords: obesity, prevention, meta-analysis, moderators

Obesity in adulthood results in an increased risk for future death from all causes, coronary heart disease, atherosclerotic cerebrovascular disease, and colorectal cancer, as well as serious medical problems including hyperlipidemia, hypertension, gallbladder disease, and diabetes mellitus (Calle, Thun, Petrelli, Rodriguez, & Heath, 1999). Obesity in childhood and adolescence has also been associated with serious medical problems, including high blood pressure, adverse lipoprotein profiles, diabetes mellitus, atherosclerotic cerebrovascular disease, coronary heart disease, colorectal cancer, and death from all causes, as well as lower educational attainment and poverty (Dietz, 1998). The prevalence of obesity has increased sharply over the last 3 decades; currently, 65% of adults are classified as overweight or obese (Hedley et al., 2004). The prevalence of obesity has risen even more sharply among adolescents and young adults (Hedley et al., 2004), which is alarming because obesity persists into adulthood for 70% of obese adolescents (Magarey, Daniels, Boulton, & Cockington, 2003). Obesity also carries a high fiscal cost; roughly $100 billion per year is spent on obesity-related health care (Wolf, 1998).

Unfortunately, successful treatments for obesity have been elusive. For adults, the current treatment of choice only results in about a 10% reduction in body weight, and virtually all patients regain this weight within a few years of treatment (Jeffery et al., 2000). Obesity treatments for children and adolescents have

yielded similar effects, though behavioral family-based interventions have produced more persistent weight loss effects (Epstein, Valoski, Wing, & McCurley, 1990; Flodmark, Ohlsson, Ryden, & Sveger, 1993). Compounding matters, only about 10% of obese children and adolescents seek weight loss treatment (e.g., French, Perry, Leon, & Fulkerson, 1994). Accordingly, much effort has been devoted to developing and evaluating obesity prevention programs, in the hope that this strategy will more effectively curb this pernicious public health problem.

Studies have evaluated four major types of interventions that were expected to produce weight gain prevention effects. These include (a) multifocus cardiovascular disease prevention programs that targeted obesity along with other risk factors for cardiovascular disease (e.g., hypertension and smoking), (b) prevention programs that focused solely on the prevention of obesity or weight gain, (c) interventions designed to solely increase physical activity, and (d) eating disorder prevention programs that promoted use of healthy weight-management skills.

Although numerous evaluations of weight gain prevention programs have been conducted, their results have not been comprehensively reviewed and analyzed with meta-analytic procedures. Several excellent narrative reviews exist (e.g., Dietz & Gortmaker, 2001; Schmitz & Jeffery, 2000; Story, 1999), but meta-analytic techniques were not used to empirically describe effect sizes or investigate potential moderators of intervention effects. One meta-analytic review has been published (Campbell, Waters, O'Meara, Kelly, & Summerbell, 2003), but it used extensive exclusionary criteria that rendered it impossible to examine moderators of intervention effects (only 10 trials were included). Thus, the overarching goal of this article is to address this important gap in the literature. The first aim of this review is to provide a summary of these prevention programs and their effects. The second aim is to examine participant, intervention, delivery, and design features that are associated with larger intervention effects. Given the heterogeneity in the effects from these interventions, it is important

to systematically consider the moderators associated with intervention that produced the largest effects. The third aim is to discuss promising directions for future research in light of the findings from completed trials.

Putative Moderators of Intervention Effects

A unique feature of meta-analyses is that they permit empirical examination of factors associated with variation in effect sizes. Elucidating factors that moderate prevention program effects is informative because it highlights aspects of the participants, intervention, program delivery, and research design that are associated with stronger intervention effects. This information should increase the yield of future prevention efforts by identifying the conditions under which optimal prevention effects occur. As well, this information might identify particular subgroups of individuals for whom alternative obesity prevention programs need to be developed. Analyses of moderators of intervention effects 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 previous literature reviews.

Participant Features

Participant Age

Researchers have hypothesized that obesity prevention programs are more effective when they are delivered to middle school or high school students versus grade school students (Baranowski, Cullen, Nicklas, Thompson, & Baranowski, 2002). Younger children may find it difficult to grasp the concepts and skills taught in the interventions. They may also be less likely to impact the food purchases made by adults (when eating at home or restaurants). Thus, we hypothesized that effects would be significantly larger for interventions offered to adolescents versus children.

Participant Gender

Results from prior trials suggest that obesity prevention programs that promoted a healthier lower calorie diet (Perry et al., 1998) and those that also attempted to increase physical activity and/or decrease sedentary behavior (Gortmaker et al., 1999; Vandongen et al., 1995) produced larger effects for females than for males. However, another obesity prevention program that promoted healthy lower calorie diets and increased physical activity found significantly stronger effects for males than for females (Kain, Uauy, Vio, Cerda, & Leyton, 2004), and one obesity treatment trial found that an intervention solely aimed at increasing activity and decreasing sedentary behaviors was more effective for boys than girls, though an intervention focusing solely on increasing activity level was equally effective for boys and girls (Epstein, Paluch, & Raynor, 2001). Although these findings may represent chance findings because most trials did not report that intervention effects were moderated by gender, there was more evidence that obesity prevention programs produced larger effects for females than for males. This finding may have emerged because sociocultural pressures for thinness are greater for females (Thompson, Heinberg, Altabe, & Tantleff-Dunn, 1999), which may amplify the

effects of obesity prevention programs for this population. In support, more females than males are dissatisfied with their bodies, and the vast majority of adolescent females with body image concerns are dissatisfied because they feel overweight (Thompson et al., 1999). In contrast, the reasons males give for body image concerns are more heterogeneous, and nearly half who indicate that they are dissatisfied with their weight actually wish to gain weight. In addition, females are at higher risk for onset of obesity than males (Solomon & Manson, 1997). Because there was more evidence that obesity prevention programs produce larger effects for females than males, we hypothesized that intervention effects for prevention programs might be larger for females.

Participant Ethnicity

There is also reason to believe that ethnicity might moderate obesity prevention effects. On the one hand, there is evidence that Black and Hispanic individuals show elevated rates of overweight and obesity as well as greater increases in weight over development, relative to other ethnic groups (e.g., Burke & Bild, 1996; Kimm et al., 2001), suggesting that programs targeting these high-risk groups might be more effective because there is a greater opportunity to show a prevention effect. On the other hand, overweight and obesity are less stigmatized and are associated with less body dissatisfaction for certain ethnic minority groups (e.g., Duncan, Anton, Newton, & Perri, 2003), particularly Black women, which might attenuate the effects of obesity prevention programs for these populations. Thus, we hypothesized that intervention effects would be different for programs primarily targeting high-risk ethnic minority participants versus those primarily targeting low-risk ethnic groups.

Risk Status of Participants

More generally, we have hypothesized (Stice & Shaw, 2004) that interventions are more effective when offered to high-risk participants (selected prevention programs) versus all individuals in a population (universal prevention programs). In the obesity prevention field, selected interventions have been directed at a variety of groups at elevated risk for future weight gain, including Black and Hispanic individuals, students with other cardiovascular disease risk factors (e.g., hypertension), overweight or obese individuals, 1st-year college students, and females with body dissatisfaction. Theoretically, these high-risk individuals are more motivated to engage in the prevention program content and thus are more likely to benefit. It is also likely that low-risk individuals have less room for change on the outcomes (a floor effect). One narrative review comparing selected and universal school-wide obesity prevention programs concluded that selected interventions may be more effective in reducing pediatric obesity than universal interventions (Resnicow, 1993). In addition, prevention programs for eating pathology (Killen et al., 1993), depression (Clarke et al., 1995), anxiety (Lowry-Webster, Barrett, & Dadds, 2001), behavior problems (Stoolmiller, Eddy, & Reid, 2000), and substance abuse (Murphy et al., 2001) have often produced stronger effects for high-risk subsamples than for the full sample of individuals enrolled in these universal prevention programs. Thus, we hypothesized that intervention effects would be larger for selected programs versus universal 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 moderator.

Intervention Features

Intervention Duration

Previous meta-analyses of prevention programs for other problem behaviors have suggested that longer duration multisession interventions produced more superior effects than very brief interventions (Rooney & Murray, 1996; Stice & Shaw, 2004). Theoretically, interventions with a longer duration afford a greater opportunity for presentation of information and behavioral change skills. We hypothesized that intervention effects would be stronger for prevention programs with a longer versus shorter duration.

Parental Involvement

It has also been suggested that parental involvement leads to more favorable results in obesity prevention, as the family is thought to be key to developing a psychosocial environment that is conducive to healthy eating and physical activity (Story, 1999). Parents are usually responsible for determining food offerings in and away from the home, at least through a certain age, as well as influencing exercise and recreation. Obesity treatment trials have suggested that both child and adolescent weight loss programs are more effective when at least one parent is involved (Epstein, Wing, Koeske, & Valoski, 1987; Golan, Weizman, Apter, & Fainaru, 1998). Therefore, we hypothesized that obesity prevention programs with parental involvement would have larger effects than those without parental involvement.

Psychoeducational Content

Because research has suggested that psychoeducational content is ineffective in producing behavioral change (Helweg-Larsen & Collins, 1997; Larimer & Cronce, 2002), we hypothesized that psychoeducational programs would be associated with weaker intervention effects. Indirect support for this hypothesis was provided by a recent meta-analysis which found that eating disorder prevention programs with psychoeducational content are less effective than those without this content (Stice & Shaw, 2004).

Dietary Improvement

One implication from the energy balance model of obesity is that a reduction in fat and sugar intake and an increase in fruit and vegetable intake will decrease the risk for future weight gain (Epstein, Gordy, et al., 2001). Although virtually all obesity prevention programs recommend consumption of low-fat diets, we differentiated between programs that directly manipulated dietary change as part of the intervention and those that did not. We reasoned that distinguishing between interventions that actually manipulated diet and those that did not would provide the most sensitive test of this moderator. Another benefit is that this coding scheme captures environmental manipulations of the food environment, which is useful because theorists have suggested that the food environment plays a key role in obesity promotion (Wadden, Brownell, & Foster, 2002). The most common example was inter-

ventions that directly changed the nutritional content of school lunches (e.g., Donnelly et al., 1996; Luepker et al., 1996). We hypothesized that interventions that involved a direct improvement to dietary intake should produce stronger intervention effects than those that did not.

Increased Activity

Another implication from the energy balance model of obesity is that increased physical activity will decrease risk for future weight gain (Wadden, Vogt, Foster, & Anderson, 1998). Although most obesity prevention programs recommend regular physical activity, we distinguish between prevention programs that directly manipulated physical activity from those that simply recommended it, because we felt this would provide a more sensitive test of this potential moderator. The most common example of programs that manipulated physical activity was school-based interventions that administered a physical education class for students in the intervention condition but not the control condition (e.g., Dwyer, Coonan, Leitch, Hetzel, & Baghurst, 1983; McMurray et al., 2002). We hypothesized that programs that directly increased physical activity would have larger intervention effects than those that did not increase activity.

Reduced Sedentary Behavior

A third implication of the energy balance model of obesity is that interventions that reduce sedentary behavior, such as TV viewing and video game use, should also decrease risk for future weight gain. Indeed, it has been theorized that more effective obesity prevention programs focused on reducing sedentary behavior (Baranowski et al., 2002), and TV viewing is considered one of the most modifiable causes of obesity in children (Robinson, 1999). We hypothesized that larger effects would emerge for programs that focused on reducing sedentary behaviors than for programs that did not target this risk factor.

Number of Behavior Targets

Our review of the literature suggested that the number of health behaviors targeted in an intervention was inversely related to the magnitude of intervention effects for obesity. Specifically, it appeared that interventions that attempted to change a broad array of health behaviors, such as body weight, blood pressure, cholesterol, and smoking, were less effective than programs that focused solely on body weight. Our clinical experience from designing and evaluating prevention programs also suggests that interventions focusing on a few concepts are more effective than those focusing on a broader array of concepts. It may be that the greater the complexity of the message relayed by the intervention, the more difficult it is for participants to process, store, and retrieve information presented in the programs. Consistent with this general impression, a review of school-based cardiovascular disease prevention trials concluded that broad-based programs targeting multiple health behaviors aimed at reducing risks for cardiovascular disease have not been effective for reducing obesity in children (Resnicow & Robinson, 1997). We hypothesized that programs targeting multiple health behaviors would have smaller effects than those solely targeting weight change.

Delivery Features

Teachers Versus Professional Interventionists

Researchers have suggested that obesity prevention programs are more effective when delivered by dedicated interventionists versus classroom teachers (Baranowski et al., 2002). Theoretically, teachers are not able to devote as much time and energy to providing interventions as dedicated interventionists because teachers have classroom responsibilities that take precedence. Moreover, dedicated interventionists are typically able to provide the intervention several times per school year, allowing them to develop and refine their presentation strategies, whereas teachers typically will only provide the intervention once per year. In addition, teachers rarely receive the amount of specialized training and detailed supervision provided to dedicated interventionists. Thus, we hypothesized that intervention effects would be significantly larger for programs delivered by dedicated interventionists versus classroom teachers.

Didactic Versus Interactive Format

Meta-analytic reviews of substance abuse (Tobler et al., 2000) and eating disorder (Stice & Shaw, 2004) prevention programs have found that interactive programs produced larger intervention effects than didactic programs. Theoretically, participants in interactive programs show greater intervention effects because this format helps participants engage 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 enhance skill acquisition (e.g., particular sports). We predicted that interactive programs would be more effective than didactic programs.

Design Features

Pilot Study

Our review of the prevention and treatment literature for obesity and eating disorders suggested that larger intervention effects were often observed for pilot trials of a new intervention relative to large demonstration trials. Such a pattern of effects might occur because interventionists are more passionate about new prevention programs or because demonstration trials are more methodologically rigorous and are therefore more immune to experimenter effects (e.g., because they more often use blinded assessors and minimal intervention control conditions). Thus, we hypothesized that intervention effects would be significantly larger for pilot evaluations of new interventions.

Recruitment Method

Our experience suggests that intervention effects are often larger when prevention programs are delivered solely to participants who have actively self-selected into trials in response to recruitment efforts, such as media advertisements, relative to when prevention programs are offered to all individuals in a defined population (e.g., a particular school). Presumably this is because the former strategy recruits individuals who are more motivated to achieve weight gain prevention effects and therefore engage more effectively in the prevention program. Thus, we hypothesized that intervention effects would be significantly larger for self-presenting volunteers than for participants recruited through population-based recruitment efforts.

Random Assignment

We theorized that trials that randomly assigned participants to condition might produce larger intervention effects than trials that used alternative approaches to allocating participants to treatment condition, such as matching. We reasoned that because random assignment is the best approach to generating groups that are equivalent on any potential confounding variables at baseline (with sufficiently large sample sizes), it should therefore minimize the chances that any of these confounding variables are correlated with treatment condition, which should thus maximize the ability to detect intervention effects if they really occur (i.e., randomization maximizes the signal-to-noise ratio reflected in inferential tests of the intervention effects). Accordingly, we hypothesized that intervention effects may be greater for trials that used random assignment relative to other approaches to assigning participants to condition. However, because the proper analysis of intervention effects involves tests of differential change across conditions, which adjusts for any initial differences at baseline on the outcome, we suspected that this effect might not reach statistical significance. Consistent with this expectation, random assignment did not emerge as a significant moderator of effects sizes in our meta-analysis of eating disorder prevention programs (Stice & Shaw, 2004).

Nested Data Modeled Incorrectly

Virtually all parametric inferential tests, such as repeated measures analysis of variance, growth curve, and survival models, used to test for intervention effects within randomized trials assume independence of errors. However, when participants are nested within schools, classes, or group-based interventions, the assumption of independence may not hold (Baldwin, Murray, & Shadish, 2005). Participants within these nested groups may be more similar than participants from across these groups, which can artificially reduce the error terms used to test for intervention effects, which increases risk for a false positive finding. Thus, we hypothesized that studies that did not model the nested nature of the data in the trial would produce artificially larger effect sizes for the interventions relative to studies that modeled the nested nature of the data.

Potential Artifacts

We also investigated three variables that might produce artifacts for the effect sizes and bias our estimates of effect size moderators, with the goal of including these variables as covariates in the models if necessary. First, our review of the eating disorder prevention field suggested that interventions tend to produce larger effect sizes when they are compared with assessment-only or waitlist control conditions relative to when they are compared with active interventions that are credible and structurally matched to the intervention in terms of contact hours (Stice & Shaw, 2004). Theoretically, this pattern of findings occurs because the active

comparison groups more effectively control for demand characteristics, participant expectances, and other nonspecific factors that contribute to intervention effects. Thus, we tested whether type of control condition was systematically related to the intervention effect sizes. Second, because effect sizes for prevention programs tend to be smaller when longer follow-up periods are examined relative to shorter follow-up periods or pretest to posttest designs for prevention programs (Stice & Shaw, 2004), we tested whether follow-up length was related to effect size magnitude. Third, because prior meta-analyses have found that unpublished studies often have smaller effects than published studies (Lipsey & Wilson, 2001), we investigated whether publication status was related to intervention effect sizes.

Method

Sample of Studies

Following the recommendations of Lipsey and Wilson (2001), we used five procedures to retrieve published and unpublished trials of obesity prevention programs. First, a computer search was performed on PsycINFO, MEDLINE, *Dissertation Abstracts International*, and Cumulative Index to Nursing and Allied Health Literature for the years 1980–2005 (through October) with the following keywords: *obesity weight*, *cardiovascular disease*, *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 and Heather Shaw reviewed the products of all three searches to identify pertinent articles. Second, the tables of content for journals that commonly publish articles in this area were reviewed for this same period (e.g., *Preventive Medicine, Journal of Pediatrics, Health Education Quarterly*). Third, we consulted narrative reviews of the obesity prevention field to search for additional citations of relevance. Fourth, the reference sections of all identified articles were examined. Finally, established obesity 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 obesity prevention program is that it results in significantly less weight gain or risk for obesity onset than observed in the control group. Thus, we only included trials that used some type of proxy measure of body fat as an outcome. Most trials used the body mass index (BMI = Kg/M^2) as the primary proxy measure of body fat, but a few studies, particularly older ones, used skinfold thickness. It is important to note that BMI is not a direct measure of body fat. Although this proxy measure tends to show high correlations with the most precise measures of body fat ($r = .80$–$.90$), such as dual energy x-ray absorptiometry (DEXA; Dietz & Robinson, 1998), it has been found to show lower agreement with DEXA measures in large epidemiology samples ($r = .71$; Ellis, Abrams, & Wong, 1999). Nonetheless, because the BMI is easy to measure, shows high test–retest reliability, is inexpensive, and correlates with health risk markers and diseases, such as elevated blood pressure, adverse lipoprotein profiles, atherosclerotic lesions, serum insulin levels, and diabetes mellitus, it is considered the measurement of choice for large-scale studies (Dietz & Robinson, 1998; Freedman & Perry, 2000).

As noted previously, we included trials that were primarily conceptualized as evaluations of obesity prevention programs, as well as trials that evaluated other interventions that were expected to result in less weight gain or risk for obesity onset but that were not primarily conceptualized as obesity prevention programs (e.g., certain physical activity interventions, eating disorder prevention programs, and psychoeducational interventions). A prior meta-analysis indicated that certain eating disorder prevention programs and psychoeducational interventions produced significant weight gain prevention effects (Stice & Shaw, 2004). We included a wide variety of interventions that were expected to produce weight gain prevention effects in the hope that it would maximize our chances of identifying participant, intervention, delivery, and design features that are associated with the most efficacious obesity prevention programs. If multiple reports of the same trial were published, we selected the one with the longest follow-up period.

This meta-analysis focused solely on effect sizes for weight gain prevention effects, as assessed by differential change in body fat measures. We did not include effect sizes for changes in self-reported dietary intake or physical activity, because numerous trials have found significant intervention effects for self-reported dietary intake and physical activity, but no significant effects for weight change (e.g., Baranowski et al., 2003; Luepker et al., 1996; Puska et al., 1982). According to the energy balance model of adiposity, any true reduction in caloric intake and/or increase in physical expenditure should be accompanied by concomitant changes in body mass. Therefore, we interpreted this pattern of findings as suggesting that self-report measures of dietary intake and physical activity are of questionable validity, at least within the context of the demand characteristics of obesity prevention trials. This interpretation dovetails with studies that have found that people underreport caloric intake and overreport activity level (Bandini, Schoeller, Dyr, & Dietz, 1990; Lichtman et al., 1992).

We focused exclusively on prevention programs that were evaluated in controlled trials. We included trials in which participants were randomly assigned to an intervention; to active interventions that were not focused on weight gain prevention (e.g., a general parent training intervention); or to usual-programming (e.g., standard physical education classes), waitlist, or assessment-only control conditions. We also included 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 at baseline (Shadish, Cook, & Campbell, 2002). Because many confounds are unknown, random assignment is preferable to the use of control groups that are matched to the intervention group on preselected 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 (i.e., controlled for initial between-group differences on the outcome; Shadish et al., 2002). We excluded trials that compared only active interventions, because it seemed inappropriate to compare them with trials that used a control condition and because it is difficult to determine whether a lack of differential change across active interventions signifies that both prevention programs were effective or that neither was effective.

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 × Condition interaction in a repeated-measures analysis of variance model, an analysis of covariance model that controlled for initial levels of the outcome variable, or a growth curve model that controlled for initial levels of the outcome (e.g., the effects were conditional upon the intercept value of the dependent variable coded to reflect the level of the outcome at baseline; Stice & Shaw, 2004). It is necessary to control for initial levels of the outcome variable because otherwise the analyses are not providing a test of differential change over time across conditions. Verifying that the groups do not differ at baseline on the outcome variable does not solve this problem because the objective is to model change from baseline to intervention termination or follow-up, rather than just to conduct between-subjects tests of the groups at termination or follow-up. If the intervention group had higher initial BMI scores than the control group, the analyses may not detect a true intervention effect (a Type II error), whereas if the control group had higher initial BMI scores than the intervention group, the analyses might erroneously suggest that an intervention effect was present when it was not (a Type I error). We also included trials that used logistic

regression or survival models to test whether the rates of onset of obesity or overweight were significantly less in the intervention condition versus the control condition if initially obese or overweight participants, respectively, were excluded from the analyses (Willett & Singer, 1993). Studies that only tested for significant changes within condition were not included because this type of analysis does not test whether the changes in the intervention condition are significantly greater than the changes in the control condition. With this latter approach, there is no way to separate the effects of the intervention from those of alternative sources, such as regression to the mean or measurement artifacts.

We excluded trials that were described as obesity treatment programs by the authors because the purpose of the present report was to provide a meta-analytic review of programs that sought to prevent future weight gain or obesity onset. Nonetheless, we included evaluations of programs that sought to prevent future weight gain in overweight or obese samples if they were not referred to as treatment programs by the authors. More generally, we did not exclude studies solely because the average BMI of participants fell above conventional cutoffs for overweight or obese (e.g., over 25 or 30 for young adult samples).

We also restricted our focus to trials that targeted children and adolescents because of our interest in determining whether effective interventions have been designed for developing individuals. We believe that obesity prevention programs should be implemented before most individuals will show onset of obesity. However, we used a broad view of adolescence and included trials with a mean participant age of up to 22 years because this captured college-based obesity prevention programs. College-aged individuals are still developing self-regulation skills, particularly with regard to dietary and exercise behaviors. In addition, many developmental psychologists consider adolescence to span from approximately age 12 through age 24 because most individuals in the United States have not settled into adult roles by their early 20s (Arnett, 2000).

Effect Size Estimation Procedures

We calculated effect sizes for tests of differential change in BMI and risk for obesity onset across the intervention and control conditions because virtually all of the prevention trials included BMI as a primary outcome. Although other proxy measures of adiposity were used in several trials, such as skinfold thickness and waist-to-hip ratios, these latter outcomes were operationalized inconsistently and were collected in only a subset of the trials. We considered averaging the effect sizes from these various adiposity proxy measures, but we noted that the intervention effects for these various outcomes were often contradictory and were concerned that averaging across diverse measures would introduce unnecessary error variance into the analyses. Furthermore, the measurement error is considerably lower for the BMI relative to alternative proxy body fat measures, including waist circumference, triceps skinfold, and subscapular skinfold measures (Freedman & Perry, 2000). In the four studies that did not collect BMI data, effect sizes were calculated for alternative proxy measures of body fat; Dwyer et al. (1983) used skin-fold measures, Eliakim, Makowski, Brasel, and Cooper (2000) used MRI estimations of percent body fat, and Gutin et al. (1995) and Gutin and Owens (1999) used DEXA estimations of percent body fat.

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; Rosenthal, 1991). Furthermore, this effect size preserved the valence of the effects (unlike measures such as eta squared). Cohen's (1988) criteria for small ($r = .10$), medium ($r = .30$), and large ($r = .50$) effects were used.[1]

If effect sizes were reported in Cohen's (1988) d, we converted them to r with the formula provided on page 20 of Rosenthal (1991). If effects were reported as odds ratios, they were converted to r with the formula provided on page 194 of Lipsey and Wilson (2001). If no effect sizes were reported,

we generated them directly by calculating Cohen's d with the means and standard deviations (from the control group at baseline) reported in the article, which we then converted to r using the Rosenthal formula, or we reconstituted the data using weighted probability values to estimate a chi-square test that provided an odds ratio, which we then converted to r using the Lipsey and Wilson formula. If none of these options for generating effect sizes was possible, we estimated effect sizes from the exact p values reported by the authors using the formula provided on page 19 of Rosenthal (1991). If exact p values were not reported, we generated them from the test statistics (e.g., F) and degrees of freedom using Microsoft Excel.

We were able to use the methods described previously to generate effect sizes or estimates of effect sizes for all trials that reported significant intervention effects and for most trials that reported nonsignificant effects. However, for the two trials that reported nonsignificant effects and did not provide any other data with which to estimate the effect size (Fardy et al., 1996; Willet, 1995), we used full information maximum likelihood estimation to impute the missing effect sizes because this approach produces more accurate and efficient parameter estimates than listwise deletion or alternative imputation approaches such as mean substitution (Schafer & Graham, 2002). We selected this approach over the more common strategy of assuming an effect size of zero (Lipsey & Wilson, 2001) because more precise estimates of these missing values can be generated using the conditional probabilities between effect sizes and effect size moderators from the trials that provided complete data on these variables.

Operationalization and Coding of Effect Size Moderators

Table 1 lists the numeric values used to code each moderator, the operationalization of each moderator, and relevant descriptive statistics describing the distribution of the moderators.[2] We coded certain moderators two ways in an effort to ensure that we were not missing the effects of a moderator, because we did not operationalize it optimally. First, in addition to coding the average age of participants in the study, we also coded the age range of participants, to determine whether studies focusing on a narrow age range may be better able to deliver an intervention that is developmentally appropriate. Second, with regard to participant ethnicity, we coded both the percentage of participants who were Black or Hispanic (a continuous variable), because these two groups are at particularly high risk for obesity, and the dominant ethnic group represented in the samples (a nominal variable). Third, with regard to intervention duration, we coded both the total amount of intervention hours and the total length of the intervention in weeks because these two aspects of intervention duration varied somewhat independently (the r between these two dimensions was only .50). Fourth, with regard to psychoeducational content, we coded both whether each intervention contained psychoeducational content (to stay parallel with the coding used for the other intervention content codes) and whether the intervention included only psychoeducational content, to explore the possibility that these latter types of interventions were uniquely associated with small intervention effects.

---

[1] We did not focus on effect sizes, such as Cohen's (1988) d, which focus on posttest mean differences across conditions without correcting for pretest mean differences. Such effect size estimates are not able to rule out the possibility that differences at baseline between the conditions, even if nonsignificant, artificially amplified or attenuated effect size estimates. This theoretically has the effect of introducing greater error variance in effect size estimates and therefore decreases power in analyses testing heterogeneity of treatment effects and moderators of treatment effects.

[2] It might be noted that only 55% of the trials that did not use random assignment to condition used matching to create the groups, suggesting that the variable reflecting random assignment was not simply a surrogate for matching, which would have complicated the interpretation of the former moderator.

One aspect of our coding system was constrained by the distribution of a certain moderator across studies. Specifically, although we were interested in testing whether the intervention effects were significantly larger for females than males, only 33% of the trials that we located reported effect sizes separately for the sexes (and only 21% provided a direct test of whether sex moderated the intervention effects). Accordingly, we tested whether interventions offered solely to females were more effective than those offered solely to males or those offered to both sexes. We took this approach because (a) this variable emerged as a significant predictor of eating disorder prevention program effects (Stice & Shaw, 2004), (b) our initial review of the findings suggested that effects were larger for female-only trials, and (c) this allowed us to include all trials in the analyses. Because only two interventions were offered solely to males, we did not feel comfortable estimating an average effect for these two trials.

There were also a number of other potential moderators that we were unable to code because insufficient information was provided in the articles and reports. We were unable to code average attendance because only 44% of the studies reported this variable. We were unable to code the socioeconomic status of the sample because parallel information (e.g., average parental income) was reported in only 35% of studies. We were unable to code the method of handling missing data (e.g., listwise deletion [completer analysis], last observation carried forward, full information maximum likelihood estimation imputation) because less than 40% of the studies reported this information.

We used a consensus approach to coding the effect size moderators. Eric Stice and Heather Shaw were each responsible for coding certain moderators but consulted with each other when questions regarding the coding of particular studies arose. Although this approach allowed for a refinement of the coding system and served to increase interrater agreement, we did not use the consensus approach on all data points or double code all studies. Thus, we examined intercoder agreement by having Eric Stice and Heather Shaw code all of the moderators for a randomly selected 30% of the trials examined in this meta-analytic review.

Results

Descriptive Statistics

The literature search identified 46 trials that met the inclusion criteria, in which 61 different obesity prevention programs were evaluated (12 trials evaluated more than 1 prevention program, and 3 prevention programs were evaluated in 2 trials), resulting in a total of 64 effect sizes for this review. Of these 64 prevention programs, 30 were universal, and 34 were selected. The majority focused on both males and females ($n = 48$), but 14 focused solely on females, and 2 focused solely on males. The majority of these interventions were school-based programs (84%). A total of 51 of the 64 prevention programs used random assignment to condition, of which 13% were randomized at the participant level, 2% were randomized at the group level, and 85% were randomized at the school level. Brief descriptions of the samples, program content, and intervention effects are provided in Tables 2 and 3 for universal and selected prevention programs, respectively. Figure 1 provides a flowchart showing the number of studies that were omitted because of the various exclusionary criteria.

To assess interrater agreement between the two coders responsible for abstracting effect sizes and moderators, we calculated the interclass correlation coefficient for continuous variables and kappa (κ) coefficients for nominal variables (see Table 4). The interclass correlation coefficients ranged from a low of .95 (for the effect size estimates) to 1.0 (for 80% of the continuous variables examined in this report). The κ coefficients ranged from .87 (for

whether nested data was modeled incorrectly) to 1.00 (for 75% of the nominal variables examined in this report). These analyses indicate that there was high interrater agreement.

Tables 5 and 6 report the magnitude of effect sizes and provide the participant, intervention, delivery, and design features that were investigated as potential moderators of intervention effects. The effect sizes reflect analyses performed on the entire samples used in these studies, versus effect sizes for various subgroups such as the different genders, because such subgroup analyses were not consistently reported across trials.

Average Effect Size and Effect Size Heterogeneity

Analyses were conducted on the effect size for change in BMI in the intervention condition versus the control condition. We first converted Pearson's rs to z scores to avoid problematic standard error estimates (Hedges & Olkin, 1985). We then used the SPSS macro developed by Lipsey and Wilson (2001) to estimate the overall inverse variance weighted average effect size for random effects models. All mean values were computed with this method.

The average effect size across all studies was very small ($r = .04$) but was significantly larger than zero ($z = 2.94$, $p < .01$). The rs for the effect sizes ranged from $-.24$ to $.50$. Only 13 of these interventions (1 of which was evaluated in two trials), or 21% of the 61 programs evaluated, found significant positive intervention effects based on an alpha level of .05 (Dwyer et al., 1983; Eliakim et al., 2000; Fitzgibbon et al., 2004; Gutin & Owens, 1999; Killen et al., 1988; Lionis et al., 1991; Manios, Moschandreas, Hatzis, & Kafatos, 2002; Robinson, 1999; Stice, Orjada, & Tristan, 2006; Stice & Ragan, 2002; Stice, Shaw, Burton, & Wade, 2006; Tamir et al., 1990). One intervention (Alexandrov, Maslennikova, Kulikovm, Propirnij, & Perova, 1992) reported a significant negative effect, which either represented a chance finding or an iatrogenic effect.

There was significant heterogeneity in effect sizes ($Q = 204.41$, $p < .001$), indicating that there was variability across the effect sizes produced by the interventions (i.e., that effects were not equivalent across trials). The heterogeneity in the effects suggests that there may be participant, intervention, delivery, and design features that account for the variability in effect sizes.

Moderator Analyses

Two moderators could not be examined because of severe restrictions in range; because only two studies used credible active control conditions, and because we located only two unpublished reports, we did not consider type of control condition or publication status[3] further. Two potential confounding variables were not examined because they did not show significant relations to effect sizes: preliminary univariate analyses indicated that length of follow-up ($z = 1.58$, $p = .11$, $\beta = 0.18$) and the age range of participants in the trials ($z = .80$, $p = .42$, $\beta = 0.10$) were not significantly related to effect size magnitude. Within this context, it should be noted that preliminary analyses also indicated that

---

[3] Even though there were only two unpublished trials included in the present meta-analysis, we confirmed that there was no evidence that the unpublished studies had significant different effect sizes relative to published studies ($z = .03$, $p = .82$, $\beta = 0.03$).

Table 1
Operationalization and Descriptive Statistics for Moderators

Moderator	Value	Coding description and criteria	Descriptive statistic
Participant Features			
Age			
M	Continuous	Continuous variable representing mean age of participants at baseline in the sample.	$M = 10.89, SD = 3.60$
Range	Continuous	Continuous variable representing the number of values in the age range covered in the sample (e.g., 12 – 15 = 4). When age range was not reported, grade range was substituted (+2 because most grade levels primarily include students from two age levels). Age range was truncated at 6 because of positive skew.	$M = 3.38, SD = 1.50$
Gender	Female only = 1, both genders or male only = 0	Dichotomous variable representing whether intervention covered just females, or both males and females or males only. Because there were so few studies with just males, just male and male/female studies were combined to create a dummy variable.	Mixed gender and males = 50; Female only = 14
Ethnicity			
% Black/Hispanic	% Black/Hispanic	Percentage of participants from the entire sample (at baseline) who were either Black or Hispanic because these are the two largest ethnic minority groups at high risk for obesity.	$M = 33.34, SD = 39.65$
Dominant ethnic group	1 = Caucasian, 2 = Black, 3 = Hispanic, 4 = Asian/Pacific Islander, 5 = Native American	Ethnic group representing the majority of the sample.	Caucasian = 42; Black = 13; Hispanic = 4; Asian/Pacific Islander = 2; Native American = 1
Risk status	0 = universal 1 = selected,	Categorical variable representing whether the study was universally implemented or whether study participants were selected because they are a group at increased risk for obesity (e.g., Black or Hispanic ethnicity, 1-year college students, females with body dissatisfaction, and students with cardiovascular disease risk factors).	Universal = 30; Selected = 34
Intervention features			
Duration			
Hr	Continuous (hr)	Continuous variable representing hours of intervention. Unless otherwise noted, each session was coded as 1 hr each. Academic years were considered to be 28 weeks.	$M = 44.34, SD = 43.53$
Weeks	Continuous (weeks)	Continuous variable representing the number of weeks that the intervention took place. Academic years were considered to be 28 weeks.	$M = 31.30, SD = 32.73$
Parental involvement	0 = none, 1 = minimal (psychoeducation only), 2 = medium (attended sessions), 3 = high (behavioral change)	Categorical variable representing the level of parental involvement. Minimal involvement included when parents received psychoeducational materials, medium involvement included when parents were invited to attend sessions, and high involvement included when the intervention included a behavioral change (weight loss) target for parents.	None = 39; Minimal (psychoeducation only) = 13; Medium (attended sessions) = 7; High (behavioral change) = 5
Psychoeducational content	1 = yes, 0 = no	Coded as yes when intervention included presentation of information.	Psychoeducational content = 52; No psychoeducational content = 12
Psychoeducational content only	1 = yes, 0 = no	Coded as yes when intervention included only presentation of information.	Only psychoeducational content = 14; More than psychoeducational content = 50
Dietary improvement	1 = yes, 0 = no	Coded as yes when the intervention included a concrete, systematic change in diet, such as modified school lunches. Interventions that included only occasional snack preparation were coded as 0.	Dietary improvement = 7; No dietary improvement = 57
Physical activity increase	1 = yes, 0 = no	Coded as yes when intervention included a concrete, systematic increase in physical activity, such as a new or modified PE class or an after-school exercise program.	Physical activity increase = 39; No physical activity increase = 25
Reduced sedentary behavior	1 = yes, 0 = no	Coded as yes when the primary focus of the intervention was reducing sedentary behaviors, such as media (TV, video games) use, in order to reduce BMI.	Reduced sedentary behavior = 6; No reduced sedentary behavior = 58

Table 1 (*continued*)

Moderator	Value	Coding description and criteria	Descriptive statistic
No. of behavioral targets	Continuous, ranging from 0 to 3	Represents the number of health behaviors targeted for change in addition to weight control. Based on frequencies, we included smoking cessation, blood pressure control, and cholesterol reduction.	0 behavioral targets = 27 1 behavioral target = 11 2 behavioral targets = 20 3 behavioral targets = 6
Delivery features			
Teacher vs. interventionist	1 = interventionist, 0 = teacher	Dichotomous variable representing whether the intervention sessions were led by a teacher or trained interventionist.	Interventionist = 21 Teacher = 42
Didactic vs. interactive	1 = interactive, 0 = didactic	Dichotomous variable representing whether intervention included an interactive component, in which participants were able to discuss intervention material with each other or the facilitator, or whether the intervention included just didactic presentation of materials (i.e., lectures).	Interactive = 52 Didactic = 10
Design features			
Pilot study	1 = yes, 0 = no	Dichotomous variable representing whether intervention was a pilot or feasibility study. When authors did not specify, if the cell size was less than 50 (in each condition), we considered it a pilot study.	Pilot study = 18 No pilot study = 46
Recruitment method	1 = self-selected, 0 = not self-selected	Categorical variable representing whether participants were recruited through a population-based strategy (e.g., at particular schools) or self-selected in response to broader recruitment efforts (e.g., media advertisements).	Self-selected = 16 Not self-selected = 48
Random assignment	1 = yes, 0 = no	Dichotomous variable representing whether participants, classes, or schools were randomly assigned to conditions.	Random assignment = 46 No random assignment = 18
Nested data modeled incorrectly	1 = incorrectly, 0 = correctly	Dichotomous variable representing whether nested data were modeled correctly (e.g., unit of randomization was at the class or school level, but nesting variable was not modeled). Focused on data nested within schools, classes, or group-based interventions.	Nested data modeled incorrectly = 43 Nested data modeled correctly = 21
Active control group	1 = yes, 0 = no	Coded yes if the trial used an active control condition that was structurally equivalent to the intervention group (e.g., the same number of sessions) and a credible intervention for producing weight gain prevention effects (e.g., psychoeducational presentations on nutrition and exercise).	Active control group = 2 No active control group = 62
Follow-up length	Continuous	Length of follow-up in years.	$M = 0.19$, $SD = 0.52$
Unpublished report	1 = yes, 0 = no	Coded yes if report was never published in a peer reviewed journal (coded 0 if report was in press).	Published = 62 Unpublished = 2

Note. PE = physical education; BMI = body mass index.

Table 2
Descriptions of the Sample, Intervention Content, and Findings From Universal Obesity Prevention Trials

Study	Sample (boys and girls)	Intervention	Findings
Donnelly et al. (1996)	100	Interactive interventions (fitness and skill groups) that included mandated increases in PA, mandated dietary changes, and a psychoeducational component promoting healthy diet. Main outcomes were BMI, blood pressure, cholesterol, diet, PA, and knowledge.	No effects for BMI compared with control group receiving regular PE and school lunch.
Dwyer et al. (1983)	328	Intervention that consisted of mandated increases in PA. Main outcomes were BMI, blood pressure, cholesterol, fitness, and academic performance.	Significant effects for BMI compared with control group receiving regular PE program.
Gortmaker et al. (1999)	1,295	Interactive psychoeducational intervention that focused on decreasing media use and promoting healthy diet and exercise. Main outcomes were BMI, diet, weight loss behaviors, and media use.	Significant effects for BMI for girls only compared with matched control group.
Harrell et al. (1996)	1,274	Interactive interventions (risk based and classroom based) that included mandated increases in PA and a psychoeducational component promoting healthy diet and smoking prevention. Main outcomes were BMI, blood pressure, cholesterol, fitness, PA, and knowledge.	No effects for BMI compared with control group receiving regular health education curriculum.
Hopper et al. (1996)	97	Interactive intervention that included mandated increases in PA and a psychoeducational component that promoted healthy diet and exercise. Main outcomes were BMI, diet, PA, and knowledge.	No effects for BMI compared to control group receiving regular health education curriculum.
Kain et al. (2004)	3,086	Interactive intervention that included mandated increases in PA and a psychoeducational component. Main outcomes were BMI and physical fitness.	Significant effects for boys only compared with control group.
Killen et al. (1988)	1,130	Interactive psychoeducational intervention that presented information on healthy diet, exercise, smoking prevention, stress, and problem solving. Main outcomes were BMI, smoking, blood pressure, diet, fitness, PA, and knowledge.	Significant effects for BMI compared with matched control group.
Lionis et al. (1991)	146	Interactive psychoeducational intervention that promoted healthy diet, exercise, and smoking prevention. Main outcomes were BMI, blood pressure, cholesterol, diet, and knowledge.	Significant effects for BMI compared with assessment only control group.
Luepker et al. (1996)	1,653	Interactive interventions (school-based, school + family) that included mandated increases in PA and mandated dietary changes and a psychoeducational component that promoted healthy diet, exercise, and smoking prevention. Main outcomes were BMI, blood pressure, cholesterol, diet, PA, knowledge, self-efficacy, social support, and changes in school lunches and PE programs.	No significant effects for BMI compared with matched control group.
Manios et al. (2002)	641	Interactive intervention that included mandated increases in PA and a psychoeducational component promoting healthy diet, exercise, dental hygiene, and accident prevention. Main outcomes were BMI, fitness, PA, and knowledge.	Significant effects for BMI compared with control group receiving no health curriculum or PE.
Mo-suwan et al. (1998)	292	Interactive intervention that consisted of mandated increases in PA. Main outcome was BMI.	No effect for BMI compared with control group receiving regular PE.
Puska et al (1982)	851	Interactive interventions (county-wide, intensive) that included mandated dietary changes and a psychoeducational component promoting healthy diet, exercise, and smoking prevention. Main outcomes were BMI, smoking, blood pressure, cholesterol, diet, knowledge, and psychosomatic symptoms.	No effects for BMI compared with matched control group.
Resnicow et al. (1992)	1,209	Interactive intervention that included mandated dietary changes and a psychoeducational component that promoted healthy diet and exercise. Main outcomes were BMI, blood pressure, cholesterol, diet, knowledge, self-efficacy, and health attitudes.	No effects for BMI compared with control group.
Robinson (1999)	192	Interactive psychoeducational intervention designed to reduce media use. Main outcomes were BMI, diet, fitness, PA, media use, food eaten while watching TV, and other sedentary behaviors.	Significant effects for BMI compared with assessment only control group.

Table 2 (*continued*)

Study	Sample (boys and girls)	Intervention	Findings
Sahota et al. (2001)	595	Interactive intervention that included mandated dietary changes and a psychoeducational component promoting healthy diet and exercise. Main outcomes were BMI, diet, PA, knowledge, self-concept, dietary restraint, and body image.	No effects for BMI compared with control group receiving regular health curriculum and school lunches.
Sallis et al. (1993)	549	Interactive interventions (interventionist led, teacher led) that included mandated increases in PA and a psychoeducational component promoting exercise through self-management. Main outcome was BMI.	Significant effects for BMI for girls in both intervention conditions compared with control group receiving regular PE.
Tamir et al. (1990)	406	Didactic psychoeducational intervention that promoted healthy diet and exercise. Main outcomes were BMI, blood pressure, and cholesterol.	Significant effects for BMI compared with matched control group.
Vandongen et al. (1995)	1,140	Interactive interventions (fitness education, fitness + school nutrition, school nutrition, school + home nutrition, home nutrition) that included mandated increases in PA and a psychoeducational component that promoted healthy diet. Main outcomes were BMI, blood pressure, cholesterol, diet, and fitness.	No effects for BMI compared with control group receiving regular PE and health curriculum.
Walter et al. (1988)	1,769	Didactic psychoeducational intervention that promoted healthy diet, exercise, and smoking prevention. Main outcomes were BMI, smoking, blood pressure, cholesterol, diet, fitness, and knowledge.	No effects for BMI compared with control group receiving regular health curriculum.

Note. BMI = body mass index; PA = physical activity; PE = physical education.

publication year, a variable commonly included in meta-analytic reviews, was not a significant predictor of effect size ($z = 1.44$, $p = .15$, $\beta = 0.17$).

We also conducted preliminary analyses to determine which operationalization of participant ethnicity and psychoeducational content we should examine in the models. First, with regard to ethnicity, analyses indicated that neither the percentage of the sample that was Black or Hispanic ($z = .48$, $p = .63$, $\beta = 0.06$) nor the predominant ethnic group in each sample (which was represented with a series of dummy-coded vectors) was significantly related to effect sizes. Dummy variables representing Black ($z = .43$, $p = .67$, $\beta = 0.05$), Hispanic ($z = .31$, $p = .76$, $\beta = 0.04$), Asian and Pacific Islander ($z = 1.51$, $p = .13$, $\beta = 0.18$), and Native American ($z = 1.53$, $p = .12$, $\beta = 0.18$) were not statistically significant. We focused exclusively on the former operationalization for this report because the latter operationalization had some very small cell sizes (e.g., predominantly Native American participants) and necessitated the use of multiple dummy-coded vectors to represent this operationalization. Second, because the code representing whether interventions had only psychoeducational content was not systematically related to the effect sizes in a univariate model ($z = -.38$, $p = .71$, $\beta = -0.04$), we limited our analyses to whether the intervention contained any psychoeducational content (which allowed us to use a parallel approach for all of our intervention content variables). Thus, although we initially coded 23 effect size moderators (see Table 4), the moderator analyses focused on the 18 effect size moderators listed in Table 7.

Parental involvement was initially analyzed as a four-level moderator with the following levels represented: no parental involvement, psychoeducational material provided to parents, parental attendance of sessions, and parental behavioral change. However, in preliminary analyses, dummy-coded variables representing psychoeducational material ($z = -1.14$, $p = .26$, $\beta = -0.14$), parental attendance of sessions ($z = -.19$, $p = .85$, $\beta = -0.02$), and parental involvement ($z = .38$, $p = .71$, $\beta = 0.05$) were not statistically significant predictors of effect size. We therefore simplified this variable into a dichotomous variable (no parental involvement or psychoeducational material = 0; parental attendance or parental behavior change = 1) so that we could use a single dummy coded vector to represent this variable.

Random effects regression models tested whether the putative moderators were related to observed effect sizes. Random effects models separate variance between effect sizes and variance attributable to individual studies. Inferentially, random effects models can be generalized to a broader set of studies or potential studies in contrast to fixed effects models that do not account for variance attributable to a particular study. Regression models were implemented using SPSS macros (Lipsey & Wilson, 2001) for inverse variance weighted regression with maximum likelihood estimation. The correlations between the moderators are presented in Table 7.

We first examined moderators in separate univariate regression models to investigate the bivariate relations between moderators and effect sizes that were not complicated by colinearity. We then entered the moderators that showed significant effects in the univariate models into a multivariate model to estimate the unique effect of each moderator controlling for the effects of the other moderators with significant effects. The five continuous modera-

Table 3
Descriptions of the Sample, Intervention Content, and Findings From Selected Obesity Prevention Trials

Study	Sample	Intervention	Findings
Alexandrov et al. (1992)	766 boys	Interactive program that involved counseling children and their parents on diet, PA, and smoking. Main outcomes were BMI, smoking, blood pressure, and cholesterol.	Significant effects for BMI compared with control group.
Baranowski et al. (2003)	35 African American girls in ≥50th percentile BMI	Interactive summer camp that involved mandated increases in PA and dietary changes, and a psychoeducational component to promote healthy diet and exercise. Main outcomes were BMI, diet, and PA.	No effect for BMI compared with control group receiving regular camp.
Bayne-Smith et al. (2004)	442 urban, multiethnic girls	Interactive intervention that involved mandated increases in PA and a psychoeducational component to promote behavioral change in CVD risk factors, heart health knowledge, and fitness levels. Main outcomes were BMI, blood pressure, and cholesterol levels.	No effect for BMI compared with control group receiving traditional PE.
Beech et al. (2003)	60 African American boys and girls ≥ 25 percentile BMI	Interactive intervention that included mandated increases in PA and a psychoeducational component on knowledge and behavior skills to promote healthy diet and exercise for children. A parallel intervention was delivered to parents that also sought to promote increased PA and an improved diet in the children. Main outcomes were BMI, diet, and PA.	No effect for BMI compared with control group receiving a self-esteem enhancing curriculum.
Burke et al. (1998)	720 boys and girls at high risk for CVD	Interactive interventions (high risk, high + low risk) that involved mandated increases in PA and a psychoeducational component to promote healthy diet. Main outcomes were BMI, blood pressure, cholesterol, diet, fitness, PA, and media use.	No effect for BMI compared with control group.
Bush et al. (1989)	471 African American boys and girls	Didactic psychoeducational interventions (full intervention, partial intervention) focused on promoting healthy diet, exercise, and smoking prevention. Main outcomes were BMI, smoking, blood pressure, cholesterol, diet, fitness, and knowledge.	No effect for BMI compared with control group.
Eliakim et al. (2000)	44 boys	Interactive intervention that involved mandated increases in PA. Main outcomes were BMI and cholesterol.	Significant effect for BMI compared with controls not receiving program.
Fardy et al. (1996)	346 inner-city girls and boys	Interactive intervention that involved mandated increases in PA and a psychoeducational component promoting good health. Main outcomes were BMI, smoking, cholesterol, diet, fitness, PA and knowledge.	No effect for BMI compared with controls receiving traditional PE.
Fitzgibbon et al. (2004)	300 African American preschool girls and boys	Interactive intervention that involved mandated increases in PA and a psychoeducational component promoting healthy eating and exercise. The main outcome was BMI.	Significant effect for BMI compared with assessment-only control group.
Flores (1995)	81 African American and Hispanic boys and girls	Interactive intervention that included mandated increases in PA and a psychoeducational component promoting healthy diet, exercise, and smoking prevention. Main outcomes were BMI, fitness, and attitudes about fitness.	Significant effect for BMI for girls only compared with control group who received regular exercise program.
Gutin et al. (1995)	22 obese African American girls	Interactive intervention that involved mandated increases in PA. Main outcome was BMI.	No effect for BMI compared with active control group.
Gutin & Owens (1999)	73 obese boys and girls	Interactive intervention that involved mandated increases in PA. Main outcome was BMI.	Significant effect for BMI compared with control group.
Hansen et al. (1991)	132 high and normal risk boys and girls	Interactive, school-based interventions (high risk, normal) that involved mandated increases in PA. Main outcomes were BMI, blood pressure, and fitness.	No effect for BMI compared with control group who received regular PE program.
Harrell et al., (1998)	422 boys and girls with multiple CVD risk factors	Didactic interventions (risk-based, classroom-based) that involved mandated increases in PA and a psychoeducational component that promoted healthy diet and smoking prevention. Main outcomes were BMI, blood pressure, cholesterol, diet, and knowledge.	No effect for BMI compared with control group receiving regular PE and health curriculum.

Table 3 (*continued*)

Study	Sample	Intervention	Findings
Harvey-Berino & Rourke (2003)	40 Native American girls and boys	Interactive psychoeducational intervention that included maternal participation designed to reduce the prevalence of obesity. Main outcome was BMI.	No significant effect for BMI compared with active control group.
Ignico & Mahon (1995)	28 low-fit boys and girls	Interactive intervention that involved mandated increases in PA. Main outcomes were BMI, cholesterol, and fitness.	No effect for BMI compared with assessment only control group.
Matheson et al. (2005)	61 Mexican American boys and girls	Interactive psychoeducational intervention with parental component designed to improve health and prevent obesity. Main outcome was BMI.	No effect for BMI compared with active control group.
Matvienko et al. (2001)	40 female college freshman	Interactive psychoeducational intervention that consisted of a nutrition science course. Main outcomes were BMI, diet, and knowledge.	No effect for BMI compared with control group receiving no intervention.
McMurray et al. (2002)	1140 boys and girls	Interactive interventions that involved mandated increases in PA (exercise, exercise + education) and a psychoeducational component (education, exercise + education) that promoted healthy diet, exercise, and smoking prevention. Main outcomes were BMI, blood pressure, and fitness.	No effect for BMI compared with control group receiving regular PE and health curriculum.
Nader et al. (1992)	206 Mexican American and White boys and girls	Interactive family intervention that included mandated increases in PA and a psychoeducational component that promoted healthy diet. Main outcomes were BMI, blood pressure, cholesterol, diet, PA, and knowledge.	No effect for BMI compared with assessment-only control group.
Neumark-Sztainer, et al. (2003)	180 girls either overweight or at risk for becoming overweight	Interactive intervention that included mandated increases in PA and a psychoeducational component that promoted healthy diet and exercise. Main outcomes were BMI, diet, PA, and knowledge.	No effect for BMI compared with control group receiving regular PE.
Robinson et al. (2003)	61 African American girls	Randomized interactive intervention that included mandated increases in PA and a psychoeducational component designed to reduce media use. Main outcomes were BMI, diet, media use, and PA.	Significant effects for BMI compared with active control group.
Stice, Orjada, & Tristan (2006)	95 female college students	Replication of Stice & Ragan (2002).	Significant effect for BMI compared with matched control group.
Stice & Ragan (2002)	66 female college students	Nonrandomized interactive, psychoeducational eating disorders class that presented information on eating disorders, obesity, putative causes, prevention, and treatment of these disorders. Main outcomes were BMI, diet, thin-ideal internalization, body dissatisfaction, dieting, eating disorder symptoms, and depression.	Significant effect for BMI compared with matched control group.
Stice, Shaw, et al. (2006)	481 body-dissatisfied girls	Interactive dissonance-based intervention that involved counter-attitudinal exercises designed to reduce thin-ideal internalization, body dissatisfaction, and eating disorder symptoms. A second interactive intervention focused on promoting lasting increases in PA and improvements to dietary intake. Main outcomes were eating disorder risk factors, bulimic symptoms, and BMI.	Significant effect for BMI compared with control group (significantly lower risk for obesity onset over 1-year follow-up).
Story et al. (2003)	53 African American girls	Interactive intervention that involved mandated increases in PA and a psychoeducational component that promoted healthy diet and exercise. Main outcomes were BMI, diet, PA, knowledge, body image, and dieting.	No effect for BMI compared with active placebo control group.
Willet (1995)	54 African American girls	Interactive psychoeducational intervention that promoted healthy diet and exercise. Main outcomes were BMI, diet, parental support, and role-modeling.	No effect for BMI compared with attention placebo control group.

Note. PA = physical activity; BMI = body mass index; CVD = cardiovascular disease.
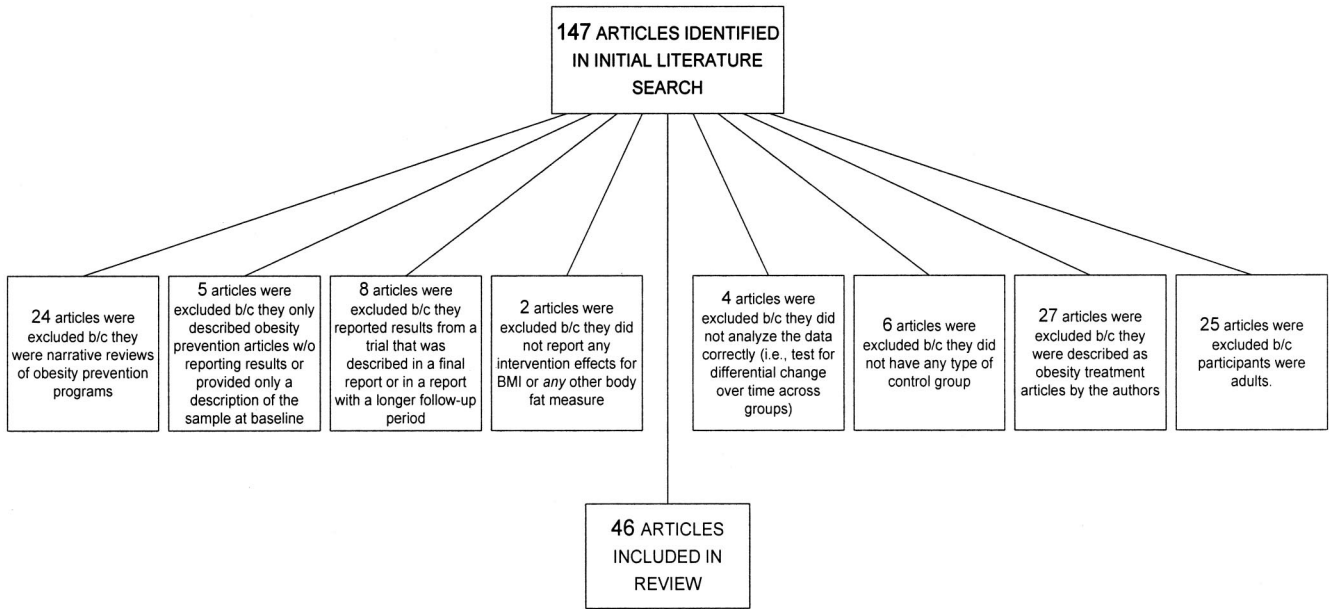
Figure 1. Flowchart illustrating the number of articles omitted for the various exclusion criteria. b/c = because; w/o = without.

Table 4

Interrater Agreement for All Variables Abstracted for the Present Meta-Analytic Review

Variable	ICC	κ
Participant age		
M	1.00	
Range	0.98	
Participant gender		1.00
Participant ethnicity		
Black and Hispanic	1.00	
Dominant ethnic group		1.00
Risk status of participants		1.00
Intervention duration		
Hr	1.00	
Weeks	1.00	
Parental involvement		1.00
Psychoeducational content		1.00
Psychoeducational content only		1.00
Dietary improvement		0.94
Physical activity increase		0.88
Reduced sedentary behavior		1.00
No. of behavioral targets	1.00	
Teachers vs. interventionist		1.00
Didactic vs. interactive		1.00
Pilot study		0.88
Recruitment method		1.00
Random assignment		1.00
Nested data modeled incorrectly		0.87
Active control group		1.00
Follow-up length	1.00	
Unpublished report		1.00
Effect size	0.95	

Note. Interrater agreement was calculated for a randomly selected 30% of the trials. ICC = interclass correlation coefficient.

tors—average age, percentage of Black and Hispanic participants, intervention duration in hours, intervention duration in weeks, and number of behavioral targets—were standardized in a z score format. We tested for linear and quadratic effects for the five continuous moderators, as statisticians recommend testing for such higher order effects to decrease the risk of model misspecification (Hosmer & Lemeshow, 2000). Effect sizes were regressed on the linear and quadratic terms. If the quadratic effect was not significant, the quadratic term was removed from subsequent models to ensure that linear effects were not obscured by colinearity between the linear and quadratic terms. When the quadratic term was significant, both the linear and quadratic terms were retained for all subsequent models.

Univariate Models

Among the five continuous moderators, the model for participant age was the only model in which the quadratic term was significant. Both the linear and quadratic age terms were significant (see Table 8). As indicated in Figure 2, larger effect sizes tended to emerge in trials involving children and adolescents, but smaller effect sizes occurred in trials involving preadolescents. To probe the form of this curvilinear pattern, we examined mean effect sizes for age tertiles: Interventions with an average less than or equal to 9.23 years exhibited effect sizes that were only marginally different from zero (mean $r = .03$, $p = .07$, $n = 21$), interventions with an average age greater than 9.23 and less than or equal to 11 did not exhibit an effect significantly different from zero (mean $r = .01$, $p = .42$, $n = 23$), and interventions with an average age greater than 11 were significantly different from zero (mean $r = .07$, $p < .05$, $n = 20$).

Significantly larger effects were observed in female-only trials than in mixed-sex and male-only trials (see Table 8). Follow-up analyses revealed that the average effect for programs focusing

solely on females was significantly different from zero (mean $r =$ .13, $p < .01$, $n = 14$), whereas the average effect for programs targeting mixed-sex samples and male-only samples was trivial and not significantly different from zero (mean $r = .02$, $p = .06$, $n = 50$).

Intervention duration was examined as a function of hours and weeks. While there was not a significant effect for duration in hours, there was a significant negative effect for duration in weeks (see Table 8). Interventions below the median of 16 weeks exhibited a mean effect size significantly greater than zero (mean $r =$ .06, $p < .01$, $n = 31$) in contrast to effect sizes of interventions at or above the median of 16 weeks that were not significantly greater than zero (mean $r = .02$, $p = .15$, $n = 33$).

The model for number of behavioral targets containing only a linear term had a significant negative coefficient (see Table 8), indicating that effect size decreased as the number of nonweight related targets increased. Interventions that targeted only weight change exhibited effect sizes greater than zero (mean $r = .09$, $p <$ .001, $n = 27$), and interventions that targeted other behavioral changes in addition to weight change were not significantly different from zero (mean $r = .01$, $p = .47$, $n = 37$).

Pilot trials of interventions exhibited significantly larger effect sizes than did fully powered demonstration trials (see Table 8). Follow-up analyses revealed that the average effect for pilot studies was significantly different from zero (mean $r = .14$, $p < .001$, $n = 18$), whereas the average effect for interventions evaluated in demonstration trials were not significantly different from zero (mean $r = .02$, $p = .07$, $n = 46$).

Trials that used a self-selected recruitment method resulted in significantly larger effect sizes than did trials that used population-based recruitment methods (see Table 8). Follow-up analyses showed that trials using self-selected recruitment were significantly greater than zero (mean $r = .14$, $p < .001$, $n = 16$), whereas the average effect for trials using population-based recruitment were not significantly different from zero (mean $r = .02$, $p = .10$, $n = 48$).

Multivariate Model

A multivariate model was estimated containing moderators that were significant predictors of effect size in previous univariate models: the linear and quadratic terms for age, participant gender, number of behavioral targets, duration in weeks, whether the trial was a pilot study, and recruitment method. Both the linear term ($z = -4.14$, $p < .001$, $\beta = -2.06$) and the quadratic terms for age showed significant univariate effects in this model ($z = 4.56$, $p <$ .001, $\beta = 2.36$). The only other moderator that remained statistically significant in the multivariate model was self-selected recruitment ($z = 2.07$, $p < .05$, $\beta = 0.30$). Participant gender ($z =$ -1.96, $p = .05$, $\beta = -0.33$), duration in weeks ($z = -.66$, $p =$ .51, $\beta = -0.08$), number of behavioral targets ($z = -.43$, $p = .67$, $\beta = -0.06$), and whether the trial was a pilot study ($z = 1.59$, $p =$ .11, $\beta = 0.19$) did not show significant unique effects in the multivariate model. The R^2 for the full model was .42.

Discussion

Summary of Effect Sizes

The first aim of this review was to summarize the effects of prevention programs that sought to produce weight gain preven-

tion effects. One of the more noteworthy findings was that although numerous prevention programs have been evaluated, most (79%) did not produce statistically reliable weight gain prevention effects. Indeed, the average intervention effect size was an r of .04, which would be considered trivial by most researchers and clinicians. This pattern of findings attests to the difficulty of altering the health behaviors that increase risk for weight gain and obesity onset, and it echoes the modest success of treatment programs for obesity in producing lasting changes in body weight (Jeffery et al., 2000).

Although it is tempting to conclude that it is particularly challenging to prevent future weight gain, the percentage of programs that produced significant intervention effects for obesity prevention programs (21%) is similar to that of prevention programs for other public health problems such as HIV (22%; Logan, Cole, & Leukefeld, 2002) and eating disorders (25%; Stice & Shaw, 2004), although smoking prevention programs have a higher rate of significant intervention effects (60%; Skara & Sussman, 2003). The average effect size for obesity prevention programs ($r = .04$) is also similar to the average effect size observed for prevention programs for other public health problems, such as smoking ($r =$ .07; Hwang, Yeagley, & Petosa, 2004), substance abuse ($r = .05$; Tobler et al., 2000), HIV ($r = .05$; Logan et al., 2002), and eating disorders ($r = .12$; Stice & Shaw, 2004). This broader pattern of modest average returns for prevention programs aimed at a variety of health behaviors implies that most prevention programs are only minimally effective in reducing maladaptive health behaviors.

The above conclusion makes it imperative to focus on the 21% of prevention programs that produced significant weight gain prevention effects. The average effect size for these 13 interventions was $r = .22$ ($p < .001$), which corresponds to a medium effect size and is of clinical significance. The effect sizes ranged from a low of .06 (Killen et al., 1988) to a high of .50 (Eliakim et al., 2000), which is a remarkably large effect size for a prevention program.

There are several noteworthy features of the interventions that produced weight gain prevention effects. First, these programs were relatively intensive: On average, they involved 40 hr of intervention time (range = 3–120 hr). However, it was not only the successful interventions that were intensive; the average number of intervention hours was 46 for the programs that did not produce weight gain prevention effects (range = 5–280 hr). Intervention duration is important to consider because it is difficult to disseminate intensive programs in schools given the competing demands for classroom time. Moreover, the long intervention duration also translates into higher dissemination costs, because both training and delivery costs will be greater. Given the range of intervention durations, it is important to consider the effect per hour of the intervention when comparing the different programs. As indicated in Tables 5 and 6, the average r per hour of interventions for those that produced significant weight gain prevention effects ranged from .001 (Robinson et al., 2003) to .063 (Stice, Shaw, et al., 2006), suggesting that certain interventions produce more apparent return per hour of intervention, which should facilitate dissemination and lower dissemination cost.

It is also noteworthy that only 2 of the 13 prevention programs that produced significant weight gain prevention effects were primarily conceptualized as obesity prevention programs (Fitzgibbon et al., 2004; Robinson, 1999). The other interventions were

Table 5
Universal Programs: Moderator Values and Effect Sizes

Study	Participant characteristics			Intervention duration		Parental involvement	Psychoeducational content	Dietary improvement	Physical activity increased
	Age	Sex	% Black/ Hispanic	Hr	Weeks				
Donnelly et al. (1996)	9.2	0	6	40.0	56	0	1	1	1
Dwyer et al. (1983)									
Fitness	10.2	0	0	87.5	14	0	0	0	1
Skill	10.2	0	0	87.5	14	0	0	0	1
Gortmaker et al. (1999)	11.7	0	13	26.5	56	0	1	0	0
Harrell et al. (1996)	8.9	0	20	40.0	8	0	1	0	1
Hopper et al. (1996)	8.9	0		30.0	10	3	1	0	1
Kain et al. (2004)	10.6	0	100	47.0	24	2	1	0	1
Killen et al. (1988)	15.0	0	8	16.1	7	0	1	0	0
Lionis et al. (1991)	13.5	0	0	20.0	28	1	1	0	0
Luepker et al. (1996)									
School based	8.8	0	27	33.0	84	0	1	1	1
School + family	8.8	0	27	33.0	84	1	1	1	1
Manios et al. (2002)	6.0	0	0	120.0	84	2	1	0	1
Mo-suwan et al. (1998)	4.5	0	0	54.0	30	0	0	0	1
Puska et al. (1982)									
County-wide	13.0	0	0		56	2	1	1	0
Intensive	13.0	0	0		56	2	1	1	0
Resnicow et al. (1992)	7.5	0	83	21.0	70	0	1	1	0
Robinson (1999)	8.9	0		15.0	24	2	1	0	0
Sahota et al. (2001)	8.4	0	0		28	0	1	1	0
Sallis et al. (1993)									
Interventionist	9.3	0	8	64.0	56	1	1	0	1
Teacher	9.3	0	8	64.0	56	1	1	0	1
Tamir et al. (1990)	6.0	0	0	35.0	56	0	1	0	0
Vandongen et al. (1995)									
Fitness education	11.0	0	0	38.0	28	0	1	0	1
Fitness + school nutrition	11.0	0	0	48.0	28	0	1	0	1
School nutrition	11.0	0	0	10.0	28	0	1	0	0
School nutrition + home nutrition	11.0	0	0	15.0	28	0	1	0	0
Home nutrition	11.0	0	0	5.0	28	0	1	0	0
Walter et al. (1988)	8.9	0	43	280.0	140	0	1	0	0

Note. BMI = body mass index.
* *p* < .05.

described as general health education interventions (Lionis et al., 1991; Manios et al., 2002), cardiovascular disease prevention programs (Killen et al., 1988; Tamir et al., 1990), physical activity interventions (Dwyer et al., 1983; Eliakim et al., 2000; Gutin & Owens, 1999), and eating disorder prevention programs (Stice, Orjada, & Tristan, 2006; Stice & Ragan, 2002; Stice, Shaw, et al., 2006). This suggests that there may be many avenues to preventing obesity beyond programs that are directly billed as weight gain prevention programs and that it would be fruitful to follow up these alternative interventions. Another benefit of these alternative interventions is that they produce effects for additional public health problems beyond obesity (e.g., smoking and eating disorders).

A third noteworthy feature of the 13 effective programs is that only 3 (5% of the total programs evaluated) of these interventions produced weight gain prevention effects that persisted over a significant follow-up period (Fitzgibbon et al., 2004; Stice, Shaw, et al., 2006; Stice, Orjada, & Tristan, 2006). The remainder of the programs produced weight gain prevention effects from pretest to posttest. Because virtually all of the weight loss effects observed in obesity treatment studies disappeared by the 3-year follow-up

(Jeffery et al., 2000), it is possible that the weight gain prevention effects did not persist. It will be important to include longer term follow-ups in future obesity prevention trials.

A fourth feature of the programs that produced weight gain prevention effects is that the positive effects for weight gain have been replicated in multiple trials for only one intervention (Stice, Orjada, & Tristan, 2006; Stice & Ragan, 2002). Given that independent replication is a necessary step in establishing that a program is efficacious—that it produces statistically reliable effects in highly controlled trials (American Psychological Association [APA], 1995)—this represents another important gap in the literature.

Moderators of Obesity Prevention Effects

The second aim of the present review was to examine participant, intervention, delivery, and design features that are associated with larger intervention effects. Results indicated that intervention effects were stronger for children and adolescents relative to preadolescents, with the strongest effects emerging for adolescents. The evidence that obesity prevention programs were most effective for adolescents generally conformed to the initial hypoth-

Reduced sedentary behaviors	No. of behavioral targets	Teacher vs. interventionist	Didactic vs. interactive	Pilot study	Recruitment method	Random assignment	Nested data modeled incorrectly	Effect size for BMI	Effect size per hr of intervention
0	2	0	0	0	0	0	1	.00	.000
0	2	0		0	0	1	0	.16*	.002
0	2	0		0	0	1	0	.04	.000
1	0	0	1	0	0	1	0	.03	.001
0	2	0	1	0	0	1	0	−.05	−.001
0	0	0	1	1	0	1	1	.00	.000
0	0	0	1	0	0	0	0	.05	.001
0	2	1	1	0	0	1	1	.06*	.004
0	2	0	1	0	0	0	1	.16*	.008
0	2	0	1	0	0	1	0	−.04	−.001
0	2	0	1	0	0	1	0	−.04	−.001
0	0	0	1	0	0	0	1	.10*	.001
0	0	1	1	0	0	1	0	.00	.000
0	3	1	1	0	0	0	1	.04	
0	3	1	1	0	0	0	1	.03	
0	2	0	1	0	0	0	1	.04	.002
1	0	0	1	0	0	1	1	.22*	.015
0	0	1	1	0	0	1	0	0	.000
0	0	1	1	0	0	1	1	−.09	−.001
0	0	0	1	0	0	1	1	−.08	−.001
0	2	0	0	0	0	0	1	.14*	.004
0	2	0	1	0	0	0	1	.04	.001
0	2	0	1	0	0	0	1	.02	.000
0	2	0	0	0	0	0	1	−.04	−.004
0	2	0	1	0	0	0	1	.00	.000
0	2	0	1	0	0	0	1	.00	.000
0	3	0	0	0	0	1	0	.05	.000

esis (Baranowski et al., 2002). Theoretically, older participants are better able to grasp intervention material and wield control over their food and physical activity choices than are younger participants. In addition, adolescence is a developmental period during which individuals often must develop self-regulatory skills because they are becoming more autonomous, and it may be particularly useful to deliver obesity prevention programs at this time. There was some evidence that obesity prevention programs are more effective for children versus preadolescents, which seems inconsistent with the suggestion that obesity prevention programs are less effective for children (Baranowski et al., 2002). An examination of the prevention programs aimed at children suggests that the interventions that produced the largest effect in this age range included a parental involvement component (e.g., Harvey-Berino & Rourke, 2003), which may be a particularly effective way to alter the food environment of children in this age range. However, it is difficult to interpret age effects in a meta-analytic review of intervention programs because the same interventions were not tested across a range of ages (i.e., the age of participants and the content of the intervention varies). Thus, it may be something about the types of programs that are delivered to adolescents,

rather than the adolescent developmental period, that explains why larger effects tended to emerge with adolescents. Nonetheless, the fact that the effect of participant age remained significant in the multivariate model suggests that this effect did not emerge because of some confound with the other moderators that showed significant effects (e.g., participant gender or intervention duration).

As expected on the basis of a prior meta-analytic review of eating disorder prevention programs (Stice & Shaw, 2004), the univariate model found that obesity prevention programs were more effective when delivered solely to female samples versus male or mixed-gender samples. We theorize that females may be more receptive in general to interventions promoting weight control because of the significantly stronger societal pressures for them to conform to a thin-beauty ideal espoused by Western cultures. It was also noteworthy that for the trials for which separate effect sizes for the two genders were available, the average effect size for females ($r = .06$) was larger than the average effect size for males ($r = .02$) and that only the former differed significantly from zero. This finding suggests that extant programs may be more effective for females and that there is a need to develop alternative interventions for males. However, participant

Table 6
Selected Programs: Moderator Values and Effect Sizes

Study	Participant characteristics			Intervention duration		Parental involvement	Psychoeducational content	Dietary improvement	Physical activity increase
	Age	Sex	% Black/ Hispanic	Hr	Weeks				
Alexandrov et al. (1992)	11.9	0	0		84	2	1	0	0
Baranowski et al. (2003)	8.0	1	100	82	12	1	1	1	1
Bayne-Smith et al. (2004)	16.0	1	75	30	12	0	1	0	1
Beech et al. (2003)									
Girls	8.9	1	100	18	12	0	1	0	1
Parents	9.0	1	100	18	12	1	1	0	1
Burke et al. (1998)									
High risk	11.0	0	0	53	10	1	1	0	1
High and low risk	11.0	0	0	53	10	1	1	0	1
Bush et al. (1989)									
Full intervention	10.5	0	100	84	140	1	1	0	0
Partial intervention	10.5	0	100	84	140	1	1	0	0
Eliakim et al. (2000)	16.0	0	9	50	5	0	0	0	1
Fardy et al. (1996)	15.9	0	68	27	11	0	1	0	1
Fitzgibbon et al. (2004)	4.0	0	99	28	14	1	1	0	1
Flores (1995)	12.6	0	100	30	12	0	1	0	1
Gutin et al. (1995)	9.2	1	100	25	10	0	0	0	1
Gutin & Owens (1999)	9.4	0	44	53	16	0	0	0	1
Hansen et al. (1991)									
High risk	10.0	0	0	20	32	0	0	0	1
Normal	10.0	0	0	20	32	0	0	0	1
Harrell et al. (1998)									
Risk based	9.0	0	22	40	8	0	1	0	1
Classroom based	9.0	0	22	40	8	0	1	0	1
Harvey-Berino & Rourke (2003)	1.8	0	0	16	16	3	1	0	0
Ignico & Mahon (1995)	9.7	0		30	10	0	0	0	1
Matheson et al. (2005)	7.5	0	100	6	24	3	1	0	0
Matvienko et al. (2001)	19.0	1	5	16	16	0	1	0	0
McMurray et al. (2002)									
Exercise only	12.5	0	24	12	8	0	0	0	1
Education only	12.5	0	24	16	8	0	1	0	0
Exercise + education	12.5	0	24	28	8	0	1	0	1
Nader et al. (1992)	12.0	0	59	12	36	3	1	0	1
Neumark-Sztainer et al. (2003)	15.4	1	33	88	24	1	1	0	1
Robinson et al. (2003)	9.5	1	100	155	12	2	1	0	1
Stice, Orjada, & Tristan (2006)	21.3	1	16	45	15	0	1	0	0
Stice & Ragan (2002)	21.0	1	17	45	15	0	1	0	0
Stice, Shaw, et al. (2006)									
Dissonance	17.0	1	25	3	3	0	1	0	0
Healthy weight	17.0	1	25	3	3	0	1	0	0
Story et al. (2003)	9.3	1	100	24	12	2	1	0	1

Note. BMI = body mass index.
* $p < .05$.

gender did not have a significant unique effect in the multivariate model because participant gender was correlated (see Table 7) with the number of behavioral intervention targets ($r = -.47$) and with whether the trial relied on self-selected recruitment ($r = .66$). This pattern of findings either implies that the effect of participant gender is actually driven by another moderator that was confounded with gender or that certain participant, intervention, and design features simply tend to co-occur naturally, which has the effect of attenuating the unique effect of each moderator in multivariate models. Although this interpretational ambiguity is not unique to meta-analytic reviews, as it arises with any correlational data (e.g., for prospective risk factor studies), it does signal that caution should be used when the moderator effects are interpreted.

Unexpectedly, interventions with a relatively shorter duration (in weeks) produced significantly larger effects than did those that were longer in duration. This might suggest that interventions that are long in duration are unappealing to participants, which causes them to drop out of the intervention or to disengage from the program. Intervention duration in weeks did not show a significant unique effect in the multivariate model, which occurred because intervention duration in weeks was correlated (see Table 7) with the number of health behavior targets ($r = .51$). Because it is logical to expect that interventions targeting multiple health behaviors would be longer in duration, this appears to be natural colinearity that simply functions to attenuate the unique effects of each of these factors.

Reduced sedentary behaviors	No. of behavioral targets	Teacher vs. interventionist	Didactic vs. interactive	Pilot study	Recruitment method	Random assignment	Nested data modeled incorrectly	Effect size for BMI	Effect size per hr of intervention
0	3	1	1	0	0	0	1	−.11*	
0	0	1	1	1	1	1	1	−.08	.001
0	2	0	1	0	0	1	1	−.04	−.001
0	0	1	1	1	1	1	1	.16	.009
1	0	1	1	1	1	1	1	.16	.009
0	1	0	1	0	0	1	0	.00	.000
0	1	0	1	0	0	1	0	.00	.000
0	3	0	0	0	0	0	1	−.08	−.001
0	3	0	0	0	0	0	1	−.08	−.001
0	1	0	1	1	1	1	1	.50*	.010
0	2	0	1	0	0	1	1	.00	.000
0	0	0	1	0	0	1	0	.15*	.005
0	0	0	1	1	0	1	1	.14	.001
0	0	1	1	1	1	0	1	.05	.002
0	0	1	1	1	1	1	1	.30*	.006
0	1	0	1	1	0	1	1	.00	.000
0	1	0	1	1	0	1	1	.00	.000
0	2	0	0	0	0	1	0	−.06	−.002
0	2	0	0	0	0	1	0	−.03	−.001
0	0	1	1	1	1	1	0	.30	.019
0	1	1	1	1	0	0	1	.00	.000
0	0	1	1	1	1	1	0	.04	.006
0	0	0	1	1	1	1	1	.14	.009
0	1	0	1	0	0	1	1	.00	.000
0	1	0	0	0	0	1	1	.00	.000
0	1	0	1	0	0	1	1	.00	.000
0	2	1	1	0	1	0	1	.00	.000
0	0		1	0	1	0	0	.00	.000
1	0	1	1	1	1	1	0	.19*	.001
0	0	1	1	0	0	0		.31*	.007
0	0	1	1	1	0	0		.32*	.007
0	0	1	1	0	1	1	1	.17*	.057
0	0	1	1	0	1	1	1	.19*	.063
1	0	1	1	1	1	1	1	−.02	.001

As hypothesized, effects were significantly larger for interventions that solely focused on obesity prevention than for interventions that focused on additional health behaviors. This finding is consistent with our suggestion that message complexity may curtail the effectiveness of health promotion interventions. It may be necessary to keep health promotion interventions relatively short and simple for maximal intervention effects. The effect for number of behavioral targets became nonsignificant in the multivariate model because (see Table 7) interventions targeting multiple health behaviors tended to be longer in duration ($r = .51$ with duration in weeks) and tended to not use self-selected recruitment ($r = -.51$) or to be pilot trials ($r = -.49$). Although it is possible that the effect for number of behavioral targets was actually driven

by one of these moderators, as noted above, it is also possible that this colinearity is natural and simply attenuates the unique effects for the moderators.

Also as expected, there was evidence that pilot trials tend to produce larger effects than large demonstration trials. Presumably, this finding emerged because interventionists and researchers are more passionate about new interventions, which contributes to larger effect sizes because of demand characteristics or because demonstration trials are more methodologically rigorous (e.g., are more likely to use blinded assessors), which makes them more immune to experimenter bias. The effect of this moderator also became nonsignificant in the multivariate model because pilot trials tended to focus on fewer behavioral targets ($r = -.51$) and

Table 7
Correlations Among the Putative Moderators of Weight Gain Intervention Effects

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Mean participant age	—	.40	−.15	.26	−.16	−.22	−.20	.08	−.11	−.24	−.12	.04	−.14	−.13	−.07	.04	−.11	.35	.25
2. Participant gender		—	.43	.50	−.03	−.31	.04	.06	−.19	−.04	.35	−.47	.39	.03	.43	.66	.08	.21	.28
3. Participant ethnicity: % Black/Hispanic			—	.49	.06	.05	.21	.10	−.12	.12	.37	−.25	.27	.00	.42	.44	.23	.08	.03
4. Risk status of participants				—	−.15	−.28	.05	−.13	−.37	.08	.09	−.29	.27	−.13	.45	.54	.32	.14	.24
5. Intervention duration: Hr					—	.50	−.04	−.05	−.05	.12	−.08	.19	−.10	−.26	−.10	−.08	.11	−.29	−.12
6. Intervention duration: Weeks						—	.06	.13	.33	−.29	−.10	.51	−.16	−.32	−.31	−.31	−.15	.03	−.32
7. Parental involvement							—	.03	.09	−.11	.12	−.07	.43	.22	.23	.28	−.15	−.01	.04
8. Psychoeducational content								—	.17	−.30	.02	.04	−.17	−.19	−.32	−.09	−.03	−.01	−.24
9. Dietary improvement									—	−.13	−.11	.31	.07	.16	−.22	−.20	−.23	.01	−.15
10. Physical activity increase										—	.04	−.18	−.12	.36	.15	.02	.28	−.08	−.06
11. Reduced sedentary behavior											—	−.33	.23	.14	.28	.31	.20	−.22	.22
12. No. of behavioral targets												—	−.29	−.31	−.51	−.49	−.28	−.00	−.38
13. Teachers vs. interventionist													—	.32	.37	.63	.08	.05	.11
14. Didactic vs. interactive														—	.18	.26	.11	.12	.10
15. Pilot study															—	.60	.16	−.01	.10
16. Recruitment method																—	.28	.22	.40
17. Random assignment																	—	.10	−.04
18. Nested data modeled incorrectly																		—	.13
19. Effect size																			—

Note. Correlations greater than ±.24 are significant at the .05 level.

more often used self-selected recruitment methods ($r = .60$), which attenuated the unique effects of the moderators.

Self-selected recruitment also showed a significant relation to the intervention effect sizes in both the univariate and multivariate models. Theoretically, this effect emerged because self-presenting participants are more motivated to engage in the program and more likely to make the recommended lifestyle changes, which contributes to the larger effect sizes observed for trials that used this recruitment method.

It was also noteworthy that a number of factors that have been hypothesized to moderate obesity prevention program effects, such as mandated improvements in diet and exercise, sedentary behavior reduction, parental involvement, and delivery by trained professional interventionists (vs. teachers), were not significantly related to larger effect sizes. This did not appear to be simply a function of limited statistical power, because on the basis of the procedures described by Hedges and Pigott (2004) for mixed-effects regression models, we had a power of .54 to detect a small effect ($r = .10$), a power of .89 to detect a medium effect ($r = .30$), and a power of greater than .99 to detect a large effect ($r = .50$). These calculations were based on two-tailed inferential tests and an assumed variance of .1, which was a conservative value that exceeded all observed error variances. The effect sizes in Table 8 confirm that we had sufficient effects to detect medium effects. Moreover, the effect sizes for certain moderators, such as participant ethnicity, participant risk status, parental involvement, physical activity increase, use of random assignment, and modeling nested data incorrectly were so small that it is unlikely that limited power accounts for these null effects. Nonetheless, the fact that the effects for other moderators, such as psychoeducational content and reduced sedentary behavior, were somewhat larger suggests that future studies should continue to investigate these potential moderators.

Another possible explanation for the null effects for dietary and physical activity changes is titration, wherein participants in interventions that directly change dietary intake or activity level in schools may compensate for these changes by altering their dietary intake and activity level at other times. In support of this speculation, Donnelly et al. (1996) found that an intervention that mandated increased physical activity during school resulted in increased activity during school, but a significant decrease in activity outside of school. If participants compensate for such mandated health behavior changes by making alterations outside of the program to keep them at a particular energy intake or energy expenditure level, these mandated changes may not produce differential effects relative to interventions without mandated behavioral change. Another implication of this possibility is that it is vital to measure behavioral change that may offset any positive behavioral change that occurs during the intervention.

Directions for Future Obesity Prevention Efforts

The final aim of this review is to explore directions for future research. First, the finding that most obesity prevention programs that have been evaluated did not produce significant weight gain prevention effects suggests that it will be vital to conduct follow-up trials of enhanced versions of the programs that produced significant weight gain prevention effects and to design new programs that build upon those that worked. This will also provide

Table 8
Univariate Regression Models for Individual Moderators

Moderator	B	$SE\ B$	β	Model R^2
Participant age				
Linear term	−0.24***	0.06	−2.02	0.29
Quadratic term	0.28***	0.06	2.24	
Participant gender	0.10**	0.04	0.30	0.09
Participant ethnicity: % Black/Hispanic	0.01	0.01	0.06	0.00
Risk status of participants	0.02	0.03	0.10	0.01
Intervention duration				
Hr	−0.01	0.01	−0.06	0.00
Weeks	−0.03*	0.01	−0.26	0.07
Parental involvement	0.01	0.03	0.03	0.00
Psychoeducational content	−0.06	0.04	−0.20	0.04
Dietary improvement	−0.04	0.04	−0.13	0.02
Physical activity increase	−0.01	0.03	−0.04	0.00
Reduced sedentary behavior	0.09	0.05	0.21	0.04
No. of behavioral targets	−0.04**	0.01	−0.33	0.11
Teachers vs. interventionist	0.02	0.03	0.10	0.01
Didactic vs. interactive	0.04	0.03	0.16	0.03
Pilot study	0.12**	0.04	0.35	0.12
Recruitment method	0.12**	0.04	0.37	0.13
Random assignment	−0.02	0.03	−0.09	0.01
Nested data modeled incorrectly	0.02	0.03	0.11	0.01

Note. The parameter estimates of the linear and quadratic effect for age of participants were drawn from a model containing both these terms. All other parameter estimates were drawn from univariate models.
* $p < .05$.   ** $p < .01$.   *** $p < .001$.

an opportunity to conduct independent replications of the most successful obesity prevention programs, which is a necessary component to establishing that these interventions are efficacious (APA, 1995).

Second, it will also be important to determine how to better design obesity prevention programs for populations that generally did not derive weight gain prevention effects from extant programs, such as preadolescents and males. Unless efficacious prevention programs are developed for a broad array of participants,

Figure 2. Relation of the average age of participants to the weight gain intervention effect sizes.

it will be difficult for obesity prevention efforts to achieve a meaningful reduction in obesity at the population level.

Third, it will be important for future trials to address methodological limitations of prior trials. Future trials should use random assignment to condition, blinded assessment procedures, direct measures of body fat, and procedures that minimize attrition. Additionally, it would be desirable if they used active control groups, rather than the assessment-only or waitlist control conditions that are commonly used, because these latter control conditions do not rule out the possibility that demand characteristics, expectancy effects, or attention contribute to any apparent intervention effects. Showing that a prevention program outperforms an active placebo control condition or alternative active intervention is also necessary for establishing that a program is efficacious (APA, 1995).

Future trials should also include multiyear follow-ups to ensure that any intervention effects persist beyond the termination of the intervention, as most programs that produced weight gain prevention effects used only pre–post designs. Given that the vast majority of individuals who show successful weight loss in obesity treatment programs regain the lost weight a few years after treatment termination (Jeffery et al., 2000), it is possible that obesity prevention effects likewise erode over time. The fact that three interventions produced weight gain prevention effects that persisted over follow-up suggests it is possible to arm individuals with the skills necessary to avoid unhealthy weight gain in the future, though it may be necessary to offer obesity prevention programs at multiple developmental periods to maximize weight gain prevention effects. It will also be important to test whether there were actually lower rates of onset of clinically significant weight gain (e.g., obesity onset), which is an outcome that is both clinically meaningful and more consistent with the concept of prevention
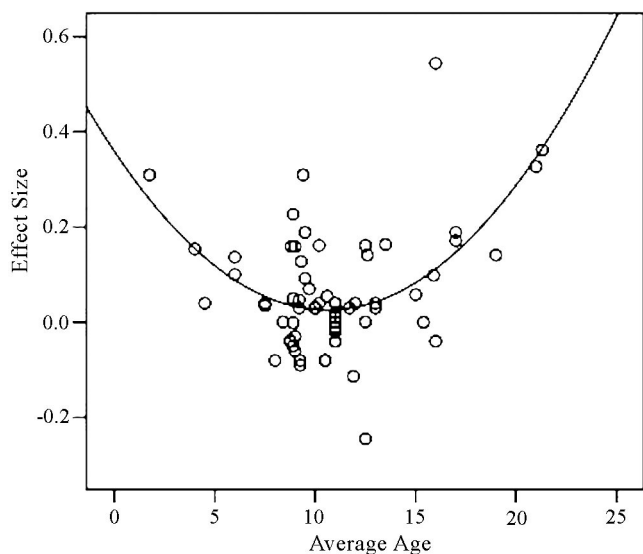
than reductions in average weight gain. Very few past trials have examined this outcome.

It will also be vital to evaluate the mediators that putatively account for any weight gain prevention effects. If the intervention produces change in putative mediators, but no weight gain prevention effects, or produces weight gain prevention effects, but the mediators do not change, this signals that the intervention model may be incorrect or that certain measures are unreliable or invalid. The fact that many obesity prevention programs reported significant effects for reductions in self-reported caloric intake and increases in self-reported exercise, but no significant effects for change in body mass, raises questions about the veracity of these self-report outcomes, as a true reduction in caloric intake or increase in physical expenditure should be accompanied by concomitant changes in BMI. However, it will be difficult to address these questions because extant measures of intake and activity level have limited validity, may not be sufficiently sensitive to detect the small changes in eating and activity promoted in most interventions, and may be too expensive for routine use in large-scale randomized trials (e.g., the double labeled water method of assessing energy intake).

Finally, the consistency regarding the limited returns of health behavior change prevention programs suggests that there is a need to develop and evaluate general theories regarding resistance to health behavior change. Such theories have the potential of increasing the return of obesity prevention programs as well as prevention programs aimed at other health behaviors. The behavioral economics model of obesity (Epstein & Saelens, 1999) seems particularly well suited to understand resistance to change as it expressly recognizes that behavior is a result of the balance between benefits and costs of the behavior. An improved understanding of the benefits of overeating and a sedentary lifestyle may imply ways to overcome barriers to health behavior change. The behavioral economics model also posits that there are individual difference factors that may cause some people to obtain more reinforcement from eating and less reinforcement from exercise, thereby increasing their risk for weight gain. This perspective suggests that it would be advantageous to measure such individual differences and adapt the interventions to better address the specific needs of various subpopulations. In addition, the assertion that health behavior arises from the balance between the benefits versus costs of the behavior identifies a key challenge of preventing obesity—the costs of this behavior will be experienced in the future, whereas the benefits of overeating and sedentary behavior occur in the present. The fact that most youth are more oriented to the current benefits of a lifestyle involving a positive energy balance rather than potential future costs may make obesity prevention a particularly challenging target for prevention programs delivered to children and adolescents.

Understanding more fully the barriers to making health behavior changes may also help improve obesity prevention programs. Research suggests that internal barriers to change, such as a lack of willpower and the perception that one is too busy to make healthy changes, predict failed attempts to change diet and exercise behaviors (Ziebland, Thorogood, Yudkin, Jones, & Coulter, 1998). It has also been suggested that difficulty in impulse control and a denial of the consequences of unhealthy behaviors undermine health change efforts (Sjoberg, 2003).

Another explanation for the relatively modest impact of extant obesity prevention programs is that environmental factors, such as the availability of high fat foods and a scarcity of pleasant places to exercise in many communities, play a key role in obesity promotion (Wadden et al., 2002). If this model is correct, it will be important to attempt to directly manipulate these environmental factors in future obesity prevention trials.

Conclusions

In sum, this meta-analytic review suggests that most interventions do not produce the hypothesized weight gain prevention effects and that the overall average intervention effect was small. Findings also indicated that for most programs that produced significant weight gain prevention effects, the effect sizes are clinically meaningful but are usually confined to pre–post effects. Additionally, results indicated that several prevention programs targeting a variety of health behaviors, such as eating pathology and smoking, produced weight gain prevention effects. These findings are encouraging because they suggest that there may be many efficacious approaches to reducing risk for weight gain and because some of these interventions produce intervention effects for multiple health behaviors. Results did not provide support for several factors that have been hypothesized to differentiate effective from ineffective prevention programs, but they did suggest that larger weight gain prevention effects were observed for programs targeting children and adolescents (vs. preadolescents), females, and self-presenting samples; programs that were relatively brief; programs solely targeting weight control versus other health behaviors (e.g., hypertension); and programs evaluated in pilot trials. Future trials should follow up promising findings and address methodological limitations of this literature (including the scarcity of long-term follow-up). Although significant progress has been made with regard to preventing this burgeoning public health problem, considerable work lies ahead.

References

Alexandrov, A. A., Maslennikova, G. Y., Kulikovm, S. M., Propirnij, G. A., & Perova, N. V. (1992). Primary prevention of cardiovascular diseases: 3-year intervention results in boys of 12 years of age. *Preventive Medicine, 21,* 53–62.

American Psychological Association. (1995). *Template for developing guidelines: Interventions for mental disorders and psychological aspects of physical disorders.* Washington, DC: Author.

Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist, 55,* 469–480.

Baldwin, S. A., Murray, D. M., & Shadish, W. R. (2005). Empirically supported treatments or Type I errors? Problems with the analysis of data from group-administered treatments. *Journal of Consulting and Clinical Psychology, 73,* 924–935.

Bandini, L. G., Schoeller, D. A., Dyr, H., & Dietz, W. H. (1990). Validity of reported energy intake in obese and nonobese adolescents. *American Journal of Clinical Nutrition, 52,* 421–425.

Baranowski, T., Baranowski, J., Cullen, K. W., Thompson, D. I., Nicklas, T., Zakeri, I., et al. (2003). The fun, food, and fitness project (FFFP): The Baylor GEMS pilot study. *Ethnicity and Disease, 13*(Suppl. 1), S1–30–S1–39.

Baranowski, T., Cullen, K., Nicklas, T., Thompson, D., & Baranowski, J.

(2002). School-based obesity prevention: A blueprint for taming the epidemic. *American Journal of Public Health, 26,* 486–493.

Bayne-Smith, M., Fardy, P. Azzollini, A., Magel, J., Schmitz, K., & Again, D. (2004). Improvements in heart health behaviors and reduction in coronary artery disease risk factors in urban teenaged girls through a school-based intervention: The PATH program. *American Journal of Public Health, 94,* 1538–1543.

Beech, B. M., Klesges, R. C., Kumanyika, S. K., Murray, D. M., Klesges, L., McClanahan, B., et al. (2003). Child and parent targeted interventions: The Memphis GEMS pilot study. *Ethnicity and Disease, 13*(Suppl. 1), S1–40–S1–53.

Burke, G. L., & Bild, D. E. (1996). Differences in weight gain in relation to race, gender, age and education in young adults: The CARDIA study. *Ethnicity & Health, 1,* 327–335.

Burke, V., Milligan, R. A. K., Thompson, C., Taggart, A. C., Dunbar, D. L., Spencer, M. J., et al. (1998). A controlled trial of health promotion programs in 11-year-olds using physical activity "enrichment" for higher risk children. *The Journal of Pediatrics, 132,* 840–848.

Bush, P. J., Zuckerman, A. E., Theiss, P. K., Taggart, V. S., Horowitz, C., Sheridan, M. J., et al. (1989). Cardiovascular risk factor prevention in Black school children: Two year results of the "know your body" program. *American Journal of Epidemiology, 129,* 466–482.

Calle, E., Thun, M., Petrelli, J., Rodriguez, C., & Heath, C. (1999). Body mass index and mortality in a prospective cohort of US adults. *New England Journal of Medicine, 341,* 1097–1105.

Campbell, K., Waters, E., O'Meara, S., Kelly, S., & Summerbell, C. (2003). Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews, 1,* 1–59.

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 and Adolescent Psychiatry, 34,* 312–321.

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

Dietz, W. H. (1998). Childhood weight affects adult morbidity and mortality. *Journal of Nutrition, 128*(Suppl.), 411S–414S.

Dietz, W. H., & Gortmaker, S. L. (2001). Preventing obesity in children and adolescents. *Annual Review of Public Health, 22,* 337–353.

Dietz, W. H., & Robinson, T. N. (1998). Use of body mass index (BMI) as a measure of overweight in children and adolescents. *The Journal of Pediatrics, 132,* 191–193.

Donnelly, J. E., Jacobsen, D. J., Whatley, J. E., Hill, J. O., Swift, L. L., Cherrington, A., et al. (1996). Nutrition and physical activity program to attenuate obesity and promote physical and metabolic fitness in elementary school children. *Obesity Research, 4,* 229–243.

Duncan, G., Anton, S., Newton, R., & Perri, M. (2003). Comparison of perceived health to physiological measures of health in Black and White women. *Preventive Medicine, 36,* 624–628.

Dwyer, T., Coonan, W. E., Leitch, D. R., Hetzel, B. S., & Baghurst, R. A. (1983). An investigation of the effects of daily physical activity on the health of primary school students in South Australia. *International Journal of Epidemiology, 12,* 308–313.

Eliakim, A., Makowski, G. S., Brasel, J. A., & Cooper, D. M. (2000). Adiposity, lipid levels, and brief endurance training in nonobese adolescent males. *International Journal of Sports Medicine, 21,* 332–337.

Ellis, K. J., Abrams, S. A., & Wong, W. W. (1999). Monitoring childhood obesity: Assessment of the weight/height2 index. *American Journal of Epidemiology, 150,* 939–946.

Epstein, L. H., Gordy, C. C., Raynorm, H. A., Beddome, M., Kilanowski, C. R., & Paluch, R. (2001). Increasing fruit and vegetable intake and decreasing fat and sugar intake in families at risk for childhood obesity. *Obesity Research, 9,* 171–178.

Epstein, L. H., Paluch, R. A., & Raynor, H. A. (2001). Sex differences in

obese children and siblings in family-based obesity treatment. *Obesity Research, 9,* 746–753.

Epstein, L. H., & Saelens, B. E. (1999). Behavioral economics of obesity: Food intake and energy expenditure. In W. K. Bickel & R. E. Vuchinich (Eds.), *Reframing health behavior change with behavioral economics* (pp. 293–311). Mahwah, NJ: Erlbaum.

Epstein, L., Valoski, A., Wing, R., & McCurley, J. (1990). Ten-year follow-up of behavioral family-based treatment for obese children. *Journal of the American Medical Association, 264,* 2519–2523.

Epstein, L., Wing, R. R., Koeske, R., & Valoski, A. (1987). Long-term effects of family-based treatment of childhood obesity. *Journal of Consulting and Clinical Psychology, 55,* 91–95.

Fardy, P. S., White, R. E., Hjaltiwanger-Schmitz, K., Magel., J. R., McDermott, K. J., Clark, L. T., et al. (1996). Coronary disease risk factor reduction and behavior modification in minority adolescents: The PATH program. *Journal of Adolescent Health, 18,* 247–253.

Fitzgibbon, M., Stolley, M., Schiffer, L., Van Horn, L., Kaufer-Christoffel, K., & Dyer, A. (2004). Two-year follow-up results for the Hip-Hop to Health Jr.: A randomized controlled trial for overweight prevention in preschool minority children. *Journal of Pediatrics, 146,* 618–625.

Flodmark, C., Ohlsson, T., Ryden, O., & Sveger, T. (1993). Prevention of progression to severe obesity in obese schoolchildren treated with family therapy. *Pediatrics, 91,* 880–884.

Flores, R. (1995). Dance for health: Improving fitness in African and Hispanic adolescents. *Public Health Reports, 110,* 189–193.

Freedman, D. S., & Perry, G. (2000). Body composition and health status among children and adolescents. *Prevention Medicine, 31,* S34–S53.

French, S. A., Perry, C. L., Leon, G. R., & Fulkerson, J. A. (1994). Dieting behaviors and weight change history in female adolescents. *Health Psychology, 14,* 548–555.

Golan, M., Weizman, A., Apter, A., & Fainaru, M. (1998). Parents as the exclusive agents of change in the treatment of childhood obesity. *American Journal of Clinical Nutrition, 67,* 1130–1135.

Gortmaker, S. L., Peterson, K., Wiecha, J., Sobol, A. M., Dixit, S., Fox, M. K., et al. (1999). Reducing obesity via a school-based interdisciplinary intervention among youth. *Archives of Pediatric Adolescent Medicine, 153,* 409–418.

Gutin, B., Cucuzzo, N., Islam, S., Smith, C., Moffatt, R., & Pargman, D. (1995). Physical training improves body composition of Black obese 7- to 11-year old girls. *Obesity Research, 3,* 305–312.

Gutin, B., & Owens, S. (1999). Role of exercise intervention in improving body fat distribution and risk profile in children. *American Journal of Human Biology, 11,* 237–247.

Hansen, H. S., Froberg, K., Hyldebrandt, N., & Nielsen, J. R. (1991). A controlled study of eight months of physical training and reduction of blood pressure in children: The Odense schoolchild study. *British Medical Journal, 303,* 682–685.

Harrell, J. S., Gansky, S. A., McMurray, R. G., Bangdiwala, S. I., Frauman, A. C., & Bradley, C. B. (1998). School-based interventions improve heart health in children with multiple cardiovascular disease risk factors. *The Journal of Pediatrics, 102,* 371–380.

Harrell, J. S., McMurray, R. G., Bangdiwala, S. I., Grauman, A. C., Gansky, S. A., & Bradley, C. B. (1996). Effects of a school-based intervention to reduce cardiovascular disease risk factors in elementary-school children: The Cardiovascular Health in Children (CHIC) study. *Pediatrics, 128,* 797–805.

Harvey-Berino, J., & Rourke, J. (2003). Obesity prevention in preschool Native-American children: A pilot study using home visiting. *Obesity Research, 11,* 606–611.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9,* 426–445.

Hedley, A. A., Odgen, C. L., Johnson, C. L., Carroll, M. D., Curtin, L. R.,

& Flegal, K. M. (2004). Prevalence of overweight and obesity among US children, adolescents, and adults, 1999–2000. *Journal of the American Medical Association, 291,* 2847–2850.

Helweg-Larsen, M., & Collins, B. E. (1997). A social psychological perspective on the role of knowledge about AIDS in AIDS prevention. *Current Directions, 6,* 23–26.

Hopper, C. A., Munoz, K. D., Gruber, M. B., MacConnie, S., Schonfeldt, B., & Shunk, T. (1996). A school-based cardiovascular exercise and nutrition program with parent participation: An evaluation study. *Children's Health Care, 25,* 221–235.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.). New York: Wiley.

Hwang, M. S., Yeagley, K. L., & Petosa, R. (2004). A meta-analysis of adolescent psychosocial smoking prevention programs published between 1978 and 1997 in the United States. *Health Education and Behavior, 31,* 702–719.

Ignico, A. A., & Mahon, A. D. (1995). The effects of a physical fitness program on low-fit children. *Research Quarterly for Exercise and Sport, 66,* 85–90.

Jeffery, R., Drenowski, A., Epstein, L. H., Stunkard, A. J., Wilson, G. T., Wing, R. R., & Hill, D. (2000). Long-term maintenance of weight loss: Current status. *Health Psychology, 19,* 5–16.

Kain, J., Uauy, R., Vio, F., Cerda, R., & Leyton, B. (2004). School-based obesity prevention in Chilean primary school children: Methodology and evaluation of a controlled study. *International Journal of Obesity, 28,* 483–493.

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.

Killen, J. D., Telch, M. J., Robinson, T. N., Maccoby, N., Taylor, C. B., & Farquar, J. W. (1988). Cardiovascular disease risk reduction for tenth graders: A multiple-factor school-based approach. *Journal of the American Medical Association, 260,* 1728–1733.

Kimm, S. Y., Barton, B., Obarzanek, E., McMahon, R. P., Sabry, Z. I., Waclawiw, M. A., et al. (2001). Racial divergence in adiposity during adolescence: The NHLBI Growth and Health Study. *Pediatrics, 107,* e34–e40. Retrieved February 18, 2003, from http://pediatrics.aappublications.org/cgi/content/full/107/3/e34

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.

Lichtman, S. W., Pisarska, K., Berman, E. R., Pestone, M., Dowling, H., Offenbacher, E., et al. (1992). Discrepancy between self-reported and actual caloric intake and exercise in obese subjects. *New England Journal of Medicine, 327,* 1893–1898.

Lionis, C., Kafatos, A., Vlachonikolis, J., Vakaki, M., Tzortzi, M., & Petraki, A. (1991). The effects of a health education intervention program among Cretan adolescents. *Preventive Medicine, 20,* 685–699.

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

Logan, T. K., Cole, J., & Leukefeld, C. (2002). Women, sex, and HIV: Social and contextual factors, meta-analysis of published interventions, and implications for practice and research. *Psychological Bulletin, 128,* 851–885.

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.

Luepker, R. V., Perry, C. L., McKinlay, S. M., Nader, P. R., Parcel, G. S., Stone, E. J., et al. (1996). Outcomes of a field trial to improve children's dietary patterns and physical activity: The child and adolescent trial for cardiovascular health (CATCH). *Journal of the American Medical Association, 275,* 768–776.

Magarey, A. M., Daniels, L. A., Boulton, T. J., & Cockington, R. A. (2003). Predicting obesity in early adulthood from childhood and parental obesity. *International Journal of Obesity and Related Metabolic Disorders, 27,* 505–513.

Manios, Y., Moschandreas, J., Hatzis, C., & Kafatos, A. (2002). Health and nutrition education in primary schools in Crete: Changes in chronic disease risk factors following a 6-year intervention program. *British Journal of Nutrition, 88,* 315–324.

Matheson, D., Robinson, T., & Killen, J. (2005). *Obesity prevention in Mexican-American families.* Unpublished manuscript, Stanford University, CA.

Matvienko, O., Lewis, D., & Schafer, E. (2001). A college nutrition science course as an intervention to prevent weight gain in female college freshman. *Society for Nutrition Education, 33,* 95–101.

McMurray, R. G., Harrell, J. S., Bangdiwala, S. I., Bradley, C. B., Deng, S., & Levine, A. (2002). A school-based intervention can reduce body fat and blood pressure in young adolescents. *Journal of Adolescent Health, 31,* 125–132.

Mo-suwan, L., Pongprapai, S., Junjana, C., & Puetpaiboon, A. (1998). Effects of a controlled trial of a school-based exercise program on the obesity indexes of preschool children. *American Journal of Clinical Nutrition, 68,* 1006–1011.

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.

Nader, P. R., Sallis, J., Abramson, I., Broyles, S., Patterson, T. L., Senn, K., et al. (1992). Family-based cardiovascular risk reduction education among Mexican and Anglo-Americans. *Family and Community Health, 15,* 57–74.

Neumark-Sztainer, D., Story, M. N., Hannan, P. J., & Rex, J. (2003). New moves: A school-based obesity prevention program for adolescent girls. *Preventive Medicine, 37,* 41–51.

Perry, C. L., Bishop, D. B., Taylor, G., Murray, D. M., Mays, R. W., Dudovitz, B. S., et al. (1998). Changing fruit and vegetable consumption among children: The 5-a-day power plus program in St. Paul, MN. *American Journal of Public Health, 88,* 603–609.

Puska, P., Vartianinen, E., Pallonen, U., Salonen, J. T., Poyhia, P., Koskela, K., et al. (1982). The North Karelia Youth Project: Evaluation of two years of intervention on health behavior and CVD risk factors among 13–15 year old children. *Preventive Medicine, 11,* 550–570.

Resnicow, K. (1993). School-based obesity prevention: Population versus high risk interventions. *Annals of the New York Academy of Sciences, 699,* 154–166.

Resnicow, K., Cohm, L., Reinhardt, J., Cross, D., Futterman, R., Kirschner, E., et al. (1992). A three-year evaluation of the Know Your Body program in inner-city schoolchildren. *Health Education Quarterly, 19,* 463–480.

Resnicow, K., & Robinson, T. N. (1997). School-based cardiovascular disease prevention studies: Review and synthesis. *Annals of Epidemiology, 7*(Suppl.), S14–S31.

Robinson, T. N. (1999). Reducing children's television viewing to prevent obesity: A randomized controlled trial. *Journal of the American Medical Association, 282,* 1561–1567.

Robinson, T. N., Killen, J. D., Kraemer, H. C., Wilson, D. N., Matheson, D. M., Haskell, W. L., et al. (2003). Dance and reducing television viewing to prevent weight gain in African-American girls: The Stanford GEMS pilot study. *Ethnicity and Disease, 13*(Suppl. 1), S1–65–S1–77.

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.

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

Sahota, P., Rudolf, M. C. J., Dixey, R., Hill, A. J., Barth, J. H., & Cade, J. (2001). Randomised controlled trial of primary school based intervention to reduce risk factors for obesity. *British Medical Journal, 323,* 1–5.

Sallis, J. F., McKenzie, T. L., Alcaraz, J. E., Kolody, B., Hovell, M. F., & Nader, P. R. (1993). Project SPARK: Effects of physical education and adiposity in children. *Annals of the New York Academy of Sciences, 699,* 127–136.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177.

Schmitz, M. K., & Jeffery, R. W. (2000). Public health intervention for the prevention and treatment of obesity. *Medical Clinics of North America, 84,* 491–512.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. New York: Houghton Mifflin.

Sjoberg, L. (2003). Neglecting the risks: The irrationality of health behavior and the quest for *la dolce vita. European Psychologist, 8,* 266–278.

Skara, S., & Sussman, S. (2003). A review of 25 long-term adolescent tobacco and other drug use prevention program evaluations. *Preventive Medicine, 37,* 451–474.

Solomon, C. G., & Manson, J. E. (1997). Obesity and mortality: A review of the epidemiologic data. *American Journal of Clinical Nutrition, 66*(Suppl.), 1044s–1050s.

Stice, E., Orjada, K., & Tristan, J. (2006). Trial of a psychoeducational eating disturbance intervention for college women: A replication and extension. *International Journal of Eating Disorders, 39,* 233–239.

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

Stice, E., & Shaw, H. (2004). Eating disorder prevention programs: A meta-analytic review. *Psychological Bulletin, 130,* 206–227.

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

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.

Story, M. (1999). School-based approaches for preventing and treating obesity. *International Journal of Obesity, 23*(Suppl. 2), S43–S51.

Story, M., Sherwood, N. E., Himes, J. H., Davis, M., Jacobs, D. R., Cartwright, Y., et al. (2003). An after-school obesity prevention program for African-American girls: The Minnesota GEMS pilot study. *Ethnicity and Disease, 13*(Suppl.), S1–54–S1–64.

Tamir, D., Feurstein, A., Brunner, S., Halfon, S., Reshef, A., & Palti, H. (1990). Primary prevention of cardiovascular diseases in childhood: Changes in serum total cholesterol, high density lipoprotein, and body mass index after 2 years of intervention in Jerusalem schoolchildren age 7–9 years. *Preventive Medicine, 19,* 22–30.

Thompson, J. K., Heinberg, L. J., Altabe, M., & Tantleff-Dunn, S. (1999). *Exacting beauty: Theory, assessment, and treatment of body image disturbance.* Washington, DC: American Psychological Association.

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.

Vandongen, R., Jenner, D. A., Thompson, C., Taggart, A. C., Spickett, E. E., Burke, V., et al. (1995). A controlled evaluation of a fitness and nutrition intervention program on cardiovascular health in 10–12 year old children. *Preventive Medicine, 24,* 9–22.

Wadden, T. A., Brownell, K. D., & Foster, G. D. (2002). Obesity: Responding to the global epidemic. *Journal of Consulting and Clinical Psychology, 70,* 510–525.

Wadden, T. A., Vogt, R. A., Foster, G. D., & Anderson, D. A. (1998). Exercise and the maintenance of weight loss: 1-year follow-up of a controlled clinical trial. *Journal of Consulting and Clinical Psychology, 66,* 429–433.

Walter, H. J., Hofman, A., Vaughan, R. D., & Wynder, E. L. (1988). Modification of risk factors for coronary heart disease: Five year results of a school-based intervention trial. *New England Journal of Medicine, 318,* 1093–1100.

Willet, S. L. (1995). *Health risk reduction among African-American females: A one year follow-up.* Unpublished doctoral dissertation, Northwestern University, Evenston, IL.

Willett, J. B., & Singer, J. D. (1993). Investigating onset, cessation, relapse, and recovery: Why you should, and how you can, use discrete-time survival analysis to examine event occurrence. *Journal of Consulting and Clinical Psychology, 61,* 952–965.

Wolf, A. M. (1998). What is the economic cost for treating obesity? *Obesity Research, 6*(Suppl. 1), 2S–7S.

Ziebland, S., Thorogood, M., Yudkin, P., Jones, L., & Coulter, A. (1998). Lack of willpower or lack of wherewithal? Internal and external barriers to changing diet and exercise in a three year follow-up of participants in a health check. *Society, Science, and Medicine, 46,* 461–465.