EATING DISORDER PREVENTION RESEARCH: A META-ANALYSIS

A Dissertation

by

MICHELLE CORORVE FINGERET

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2004

Major Subject: Psychology

EATING DISORDER PREVENTION RESEARCH: A META-ANALYSIS

A Dissertation

by

MICHELLE CORORVE FINGERET

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Approved as to style and content by:

David Gleaves (Chair of Committee)	Mary Meagher (Member)
Emily Davidson (Member)	Linda Castillo (Member)
Steve Rholes (Head of Department)	

May 2004

Major Subject: Psychology

ABSTRACT

Eating Disorder Prevention Research: A Meta-Analysis. (May 2004)

Michelle Cororve Fingeret, B.A., Washington University; M.S., Texas A&M University

Chair of Advisory Committee: Dr. David Gleaves

The purpose of this study was to quantitatively evaluate the overall effectiveness of eating disorder prevention programs and to investigate potential moderating variables that may influence the magnitude of intervention effects. Meta-analysis was used to conduct a comprehensive and systematic analysis of data across 46 studies. Effect size estimates were grouped into outcome sets based on the following variables: knowledge, general eating pathology, dieting, thin-ideal internalization, body dissatisfaction, negative affect, and self-esteem. *Q* statistics were used to analyze the distribution of effect size estimates within each outcome set and to explore the systematic influence of moderating variables. Results revealed large effects on the acquisition of knowledge and small net effects on reducing maladaptive eating attitudes and behaviors at posttest and follow-up. These programs were not found to produce significant effects on negative affect, and there were inconsistent effects on self-esteem across studies. Population targeted was the sole moderator that could account for variability in effect size distributions. There was a tendency toward greater benefits for studies targeting participants considered to be at a relatively higher risk for developing an eating disorder. Previous assumptions regarding the insufficiency of “one-shot” interventions and

concerns about the iatrogenic effects of including information about eating disorders in an intervention were not supported by the data. These findings challenge negative conclusions drawn in previous review articles regarding the inability of eating disorder prevention programs to demonstrate behavioral improvements. Although these findings have implications for the prevention of eating disorders, it was argued that a clear link between intervention efficacy and a decreased incidence of eating disorders was not demonstrated. Rather, only direct information was offered about the ability to influence eating disorder related knowledge, attitudes, and behaviors. Specific recommendations related to intervention content, reasonable goals/expectations, and outcome criteria were offered for improving research in this area.

TABLE OF CONTENTS

Page

LIST OF TABLES

INTRODUCTION

Increased attention to the prevention of eating disorders has resulted in a rapidly growing body of research. Empirical studies of eating disorder preventive interventions are being published at a steady rate, in part due to the emergence of a journal largely devoted to this subject matter. *Eating Disorders: The Journal of Treatment and Prevention* was created in 1993 to contend with increased interest by the scientific community in the prevention of these disturbances and to report new developments in the area of treatment. Since the emergence of this journal, a concomitant increase in eating disorder prevention studies has been found in other scholarly journals, and there has been a substantial rise in the number of dissertations devoted to developing and testing preventive interventions in this area.

Given the serious health concerns and myriad of harmful physiological and psychological sequelae that frequently accompany eating disorders, increased attention to the prevention of these disturbances seems warranted. Agras (2001) highlighted the costs and consequences associated with eating disorders as revealed by medical complications, comorbid psychopathology, and financial costs of inpatient hospitalization and outpatient treatment. Anorexia nervosa has been reported to have the highest mortality rate of any psychiatric disorder with rates of 5% and up to 18% being recorded (Agras, 2001; Cassell & Gleaves, 2000). Although the majority of deaths are related to physical complications of the disorder (Gary, 2001), death also results from suicide. The mortality rate alone indicates the need to develop and implement successful preventive interventions for these disturbances. Moreover, concerns have been raised

---

This dissertation follows the style and format of *Psychological Bulletin*.

regarding the degree to which these disorders can be obscured by secrecy, shame, guilt, and denial with many sufferers attempting to conceal their symptoms or refusing to acknowledge the existence of a problem. Garvin and Streigel-Moore (2001) reviewed research suggesting that many individuals with eating disorders do not seek treatment or are not being identified or referred for treatment. These authors also pointed to changes in the health care system that have resulted in reduced access to and limited insurance coverage for eating disorder care. These findings further highlight the importance of conducting prevention research in this field.

With the rapid growth of eating disorder prevention research, efforts to integrate the findings across studies become increasingly important. Findings from individual studies in this area have been widely divergent. Although many researchers reported being pleased with the results they obtained and concluded that the intervention tested was effective and successful in achieving its intended goals (e.g., Irving, 2000; O'Dea & Abraham, 2000; Stice & Ragan, 2002), other researchers concluded that the intervention was unsuccessful or ineffective (e.g., Killen et al., 1993, Nelson, 1996; Paxton, 1993). Of particular concern are the few studies that reported a worsening of eating disorder symptomatology following an intervention (Carter, Stewart, Dunn & Fairburn, 1997; Mann, Nolen-Hoeksema, & Burgard, 1997; Wolfe, 1992). Based on individual findings in this field, a controversy has ensued regarding the degree to which eating disorder preventive efforts are wholly ineffective or perhaps even harmful. This controversy has led to arguments for eliminating didactic material on eating disorders from interventions

(e.g., O'Dea & Abraham, 2000; Rosenvinge & Borresen, 1999) and to focus efforts only on high-risk populations (Killen et al., 1993).

Although a few individual research studies have reported a worsening of eating disorder symptoms following an intervention (Carter et al., 1997; Mann et al., 1997; Wolfe, 1992), it is widely recognized that single studies offer limited information on a given research question. Schmidt (1996) pointed to the critical role of sampling error, measurement error, and other artifacts in determining the observed findings and the statistical power of individual studies. He argued that a single primary study can rarely resolve an issue or answer a question and rather should be considered as a data point to be contributed to a larger literature review. Integrating the findings across studies allows researchers to more effectively deal with conflicting findings. Within the eating disorder prevention field, it appears particularly important to deal with the contrasting and contradictory findings involving detrimental effects, given the potential implications of these results. Furthermore, a comprehensive understanding of the findings and the nature of the research being conducted in this field are vital to the continued advancement of this burgeoning area of research.

Previous Review Articles

Previous efforts to integrate findings across eating disorder prevention studies have consistently concluded that the results are disappointing and discouraging (Austin, 2000; Fairburn, 1995; Mussell, Binford, & Fulkerson, 2000). Several articles have focused on the apparent inability of programs to successfully meet prevention goals (Franko & Orosan-Weine, 1998; Mussell et al.; Rosenvinge & Borrensen, 1999). The

synopsis provided by Musell et al. is representative of this perspective: "With only a few exceptions, most of these programs have failed to demonstrate efficacy though improvements in or prevention of eating disordered-related behavior" (p. 277). These negative conclusions were drawn from qualitative reviews of the literature and were primarily based on the importance of behavioral outcome variables in determining intervention efficacy. As such, they are considered to be limited in scope. These conclusions do not necessarily reflect the degree to which interventions effectively influence other commonly studied indicators of outcome such as knowledge, attitudes, and affect. Moreover, the interpretation of findings and conclusions reached in narrative literature reviews have been widely criticized for being dependent upon subjective judgment. Wood (1995) argued that narrative reviews are particularly prone to bias, selective inclusions of evidence, and misinterpretation of research findings. He further commented that narrative reviews have difficulty dealing with conflicting findings because the manner in which findings are distorted by artifacts is not accounted for and conclusions from narrative reviews can be difficult to replicate when procedures for synthesizing the data have not been made explicit.

Although it appears clear that more rigorous methods are needed to integrate the data in this field, Fingeret (2002) provided preliminary evidence to challenge conclusions drawn in previous review articles. A narrative review of 59 published empirical studies of eating disorder preventive interventions revealed that positive effects on a variety of outcome variables were reported in the majority of studies. Specifically, 42 of 56 studies assessing attitudes reported positive attitudinal

improvements following an intervention, 30 of 47 studies assessing behavior reported positive behavior changes, 21 of 24 studies assessing knowledge reported improvements in this domain following an intervention, and 6 of 9 studies evaluating affect reported positive changes on relevant outcome measures. Moreover, 75% of the studies reviewed reported that study goals were wholly or partially met. Fingeret incorporated a broader array of studies compared to previous review articles including more recent studies utilizing innovative intervention strategies. Furthermore, the author explored trends in study methods and intervention techniques, and revealed that a variety of factors could potentially influence intervention efficacy. The four primary factors that were of interest to the current study were: length of intervention, type of intervention strategy, population targeted, and outcome variables. The manner in which each of these factors may moderate the magnitude of intervention effects in eating disorder prevention studies is further explored below.

Length of intervention. Across eating disorder prevention studies, the number of discrete sessions used to implement an intervention varies widely, and some studies employ ongoing and continuous programming efforts for several weeks, months or years. The degree to which an intervention implemented in a single session can have a meaningful impact on eating disorder related attitudes and behaviors has been strongly questioned (e.g., Martz & Bazinni, 1999). Interventions implemented over multiple sessions and those involving ongoing programming efforts are likely to be more intense in nature compared to "one-shot" prevention programs and offer more opportunity for internalizing intervention material. However, assumptions about the insufficiency of

single session programs and the superiority of multiple sessions or continuous programming efforts over single sessions are not currently supported by data. It is suggested that a more precise exploration of the relationship between intervention duration and intervention efficacy would have important implications for the guidance and dissemination of preventive interventions on a large-scale and community-wide basis.

Intervention strategies. There is also a great deal of variability regarding the types of intervention strategies employed across studies in this area. Although psychoeducation appears to be the most widely used intervention strategy, many studies enhance traditional psychoeducational approaches to eating disorder prevention by incorporating more interactive techniques. Experiential games and activities are commonly employed with younger populations to provide an interesting and exciting learning environment and to promote cooperative learning (e.g, Coller & Neumark-Sztainer, 1999). Another method for enhancing psychoeducational strategies is to incorporate skills-based techniques such as those used in the cognitive behavioral treatment of eating disorders (e.g., self-monitoring of eating habits, identifying and challenging cognitive distortions related to eating, shape, and weight, stress inoculation techniques). Media literacy training has emerged as a type of enhanced psychoeducational intervention strategy that combines didactic, interactive, and skills-based approaches to intervention. This intervention strategy involves teaching participants to become more active and critical consumers of appearance-related media

and providing tools for resisting pressures to conform to the images and messages conveyed in the media.

Concerns have been raised about the potential iatrogenic effects of a psychoeducational approach to eating disorder prevention. As discussed by O'Dea and Abraham (2000), "The information-giving approach has the potential to create adverse effects such as the glamorization and normalization of eating disorders and to introduce young people to dangerous practices by providing information about dangerous methods of weight control (starvation, vomiting, laxative abuse)" (p. 44). Mann et al. (1997) similarly warned that presenting information about eating disorders may inadvertently promote eating problems by reducing the stigma of these disorders. Thus, several researchers in this field have opted to eliminate psychoeducational material on eating disorders from their interventions. Self-esteem based approaches to eating disorder prevention have emerged as an alternative type of intervention strategy that aims to improve body image by targeting and building self-esteem rather than focusing on pathological eating behavior (e.g., McVey & Davis, 2002; O'Dea & Abraham, 2000). The degree to which interventions incorporating psychoeducational material on eating disorders may be harmful is an empirical question that merits exploration. Moreover, given the variability in intervention strategies, it would be valuable to understand the manner in which purely psychoeducational, enhanced psychoeducational, or purely interactive approaches (i.e., specifically eliminating a psychoeducational component) to intervention influence outcome. These findings would have important implications for guiding intervention strategies to be used in future studies.

Population targeted. The value of conducting untargeted eating disorder preventive interventions has begun to be questioned. For instance, Killen et al. (1993) concluded that it may be more worthwhile to focus on high-risk populations after they failed to find any meaningful differences following an 18 week untargeted intervention program with a 24 month follow-up period. Enthusiasm about the relative benefits of targeting individuals at higher risk for developing an eating disorder received support in the narrative review by Fingeret (2002). This study found striking trends regarding the relationship between population targeted and conclusions drawn about intervention efficacy. All of the interventions directed toward individuals with minimal signs of an eating disorder (i.e., indicated interventions) were reported to be at least partially or even wholly effective in altering attitudinal and behavioral risk factors for eating disorders, while 14% of untargeted interventions were concluded to be ineffective or harmful. However, these findings are limited by the narrative approach to integrating the data, and thus require further exploration. A more precise understanding of the relationship between population targeted and intervention efficacy in concert with findings related to optimal intervention strategies may offer important information related to customizing intervention strategies to population groups. These relationships have yet to be thoroughly explored in the eating disorder prevention literature.

Outcome variables. The selection and measurement of outcome variables are particularly important to examine given the manner in which they directly influence results and conclusions drawn about intervention efficacy. Although researchers in this field tend to rely heavily on attitudinal and behavioral outcome variables as indicators of

intervention efficacy, knowledge and affect also tend to be commonly assessed. As discussed in Fingeret (2002), attitudinal measures used across eating disorder prevention studies tap into an assortment of views about the self, one's body, and eating. Types of attitudes targeted in the reviewed studies included drive for thinness, body satisfaction, weight and shape concerns, internalization and societal appearance standards, and self-esteem. Behavioral measures typically examined type and frequency of weight control methods being used, food consumption, physical activity and media habits. Type of knowledge assessed involved topics specifically explored in the interventions (e.g., knowledge related to nutrition, the dangers of dieting, the symptoms and consequences of eating disorders, the use of deceptive media tactics). Type of affect assessed included anxiety, depression, or general distress.

Previous review articles have focused exclusively on behavioral outcome variables as indicators of intervention efficacy. However, it seems clear that data on a variety of outcome variables are available for exploration. There are unique implications for understanding the degree to which interventions can effectively influence specific outcome variables and whether they can affect a combination of these variables. It is also important to reiterate that previous review articles have not explored the precise magnitude of intervention effects on behavioral outcome variables, which is considered vital to determining the general utility of the prevention program.

Current Study

The primary purpose of the current study was to quantitatively evaluate, via meta-analysis, the effectiveness of eating disorder prevention programs. The investigator

aimed to provide a more comprehensive and systematic exploration of the findings from eating disorder prevention programs compared to previous review articles. One method for achieving this goal was to conduct a statistical analysis of results from a large number of individual research studies. Meta-analysis was used to conduct a comprehensive and systematic synthesis and analysis of data across a broad range of eating disorder prevention studies. The general strategy underlying this procedure is to quantify findings from a set of research studies in a standardized form that allows for meaningful numerical comparison and statistical analysis across studies. Various effect size statistics can be used to quantify study results in a manner that is consistent across all variables and measures involved.

Meta-analysis has been touted as a methodology that allows for study findings to be represented in a more accurate and sophisticated manner compared to conventional review procedures that rely on qualitative summaries. As discussed by Wood and Christensen (in press), conclusions reached in narrative reviews are often based on the statistical significance of individual study findings in the reviewed literature. These authors warned that such findings can be misleading, because reviewed studies that use small samples with insufficient power to detect effect are not adequately interpreted. This issue is particularly germane to the eating disorder prevention field where small sample sizes are frequently employed to test interventions and statistical significance is relied upon to interpret study findings. Meta-analysis allows for conclusions to be drawn based on the direction and magnitude of an effect apart from its statistical significance, given the emphasis on interpreting effect size indices. The sophisticated

nature of meta-analysis involves the ability to find effects or relationships between study findings and study features that are likely to be obscured by narrative approaches to summarizing research (Lipsey & Wilson, 2001). Key aspects of this procedure that allow for increased sophistication in examining study findings include using systematic decision rules and procedures, statistically synthesizing and combining data and effect sizes from studies, and exploring the extent and determinants of variability in effect sizes.

The application of these procedures to the eating disorder prevention literature is believed to offer a means of integrating the data in a more accurate, sophisticated, and organized manner than has been previously attempted. Meta-analysis procedures were used to resolve important questions in this field regarding the overall effectiveness of eating disorder prevention programs and whether there were specific factors that influenced the magnitude of intervention effects. The following were posed as explicit questions to be answered by the present study: a) Can one-session prevention programs have a meaningful impact on eating disorder related attitudes and behaviors? b) Are multiple session or continuous programming efforts intrinsically more efficacious compared to "one-shot" prevention programs? c) Does presenting psychoeducatonal material on eating disorders produce iatrogenic effects on eating attitudes and behaviors? d) Are targeted approaches to eating disorder prevention more beneficial than untargeted approaches? e) Which outcome variables are most affected by intervention efforts? f) To what degree can interventions effectively influence behavioral outcome variables?

METHOD

The present study was primarily guided by meta-analytic procedures described in Lipsey and Wilson (2001) and Rosenthal (1991).

Selection of Studies

Empirical studies deemed relevant to this analysis included those testing interventions related to the prevention of eating disorders. At the outset any study testing an intervention aimed to reduce eating disorder risk and/or bolster protective factors for these disorders was considered for inclusion with the following restrictions: studies needed to utilize a nonclinical sample and include some type of comparison group. Clinical samples were excluded so as to not confound the effects of treatment with the effects of a preventive intervention. Uncontrolled study designs were excluded from this meta-analysis to ensure that intervention effects were consistently interpreted in reference to a control group. An additional exclusionary criteria set prior to reviewing relevant studies involved the elimination of studies that did not report outcome data in a sufficient format for effect sizes to be calculated. In the event that data were determined to be insufficient, study authors were contacted to request data needed to compute the effect sizes. During the process of reviewing studies, one modification was made to the inclusion/exclusion criteria. That is, given the limited number of studies including males as participants, analyses were restricted to studies reporting outcome data for female participants (i.e., studies reporting data only on a mixed group of males and females were excluded). Again, study authors were contacted to request the data needed so as to not exclude the study unless all options for inclusion were exhausted.

Locating Relevant Studies

Studies were initially identified via the computerized databases PsycINFO, Web of Sciences, Dissertation Abstracts International, and ERIC. All computer searches were conducted using the following key words and phrases in various combinations: eating disorders, prevention, intervention, eating, attitudes, and behaviors. The reference lists of studies obtained from the computerized search were examined to locate additional relevant studies. In addition, active researchers in the area who had relevant data that were unpublished or in the process of being published were contacted for their results.

Abstracting and Coding Research Results

Each study was coded with respect to a variety of study features and statistical findings. Appendix A provides an example of the coding form used by the raters. Primary categories of study features that were coded included 1) study design, 2) population targeted, 3) length of intervention, 4) length of follow-up, 5) intervention strategies, 6) gender of study participants, and 7) nature of comparison group. Study design was divided into three general categories: a) uncontrolled, b) quasi-experimental controlled, and c) experimental controlled. The first category was used to exclude studies from further analysis, while the latter two categories were used to consider the methodological quality of findings. Gordon's (1983) classification system (i.e., universal, selective, and indicated) was utilized as a framework for classifying the

population targeted in a given study[1]. Studies were coded according to whether interventions were implemented: a) in a single session, b) across multiple sessions, or c) involved continuous or ongoing programming efforts spread across a specified period of time. Studies were also coded according to the total number of sessions (frequency), duration of each session (duration), as well as number of sessions per week (intensity). Intervention strategies were first coded as to whether or not information related to eating disorders was included in the intervention. Studies were also coded as: a) purely psychoeducational, b) enhanced psychoeducational/CBT psychoeducational (i.e., incorporating skills-based techniques drawn from CBT therapy), or c) purely interactive/nonpsychoeducational (i.e., purposeful elimination of psychoeducation). Nature of comparison group was divided into 3 general categories: a) no treatment, b) delayed treatment, and c) alternative treatment. In addition to recording the raw data for each outcome variable employed (e.g., means and standard deviations, *t* value, *p* value), explicit information was coded regarding the name of measure used, whether psychometric properties were reported, treatment group sample size, and comparison group sample size.

Two raters, the author and another clinical psychology graduate student, coded the studies independently. A coding manual (see Appendix B) was created for training purposes, to provide clear operationalizations of each coding category and response

[1] Universal interventions are efforts directed toward the general public or to a whole population group that has not been identified on the basis of increased risk for developing a mental disorder. Selective interventions are those targeted to individuals or specific groups whose risk of developing a clinically diagnosable mental disorder is higher than average. Indicated interventions encompass those targeted to high-risk individuals who, although they do not currently meet DSM-IV criteria for a mental disorder, are identified as having minimal but detectable signs or symptoms of a particular disorder.

option, and to offer guidelines for handling ambiguous cases (Lipsey and Wilson, 2001). Interrater reliability was tested using the kappa statistic. Interrater disagreements were resolved through discussion between the coders.

Calculating Effect Sizes

Effect size statistics of the standardized mean difference variety (i.e., Cohen's *d*) were used in this study. The standardized mean difference is commonly used as the effect size when employing controlled studies in a meta-analysis (Wood & Christensen, in press). The *d* statistic, which is the effect size calculated in its raw form, generally represents the average difference between the treatment group and control group in standard deviation units. Becker's approach to computing *d* (Becker, 1988) was used for all studies employing an independent-groups pretest-posttest design. Becker's technique allows one to capture both pretest to posttest changes and between group differences. With this approach, within group effect sizes are computed as the pretest-posttest change divided by the pretest standard deviation for each treatment condition. The effect size for the between group comparison is calculated by subtracting the within group estimate obtained for the comparison group from that obtained for the experimental group. For studies employing an independent-groups posttest only design, the *d* statistic was calculated with the more traditional and widely used formula, defined as the raw mean difference between the treatment and comparison posttest scores divided by the pooled within group standard deviation of the scores. Morris and DeShon (2002) discussed the suitability of combining effect size estimates calculated with Becker's approach with those calculated from the more traditional method. Effect sizes were calculated so that

positive and negative signs indicated a more and less desirable outcome in the intervention group, respectively. When means and standard deviations were not reported, effect sizes were estimated from the reported *t* and *F* statistics.

DSTAT 1.10 (Johnson, 1993) and *Comprehensive Meta-Analysis* (Borenstein & Rothenstein, 1999) software programs were used to assist in the calculation, storage, and analysis of effect size estimates. Correction for small sample bias was made according to commonly employed formula supplied by Hedges (Lipsey & Wilson, 2001). All subsequent analyses used the corrected unbiased effect size estimates rather than the raw effect size estimates.

Analyzing Data

Effect size estimates were grouped into sets according to the outcome variables they were purported to represent. The most frequently and consistently measured outcome variables were selected to assess intervention efficacy. Moreover, these variables signify many of the commonly studied risk and protective factors in the eating disorder literature and include the following: general eating pathology, dieting, body dissatisfaction, thin-ideal internalization, negative affect, self-esteem, and knowledge. Unique outcome sets were created for each of these seven variables with effect sizes calculated separately on posttest and follow-up data. Independent analyses were conducted at each time point to ensure that the assumption of independence of observations was not violated. Furthermore, within each outcome set, a given study could contribute only one effect size per time point. When multiple outcomes within one study could potentially be included in a particular set of analysis (e.g., multiple measures

of body dissatisfaction were employed), an average value of all relevant effect sizes was calculated to represent the study effect.

Within each outcome set, weighted mean effect sizes were calculated at posttest and follow-up, and the distribution of effect size estimates was visually examined. Each effect size distribution was specifically inspected for the presence of outliers. A technique outlined by Hutcuff and Arthur (1995) was used to identify extreme values, which involves calculating a *sample-adjusted meta-analytic deviancy statistic (SAMD)* for each effect size. Although the *SAMD* statistic was calculated for each effect size relative to the effect size distribution of its specific outcome set, all *SAMD* values were pooled together and plotted to identify outliers. Thus, the identified outliers represent extreme values compared to the effect sizes calculated across the entire pool of studies and all sets of analyses. Analyses with the outliers removed (trimmed results) in addition to analyses with the outliers included (untrimmed results) are both provided below.

The variability of effect sizes was evaluated with the homogeneity statistic *Q*. The *Q* statistic has an approximate chi-square distribution with k-1 degrees of freedom, in which k represents the number of effect size estimates. A significant *Q* rejects the null hypothesis of homogeneity, and supports the search for possible moderating variables. Overall effect size estimates and systematic influence of moderating variables were examined with fixed-effect models. Moderators of a categorical nature were further evaluated by using a technique analogous to ANOVA. With this method, effect sizes were grouped into mutually exclusive categories based on the moderating variable, and the homogeneity among the effect sizes within the categories (Q_W) and the differences

between the categories (Q_B) were tested. Moderators of a continuous nature were further evaluated using a weighted regression analysis. Because statistical software packages estimate the standard error of the regression coefficient by interpreting the weights as representing multiple effect sizes rather than the weightings of single effect sizes, all computer output of standard error and test statistics related to the regression coefficients were recalculated by hand (Lipsey & Wilson, 2001). The overall fit of the regression models were evaluated with regression Q-statistics, Q_R and Q_E, both computed from regression sums of squares estimates (Lipsey & Wilson). A significant Q_R provides evidence that the regression model explains significant variability in the effect sizes, while a significant Q_E is indicative of continued heterogeneity of the effect size distribution that cannot be accounted for by the model.

Interpreting Effect Size Statistics

Particular attention was given to examining the practical meaning of the effect sizes obtained in this study. Conventional rules of thumb were generally used to describe effect sizes (small: $d \leq .20$; medium: $d = .50$; large: $d \geq .80$). However, the Binominal Effect Size Display (*BESD*) was also utilized for the interpretation of effect sizes. The BESD essentially depicts the overlapping distributions of scores from the experimental and comparison groups so that the difference between them can be visualized. More specifically, the *BESD* approach involves setting a "success threshold" at the median of the distribution of scores on the outcome variable for both the treatment and control groups. The proportion of each group that is above the success threshold is

then directly compared to obtain an image of the size of the differential effect in simple success rate terms.

RESULTS

For this study, 54 independent research reports met inclusion criteria and were coded by both raters at the outset. These reports yielded a total pool of 57 separate studies (39 published, 18 unpublished). Eleven of these studies did not contain sufficient information and were excluded from the meta-analysis. The final pool of studies included 32 published and 14 unpublished studies (see Appendix C for a complete list of studies with information on their primary study features). A total of 196 separate effect sizes were calculated and grouped into outcome sets according to the following variables: knowledge, general eating pathology, dieting, internalization, body dissatisfaction, negative affect, and self-esteem. Table 1 provides a summary of the average weighted effect size estimates for each outcome set. Effect size estimates from individual studies can be found in Appendix D.

The general analytic approach included the initial examination of the weighted mean effect size and variability of the effect size distribution within each outcome set at posttest and follow-up. Moderator analyses involving population targeted, type of intervention strategy, and duration of intervention were conducted on a given outcome set only in the event that significant heterogeneity of variance was detected. Each moderator was evaluated independently. However, two unique aspects of intervention duration were entered into the regression model and examined simultaneously: total number of intervention sessions and number of sessions per week. Finally, specific research questions posed by the investigator regarding the insufficiency of single session interventions and the potential iatrogenic effects of including eating disorder information

were explored across all relevant outcome sets, regardless of the results of initial heterogeneity tests.

Table 1

Summary Data for Weighted Mean Effect Size Estimates

Outcome Variable	Posttest				Follow-up			
	N	Average weighted *d*	*Q*	*P*	*N*	Average weighted *d*	*Q*	*p*
Knowledge	8[a]	1.27[a]	12.77[a]	.08[a]	7	0.75	39.86	<.001
General Eating Pathology	20	0.17	13.61	.81	14	0.13	13.39	.42
Dieting	19[b]	0.20[b]	20.45[b]	.31[b]	13	0.18	5.65	.93
Internalization	14	0.21	34.05	<.01	10	0.18	25.33	<.01
Body Dissatisfaction	34[c]	0.13[c]	43.34[c]	.11[c]	18[d]	0.07[d]	21.00[d]	.23[d]
Negative Affect	7	0.05	10.96	.09	5	0.21	0.63	.18
Self-Esteem	14	0.17	37.33	<.001	7	0.05	24.45	<.001

Note. Data with superscripts denote results based on trimmed effect size distributions.

[a]With inclusion of the two outliers involving posttest knowledge, $n = 10$, $d = 1.21$, $Q = 84.36$, $p < .001$.

[b]With inclusion of the outlier involving posttest dieting, $n = 20$, $d = 0.24$, $Q = 61.92$, $p < .001$. [c]With inclusion of the two outliers involving posttest body dissatisfaction, $n = 36$, $d = 0.17$, $Q = 101.99$, $p < .001$.

[d]With inclusion of the outlier involving body dissatisfaction at follow-up, $n = 19$, $d = 0.08$, $Q = 42.10.$, $p < .01$

Reliability Analyses

Prior to analyzing the effect size data, intterrater agreement for the primary coding categories was evaluated. Kappa statistics for these categories ranged from .51 to .96 (see Table 2), which are considered to be within an acceptable range (Landis & Koch, 1977). As discussed above, all interrater disagreements were resolved through

discussion between the coders and final ratings were assigned based on 100% agreement.

Table 2

Interrater Agreement for Primary Coding Categories

Category	Kappa
study design	.96
gender of study participants	.95
population targeted	.81
type of intervention strategy	.51
inclusion of eating disorder information	.75
nature of comparison group	.59
duration of intervention	.92

Knowledge

There were ten studies included in this outcome set with effect size estimates at posttest ranging from $d = .31$ to $d = 3.81$. Two of the outliers were located within the data set for posttest knowledge and with their removal the trimmed distribution of effect size estimates ranged from $d = 1.00$ to $d = 2.00$. As found in Table 1, the average weighted mean effect sizes at posttest were similar for the untrimmed and trimmed data sets, 1.21 (95% CI = 1.13 to 1.29) and 1.27 (95% CI = 1.19 to 1.36), respectively. However, the variability of the effect size distribution was significantly affected by the inclusion of the two outliers. Significant heterogeneity of variance was found for the untrimmed data set at posttest, $Q(9) = 84.36$, $p < .001$. Removing the outliers reduced

the variability of the effect size distribution markedly, to the extent that the distribution became homogeneous, Q(7) = 12.77, $p = .08$.

Moderator analyses conducted on the untrimmed data set at posttest revealed the following: 1) a marginally significant effect for type of intervention strategy, $Q_B(1) = 3.47$, $p = .06$, with continued heterogeneity of variance within each subgroup, 2) a significant effect for population targeted, $Q_B(2) = 18.28$, $p < .001$, with continued heterogeneity of variance within each subgroup, 3) and a significant effect for number of sessions per week ($B = .12$, *se* $B = .03$ $p < .001$), with continued heterogeneity of the effect size distribution, ($Q_R(2) = 17.65$, $p < .001$, $Q_E(7) = 451.01$, $p < .001$). Removing the outliers eliminated the between group effects as well as the significance of the regression model.

At follow-up, effect sizes for knowledge ranged from $d = .18$ to $d = 1.40$ across seven studies. The average weighted mean effect size of $d = .75$ (95% CI = .65 to .84) did not adequately summarize the data in this set as significant heterogeneity of variance was found, $Q(6) = 39.86$, $p < .001$ (see also Table 1). However, results of the moderator analyses were not straightforward. Although a marginally significant effect was found for intervention strategy, $Q_B(1) = 3.30$, $p = .07$, and a significant effect was found for population targeted, $Q_B(2) = 6.82$, $p = .03$, there was still significant heterogeneity of variance within all of the subgroups across both analyses. Thus, some differences were found based on the categorical moderators tested. However, a significant amount of variability among the effect sizes remained unaccounted for and rendered the within

group effect size estimates at follow-up less meaningful. No significant effects were found for moderators related to duration of intervention, Q_R (2)= 1.26, $p > .10$.

General Eating Pathology

At posttest, effect sizes for general eating pathology ranged from $d = -.11$ to $d = .62$ across 20 studies. The average weighted mean effect size at posttest of $d = .17$ (95% CI = .09 to .24) appeared to adequately summarize the data as the variance of the effect size distribution was found to be homogeneous, $Q(19) = 13.61$, $p = .81$. As such, moderator analyses were not conducted.

Effect size estimates for general eating pathology at follow-up ranged from $d = .21$ to $d = .82$ across 14 studies. As seen in Table 1, the average weighted mean effect size of $d = .13$ (95% CI = .04 to .22) was similar to the effect found at posttest and also represented a homogeneous distribution of effect size estimates, $Q(13) = 13.39$, $p = .42$.

Dieting

Effect size estimates for dieting at posttest ranged from $d = -.16$ to $d = 1.92$ across 20 studies. One of the outliers was located within the data set for posttest dieting, and with its removal the average weighted mean effect size diminished from $d = .24$ (95% CI = .16 to .32) to $d = .20$ (95% CI = .16 to .28). The variability of the effect size distribution was significantly affected by the inclusion of the outlier. In the trimmed data, effect size estimates were found to be homogeneous, $Q(18) = 20.45$, $p = .31$. Significant heterogeneity of variance was found for the untrimmed data set, $Q(19) = 61.92$, $p < .001$.

Moderator analyses conducted on the untrimmed data set at posttest revealed the following: 1) no significant effect for intervention strategy, $Q_B(2) = .35$ $p = .84$, 2) a significant effect for population targeted, $Q_B(2) = 11.88$, $p = < .001$, but continued heterogeneity of variance in the group that contained the outlier, and 3) significant effects for both total number of intervention session ($B = .01$, *se* $B < .01$, $p < .001$) and number of sessions per week ($B = -.14$, *se* $B = .02$, $p < .001$), with continued heterogeneity of the effect size distribution ($Q_R(2) = 66.93$, $p < .001$, $Q_E(13) = 993.84$, $p < .001$). Removing the outlier eliminated the effects involving intervention duration. However, a significant effect remained for population targeted in the trimmed data set. This effect was explained by the significantly higher mean effect size for the selective subgroup ($d = .28$) compared to both the indicated ($d = .07$) and universal subgroups ($d = -.01$).

At follow-up, effect size estimates for dieting ranged from $d = -.36$ to $d = .46$ across 13 studies. As seen in Table 1, the average weighted mean effect size of $d = .18$ (95% CI = .10 to .27), appeared to adequately summarize the data for dieting at follow-up as the variance of the effect size distribution was found to be homogeneous, Q(12) = 5.65, p = .93.

Internalization

There were 14 studies included in this outcome set, with effect size estimates at posttest ranging from $d = -.28$ to $d = 1.40$. The average weighted mean effect size at posttest of $d = .21$ (95% CI = .11 to .31) did not adequately summarize the data contained within this set as significant heterogeneity of variance was found, $Q(13) =$

34.05, $p < .01$ (see also Table 1). The search for moderators on posttest data revealed: a) no significant effect for intervention strategy, $Q_B(1) = 1.02$, $p = .31$, and b) a significant effect for population targeted $Q_B(2) = 6.79$, $p = .03$ with homogeneity of variance for the indicated and universal subgroups but continued heterogeneity of variance within the selective subgroup [$Q_W(4) = 16.23$, $p < .01$]. The effect for population targeted was explained by the significantly higher mean effect size for the indicated subgroup ($d = .48$) compared to both the selective ($d = .13$) and universal subgroups ($d = .18$). Finally, there was a significant effect of number of sessions per week ($B = .01$, *se* $B < .01$, $p < .001$); however, continued heterogeneity of the effect size distribution remained after controlling for this effect ($Q_R(2) = 9.88$, $p < .001$; $Q_E(11) = 545.06$, $p < .001$.)

At follow-up, effect sizes for internalization ranged from $d = -.38$ to $d = 1.18$ across 10 studies. The average weighted mean effect size of $d = .18$ (95% CI = .07 to .30) did not adequately summarize the data contained in this set as significant heterogeneity of variance was found, $Q(9) = 25.33$, $p < .01$ (see also Table 1). There was not enough variability to test for intervention strategy as a moderator as only one study employed a purely psychoeducational strategy and only one study employed a purely interactive strategy. There was not a significant effect for population targeted, $Q_B(1) = 1.23$, $p = .27$. However, the trend toward higher effect sizes for studies employing an indicated population compared to studies employing either a universal or selective population remained ($d = .30$ for indicated subgroup, and $d = .15$ for non-indicated subgroup). Finally, there were significant effects for both total number of intervention session ($B = -.02$, *se* $B < .01$, $p < .001$) and number of sessions per week (B

= -.08, *se B* = .01, $p < .001$), with continued heterogeneity of the effect size distribution ($Q_R(2) = 47.08$, $p < .001$; $Q_E(7) = 402.00$, $p < .001$).

Body Dissatisfaction

Body dissatisfaction was the most frequently employed outcome variable across the studies in this meta-analysis. Thirty-six studies contributed to the effect size distribution of posttest body dissatisfaction with estimates ranging from $d = -.30$ to $d = 2.57$. Two of the outliers were located within this data set, and with their removal the average weighted mean effect size diminished from $d = .17$ (95% CI = .10 to .23) to .13 (95% CI = .06 to .20). The variability of the effect size distribution was significantly affected by the inclusion of the outliers. In the trimmed data set, effect size estimates were found to be homogeneous, $Q(33) = 43.34$, $p = .11$. Significant heterogeneity of variance was found for the untrimmed data set, $Q(35) = 101.99$, $p < .001$.

Moderator analyses conducted on the untrimmed data set at posttest revealed no significant effect for intervention strategy, $Q_B(2) = .94$, $p = .62$. There was a significant effect for population targeted, $Q_B(2) = 6.64$, $p = .04$ with homogeneity of variance for the universal subgroup but continued heterogeneity of variance for the selective and indicated subgroups (the groups containing the outliers). This effect was explained by the significantly higher mean effect size for the indicated subgroup ($d = .36$) compared to both the selective ($d = .16$) and universal ($d = .08$) subgroups. There were no significant effects of intervention duration in the untrimmed data set, $Q_R(2) = 2.89$, $p > .10$. Removing the outliers served to eliminate the between group effect for population targeted, $Q_B(2) = 4.55$, $p = .10$; however, the trend toward a higher mean effect size for

the indicated subgroup (d = .30) compared to the selective (d = .11) and universal (d = .08) subgroups remained.

At follow-up, effect sizes for body dissatisfaction ranged from d = -.41 to d = 2.75 across 19 studies. One of the outliers was located within this set and with its removal the trimmed distribution of effect size estimates at follow-up ranged from d = -.41 to d = .73. The average weighted mean effect sizes were similar for the untrimmed and trimmed data sets, .08 (95% CI = -.01 to .17) and .07 (95% CI = -.02 to .15) respectively. The variability of the effect size distribution was significantly affected by the inclusion of the outlier. In the trimmed data set, effect size estimates were found to be homogeneous, $Q(17) = 21.00$, $p = .23$. Significant heterogeneity of variance was found for the untrimmed data set, $Q(18) = 42.10$, $p = .001$.

Moderator analyses conducted on the untrimmed data set at follow-up revealed no significant effect for intervention strategy, $Q_B(2) = 0.55$, $p = .76$, and a significant effect for population targeted, $Q_B(2) = 7.83$, $p = .02$ with continued heterogeneity of variance in the subgroup containing the outlier. The effect for population targeted was explained by the significantly higher mean effect size for the indicated subgroup (d = .30) compared to the selective subgroup (d = -.01) with no differences between these two groups and the universal subgroup (d = .16). In the untrimmed data set, there was also a significant effect for total number of sessions (B = -.01, $se\ B < .01$, $p < .001$) with continued heterogeneity of the effect size distribution, ($Q_R(2) = 18.11$, $p < .001$; $Q_E(14) = 476.65$, $p < .001$). Removing the outlier eliminated the effect involving intervention duration; however, the trend toward a higher mean effect size for interventions targeting

an indicated population (d = .22) compared to non-indicated populations (d = .03) remained.

Negative Affect

There were seven studies included in this outcome set, with effect size estimates at posttest ranging from d = -.30 to d = .86. The average weighted mean effect size of d = .05 (95% CI = -.14 to .24) appeared to adequately summarize the posttest data as the variance of the effect size distribution was found to be homogeneous, $Q(6) = 10.96$, p = .09 (see also Table 1).

At follow-up, effect sizes for negative affect ranged from d = -.01 to d = 69. The average weighted mean effect size of d = .20 (95% CI = -.001 to .41) also represented a homogeneous distribution of effect size estimates, $Q(4) = 6.28$, p = .1793.

Self-Esteem

Effect size estimates for self-esteem at posttest ranged from d = -.17 to d = 1.68 across 14 studies. The average weighted mean effect size of d = .17 (95% CI = .08 to .26) did not adequately summarize the posttest data as significant heterogeneity of variance was found, $Q(13) = 37.33$, $p < .001$. Moderator analyses revealed no significant effects for intervention strategy, $Q(1) < .01$, p = .97 or population targeted, $Q(2) = .16$, p = .92. There were no also significant effects involving intervention duration, ($Q_R(2) = 5.74$, $p > .05$, $Q_E(8) = 79.03$, $p > .10$)

At follow-up, effect sizes for self-esteem ranged from d = -.73 to d = 1.47. Consistent with the posttest data, the average weighted mean effect size estimate (d = .05, 95% CI = - .05 to .15) did not appear to adequately summarize the data contained

within this set because significant heterogeneity of variance found, $Q(6) = 24.45$, $p < .001$ (see also Table 1). Moderator analyses revealed no significant effects for intervention strategy, $Q_B(1) = 1.58$ $p = .21$ or population targeted, $Q_B(2) = 2.25$, $p = .32$. Results from the weighted regression analysis were not as straightforward. Significant effects were found for total number of sessions ($B = -.04$, *se* $B = .01$, $p < .001$) and number of sessions per week ($B = -.16$, *se* $B = .05$, $p < .01$; however, significant heterogeneity of variance remained in the effect size distribution after accounting for these effects ($Q_R(2) = 76.48$, $p < .001$; $Q_E(3) = 108.37$, $p < .001$).

Intervention Duration

In this set of analyses, data involving the duration of intervention were treated as categorical to explore the fundamental question of whether interventions implemented in multiple sessions are intrinsically more effective than single session interventions. Attempts were made to examine this specific research question across all outcome sets. However, the moderator was unable to be tested in the following data sets due to insufficient variability (i.e., only one study included in one of the subgroups): knowledge at follow-up, general eating pathology at follow-up, dieting at follow-up, internalization at follow-up, negative affect at posttest and follow-up, and self-esteem at follow-up. Table 3 provides a summary of the between-group effect size estimates, *Q* statistics, and *p* values associated with each outcome set with sufficient variability to analyze the effects of intervention duration. The only significant between group effect found was for knowledge at posttest, $Q_B(1) = 8.49$, $p < .01$. This finding can be better understood when one considers that the analyses were run on the untrimmed data set, which includes

two outliers. Analyses on the trimmed data set could not be conducted as this would reduce the number of studies in the single session subgroup to one. No other significant between group differences were found within the outcome sets.

Table 3

Summary Data for Analysis of Intervention Duration (Single Session vs. Multiple Session)

Outcome Variable	Posttest				Follow-up			
	Single Session *d*	Multiple Session *d*	*Q*	*p*	Single Session *d*	Multiple Session *d*	*Q*	*p*
Knowledge	1.62	1.17	8.49	<.01				
General Eating Pathology	0.14	0.17	0.02	.88				
Dieting	0.19[a]	0.20[a]	0.02[a]	.88[a]				
Internalization	0.13	0.24	1.02	.31				
Body Dissatisfaction	0.10[b]	0.14[b]	0.28[b]	.59[b]	-0.08[c]	0.09[c]	1.74[c]	.19[c]
Self-Esteem	0.19	0.17	0.01	.92				

Note. Cells left blank indicate that there was insufficient variability to analyze the effects of intervention duration. Data with superscripts denote results based on trimmed effect size distributions.
[a]With inclusion of the outlier involving posttest dieting, single session $d = 0.34$, multiple session $d = 0.20$, $Q = 2.28$, $p = .13$. [b]With inclusion of the two outliers involving posttest body dissatisfaction, single session $d = 0.20$, multiple session $d = 0.15$, $Q = 0.33$, $p = .57$. [c]With inclusion of the outlier involving body dissatisfaction at follow-up, single session $d = -0.80$, multiple session $d = 0.11$, $Q = 2.11$, $p = .15$.

Effects of Including Eating Disorder Information

The outcome sets deemed relevant to analyzing the potential iatrogenic effects of including eating disorder information in an intervention involved the constructs most directly related to eating disorder symptomatology, namely: general eating pathology, dieting, internalization, body dissatisfaction, and self-esteem. Table 4 provides a

summary of the between group effect size estimates, *Q* statistics, and *p* values associated with each outcome set. Within the majority of outcome sets, there were no significant between group effects. However, when significant differences were found, they were explained by higher mean effect sizes for interventions that included eating disorder information compared to interventions that did not incorporate such psychoeducational material. No harmful effects were found regardless of whether eating disorder information was included or not.

Table 4

Summary Data for the Analysis of the Potential Iatrogenic Effects of Including Eating Disorder Information

Outcome Variable	Post-test				Follow-up			
	Inclusion (*d*)	Exclusion (*d*)	*Q*	*p*	Inclusion (*d*)	Exclusion (*d*)	*Q*	*P*
General Eating Pathology	0.19	0.12	0.60	.44	0.14	0.13	0.01	.92
Dieting	0.24[a]	0.13[a]	1.49[a]	.22	0.18	0.22	0.16	.69
Internalization	0.56	0.15	8.53	<.01	0.21	0.01	1.57	.21
Body Dissatisfaction	0.17[b]	0.11[b]	0.79[b]	.37	0.06[c]	0.07[c]	0.01[c]	.92
Self-Esteem	0.16	0.19	0.10	.76	0.11	-0.04	2.02	.16

Note. Data with superscripts denote results based on trimmed effect size distributions.

[a]With inclusion of the outlier involving posttest dieting, inclusion $d = 0.30$, exclusion $d = 0.13$, $Q = 3.87$, $p < .05$. [b]With inclusion of the two outliers involving posttest body dissatisfaction, inclusion $d = 0.28$, exclusion $d = 0.11$, $Q = 5.81$, $p < .05$. [c]With inclusion of the outlier involving body dissatisfaction at follow-up, inclusion $d = 0.10$, exclusion $d = 0.07$, $Q = 0.10$, $p = .75$.

File Drawer Analysis

To assess for the potential effects that unpublished or unretrieved studies could have on the findings in this meta-analysis, fail-safe N analyses were computed (Rosenthal, 1979). The fail-safe N estimates the number of unretrieved studies reporting null results that would be needed to reduce the cumulated effects across studies to the point of statistical non-significance. These estimates were calculated for the outcome variables with average weighted effect sizes significantly larger than zero (i.e., 95% confidence interval did not include zero). As seen in Table 5, fail-safe N values ranged from 23 to 232, and generally indicated that more than double the amount of studies with null findings would be needed within any set of analyses to reduce the effects found to the point of nonsignificance.

Table 5

Results of the File Drawer Analysis

Outcome Variable	Post-test		Follow-up	
	N	*Fail Safe N*	*N*	*Fail Safe N*
Knowledge	10	232	7	97
General Eating Pathology	20	46	14	23
Dieting	20	76	13	35
Internalization	14	45	10	26
Body Dissatisfaction	36	84		
Self-Esteem	14	34		

Note. *N* refers to the number of studies in the original distribution of effect sizes. The criterion effect size level used to calculate Fail Safe N was .05. Fail safe Ns were not calculated for Body Dissatisfaction or Self-Esteem at follow-up because the average weighted effect sizes were found to be statistically similar to .05.

DISCUSSION

In this study, a quantitative evaluation of the overall effectiveness of eating disorder prevention programs and an investigation of potential moderating variables that may influence the magnitude of intervention effects was accomplished by using meta-analysis. Outcome was assessed with a variety of indicators thereby allowing for detailed information about the manner in which these programs influenced particular outcome variables to be discussed. Questions that have been raised in the literature regarding the potential benefits of targeted approaches compared to untargeted intervention approaches, potential iatrogentic effects of including eating disorder information, and insufficiency of "one shot" intervention programs were also addressed. The findings presented challenge the negative conclusions drawn in previous review articles regarding the inability of such programs to demonstrate behavioral improvements. Moreover, these data have important implications for the guidance and dissemination of preventive interventions on a large-scale and community-wide basis.

Magnitude of Effects on Specific Outcome Variables

The seven outcome variables used in this meta-analysis were not only among the most frequently and consistently measured variables across the pool of studies but they also represented many of the widely recognized risk and protective factors in the eating disorder literature. Effect sizes were calculated separately at posttest and at follow-up for each of the following variables: knowledge, general eating pathology, dieting, body dissatisfaction, thin-ideal internalization, negative affect, and self-esteem. Significant change from posttest to follow-up within an outcome set was established when the

average weighted effect size at each time point fell outside of the range of confidence intervals for one another. It is important to note that there were no overall negative effects in any outcome set at posttest or follow-up. Across the outcome sets, the average weighted effect size estimates were classified in the following manner: no effect, small positive effect, medium positive effect, or large positive effect. A classification of no effect was given to an outcome set when the confidence interval surrounding the average weighted effect size estimate ranged from a negative value to a positive value (i.e., the confidence interval included zero). Classifications of small, medium, and large effects were based on conventional rules of thumb, small: $d \leq .20$; medium: $d = .50$; large: $d \geq .80$. (Cohen, 1992).

The outcome variable that was most affected by the prevention programs was knowledge. At both posttest and follow-up, eating disorder prevention programs produced large positive effects on the acquisition of knowledge. The average weighted effect size for knowledge at posttest was approximately 1.2 (1.21 and 1.27 for untrimmed and trimmed data sets respectively). All of the individual effect sizes within this distribution were above 1.0 with one exception, and that remaining effect size ($d = .31$, Neumark-Sztainer et al., 2000) was identified as an outlier from the total pool of studies. At follow-up, the average weighted effect size for knowledge was .75, which represents a reduction in the magnitude of the effects found for knowledge at posttest. However, there was significant heterogeneity of variance found in this distribution of effect sizes. Four out of seven individual effect sizes in the follow-up data set were greater than 1.0, with remaining effect size estimates in the range of .6, .4, and .2.

Although the effects for knowledge at follow-up were less consistent compared to the effects found at posttest, the effect size values were still generally larger compared to those found for other indicators. It is thus concluded that the eating disorder prevention programs evaluated in this study were successful in promoting knowledge related to eating disorders.

Remaining outcome variables that demonstrated positive effects following eating disorder prevention programs included general eating pathology, dieting, internalization, body dissatisfaction, and self-esteem. Small net effects were found for each of these variables at posttest and/or follow-up. The effects for general eating pathology at posttest and follow-up were the most straightforward to interpret given the homogeneity of the effect size distributions. As such, the average weighted effect sizes of .17 at posttest and .13 at follow-up provided an adequate summary of the data contained within each set. These effect size estimates suggested that eating disorder prevention programs consistently had small positive effects on the general eating pathology of its participants. Given that general eating pathology was assessed with broad screening instruments such as the EAT and EDE-Q, a wide range of eating disordered behaviors and attitudes were captured by these instruments. Thus, there was consistent demonstration of improvements in eating disorder related behaviors and attitudes across a variety of intervention strategies and population groups. Moreover, these effects were maintained at follow-up, which for individual studies ranged from 4 to 52 weeks.

The small effects found for dieting also showed evidence for homogeneity of variance at posttest and follow-up. At posttest, the effect size distribution was found to

be homogeneous after removing an outlier, resulting in an average weighted effect size of .20. The average weighted effect size of .18 at follow-up was of a similar magnitude to the effect found at posttest and also represented a homogeneous distribution of effect size estimates. Although the search for moderators revealed some significant findings for dieting at posttest (as discussed below), the results generally indicate that small improvements in dieting behavior were found across the studies at both posttest and follow-up. These findings offer evidence of improvements on a specific behavioral outcome variable following an eating disorder prevention program thereby directly challenging conclusions drawn in previously published qualitative reviews of this literature (Austin, 2000; Fairburn, 1995; Franko & Orosan-Weine, 1998; Musell et al., 2000).

Regarding the effects found for body dissatisfaction, internalization, and self-esteem, the pattern of results was relatively more complicated to explain. Small overall effects were demonstrated for body dissatisfaction at posttest, but by follow-up the overall effect size was essentially reduced to zero. Moderator analyses revealed a different pattern of results, showing small positive effects on the body dissatisfaction of a particular subgroup of participants in the long-term. The effects on internalization were better explained by moderating variables at both posttest and follow-up. With respect to self-esteem, the overall effect size found at posttest was not considered to be representative of the data set, and no moderators were able to explain the heterogeneity of variance of the effect size distribution. As such, no firm conclusions could be drawn

regarding the manner in which eating disorder prevention programs affected the self-esteem of participants at posttest.

In general, no significant effects were found for negative affect at posttest and follow-up, and self-esteem at follow-up as the confidence interval for the overall effect sizes included zero. With these findings, it is important to consider the number of individual effect size estimates within each distribution and the total number of participants in the data set. The analyses for the effects of negative affect at posttest and follow-up and for self-esteem at follow-up were conducted on seven studies or less. With respect to negative affect at posttest, seven studies were included in the data set, and there were a total of 474 participants. At follow-up, the data set for negative affect included only five studies with a total of 368 participants. Although both of these effect size distributions were found to be homogeneous, the reduced number of studies and relatively low sample size per study limit the conclusions which can be drawn about the lack of a significant effect for this outcome variable. Examination of the individual effect size estimates for negative affect at posttest revealed no consistent pattern as there was one estimate each valued approximately at .9, .5, .3, .2, 0, -.1, -.3. There was greater consistency for the effect size estimates at follow-up with 3 values of approximately .1, and the other 2 values of -.1 and .7.

A potentially large contributor to the null findings for negative affect was the small samples sizes for each study as this tends to increase the width of the confidence interval for the effect size estimate. Much greater confidence was placed in interpreting the findings for knowledge at follow-up even though this data set also included only

seven studies, chiefly because the results were based on a total sample of 1736 participants. Regarding the null effects found for self-esteem at follow-up, this data set included seven studies with a total of 1488 participants. Thus, relatively greater confidence can be placed in interpreting the results from these data due to the sample size on which they are based. However, significant heterogeneity of variance was found for the effect size distribution rendering the overall effect size estimate meaningless. As with the effects at posttest, no moderators were able to explain the heterogeneity. Thus, it was generally concluded that eating disorder prevention programs produced inconsistent effects on self-esteem at posttest and at follow-up.

Potential Benefits of Targeted Approaches Compared to Untargeted Approaches

As described above, a variety of moderating variables were tested to explore the variability of the effect size distributions for each outcome set. Population targeted was found to explain variability within several outcome sets containing significant heterogeneity of variance, and typically served as the sole moderator that could account for the variability. There was a tendency toward greater benefits in studies targeting participants considered to be at a relatively higher risk for developing an eating disorder (i.e., indicated or selective population groups) compared to those employing participants not identified as at-risk (i.e., universal group). The highest risk groups were not always associated with the best outcome, which points to the need to review the results for each outcome set independently.

Significant moderator effects for population targeted were found when exploring the effects of dieting, internalization, and body dissatisfaction. The effects for

internalization and body dissatisfaction showed higher mean effect sizes for indicated subgroups compared to selective and universal subgroups. These findings indicate that studies targeting participants with minimal signs or symptoms of eating disorders tended to be more effective at reducing thin-ideal internalization and body dissatisfaction compared to studies employing participants who were considered to be at a higher than average risk for developing an eating disorder or those who were not previously identified as being at risk. For internalization, these between group differences were statistically significant at posttest (indicated d = .48, selective d = .13, universal d = .18) with trends toward a higher mean effect size for indicated (d = .30) compared to non-indicated (d = .15) groups remaining at follow-up. Moderator effects involving body dissatisfaction were explored on trimmed and untrimmed data sets and revealed the following: a) significant between group differences in favor of the indicated group at both posttest (indicated d = .36, selective d = .16, universal d = .08) and follow-up (indicated d = .30, selective d = -.01, universal d = .16) in the untrimmed data set, and b) trends toward higher mean effect sizes in the indicated compared to non-indicated groups at both time points in the trimmed data set.

A significant moderator effect for population targeted was also found for dieting at posttest; however, results indicated a higher mean effect size for the selective group (d = .30) compared to both the indicated (d = .07) and universal groups (d = -.01). This between groups effect for population targeted was no longer significant at follow-up. Thus, the effects for dieting were better explained by population targeted in the short-term, but by follow-up the small effects found were no longer related to this moderator.

Although these results are somewhat difficult to explain, they provide some indication of the benefits of targeting a high-risk group that is not yet demonstrating minimal signs and symptoms of an eating disorder. Of particular interest is the demonstration of significantly higher effects for dieting in a high-risk group without known eating disorder symptoms compared to a high-risk group with identifiable signs and symptoms.

These findings suggest that the interventions studied here did not influence intervention participants in the exact same manner on outcome variables based on their degree of risk for developing an eating disorder. Results however generally support the notion that greater benefits were found with studies employing higher risk populations. As such, these findings suggest that it may be more worthwhile to focus on higher risk populations when conducting interventions as remarked by Killen et al., (1993). An alternative explanation for these findings is that interventions are currently designed to optimally detect effects for higher risk groups. At the present time, significant change on attitudinal and behavioral outcome variables are employed as the definitive outcome indicators in eating disorder prevention programs regardless of population targeted. Perhaps goals and expectations regarding attitudinal and behavioral change may not be reasonable or realistic for universal populations of children and adolescents who are not initially displaying problematic eating attitudes or behaviors. Researchers conducting prevention studies with universal populations have noted the difficulties with demonstrating statistically significant changes on eating disorder variables due to low frequencies of endorsement rates on the variables at baseline (Huon, Roncolato, Ritchie & Braganza, 1997; Phelps, Sapia, Nathanson, & Nelson, 2000; Smolak, Levine, &

Schermer, 1998a). In this vein, one then becomes concerned about the issue of range restriction when employing universal populations within such a framework. Alternative indicators of outcome may therefore need to be considered if one wishes to target individuals who are not identified on the basis of being at risk for developing an eating disorder. It is important to note; however, that population targeted did not emerge as a significant variable related to effects involving knowledge or more general aspects of eating pathology.

Effects of Including Eating Disorder Information and Other Variations in Intervention Strategy

In this study, there was no evidence to support concerns raised regarding potential iatrogenic effects of incorporating psychoeducational material on eating disorders. No harmful effects were found regardless of whether eating disorder information was incorporated into an intervention or not. Within the majority of outcome sets, there were no significant differences in effect sizes between groups of studies including or excluding such material. When significant differences were found, they were actually explained by higher mean effect sizes for interventions including eating disorder information compared to interventions excluding such material. Thus, findings from this study tend to support the inclusion of such material in eating disorder prevention programs.

With respect to other variations in intervention strategy tested as moderators, no differences were found based on intervention content. Specifically, purely psychoeducational approaches compared to interventions incorporating skills based

techniques drawn from cognitive-behavioral therapy (i.e., enhanced psychoeducational) appeared to be equally effective in influencing knowledge, general eating pathology, dieting, body dissatisfaction, and self-esteem. There were no harmful effects based on intervention content. In addition to a lack of significant findings for intervention strategy, no significant effects emerged for intensity or duration of intervention. As such, assumptions about the insufficiency of single session programs were not supported by the data. In this study, "one-shot" prevention programs were found to be equally effective across all outcome variables compared to interventions implemented over multiple sessions. These results suggest that interventions of varying lengths and contents can have small but meaningful effects on attitudinal and behavioral risk factors for eating disorders and large effects on eating disorder-related knowledge. It is important to note that studies testing the effects of one session interventions typically did not include a follow-up period. Thus, no conclusions regarding the effectiveness of "one-shot" prevention programs beyond posttest could be drawn in this study.

Practical and Clinical Significance of Findings

To aid in the interpretation of the practical significance of these findings, a Binomial Effect Size Display (*BESD*; Rosenthal & Rubin, 1982) was calculated for each outcome set with a statistically significant effect size that did not have significant heterogeneity of variance within the effect size distribution. Success rates between the comparison and treatment groups were contrasted to quantify the differential rate of improvement in the treatment group that was found. For knowledge, results revealed success rates for the comparison group in the range of 23-32% with a corresponding

increase to 68-77% for the treatment group. A *BESD* was calculated for the effects of internalization on the indicated subgroup given that the overall distribution of effect sizes was found to be heterogeneous. A differential rate of improvement of 24% was found between the treatment and control group, which indicates an optimal success rate on this outcome variable. The small effects found for general eating pathology, dieting, and body dissatisfaction were associated with a relatively reduced overall rate of improvement compared to those reported for knowledge and internalization. However, differential rates of improvement between 6-10% were found these variables at posttest and/or follow-up.

With respect to the small effects found for general eating pathology, dieting, and body dissatisfaction in this study, a number of methodologists have discussed the practical meaning of effect size estimates and argued that quantitatively small effects can be quite important (Abelson 1985; Prentice & Miller, 1992). Rosenthal and DiMatteo (2001) reviewed research where effect sizes even when they are so small as to have r^2 = .0012 were associated with a treatment method that could prevent hear attacks in 34 out of every 1000 patients similar to those represented in the study. Considering the adverse medical consequences, poor treatment outcome, and mortality rates associated with eating disorders, a reduction of risk for developing these disturbances is considered clinically valuable. In this light, the differential rates of improvement between 6-10% can be considered a meaningful contribution to the field of eating disorder prevention research.

It is also important to consider the broader context in which disordered eating attitudes and behaviors are developed and maintained. A variety of biological, psychological, and social factors influence the eating attitudes and behaviors of individuals. Perhaps it is unrealistic to expect medium or large effects on such attitudes and behaviors that are shaped and continuously affected by familial, societal, and peer factors, particularly when the interventions being tested are of a fixed and relatively brief duration. Nearly 24% of the interventions included in this analysis were implemented in one session, with small effects holding across these minimal manipulations. These findings point to the pervasiveness of the effect even with the most minimal manipulation, which can perhaps be even more impressive than demonstrating that a given effect accounts for a great deal of variance (Prentice & Miller, 1992).

Precautions Taken During Analyses

Given the intricacy of the data analytic process undertaken in this study, it is important to point out the numerous precautions that were taken to enhance the accuracy of the findings. Particular attention was paid to the potential upward bias of the mean effect size due to sampling bias or systematic omission of unpublished research. Considerable attempts were made to contact active researchers in the area for unpublished research and to include database searches that report unpublished master's theses or doctoral dissertations. In addition, a file drawer analysis was conducted to assess for the potential effects that unpublished or unretrieved studies could have on the statistically significant findings. The range of fail-safe N estimates was from 23 to 232 and suggested that more than double the amount of studies with null findings would be

needed within outcome sets with significant results to reduce the cumulated effects across the studies to the point of statistical non-significance. The comprehensive nature of the search strategy argues against there being this many unretrieved studies. As such, greater confidence can be placed that a reliable estimate of mean effect sizes were obtained in this study and that the observed results are not likely to be substantially biased due to sampling.

Abstraction and coding of results were carefully conducted, and repeated contact was made with researchers to clarify findings in published documents or to ask for additional data. Two coders independently rated each study in the meta-analysis with respect to study features and statistical findings in order for reliability analyses to be conducted. Although in-depth analyses of methodological features were not conducted, two aspects of methodological quality were attended to during the coding and analysis phases: study design and nature of comparison group. Neither of these variables appeared to be systematically related to effect size variability within outcome sets. Finally, it is important to note that conclusions regarding variability of effect size distributions were not simply made from the value of the *Q* statistic. Visual inspections of individual effect sizes within distributions were conducted for each outcome set to supplement the statistical findings.

Implications for the Guidance of Future Intervention Strategies

Findings from this study support the continued development and implementation of eating disorder prevention programs, as the interventions studied here were found to be effective in influencing eating disorder related knowledge, attitudes, and behaviors.

Results further suggest that the provision of information about eating disorders should be considered for inclusion, and that skills-based and/or psychoeducation-based intervention strategies can be equally effective. One of the most important factors to consider when designing future studies appears to be population targeted. As discussed above, numerous explanations can be provided to account for the results demonstrating a tendency toward greater benefits for high-risk participants. It may be beneficial to customize goals/expectations and interventions to the population targeted for intervention as significant change on attitudinal and behavioral outcome variables may not be realistic or reasonable for universal populations of children and adolescents.

Although no differences in outcome were found between one-session and multiple session interventions in this study, it is important to note that these effects were only examined at posttest. The lack of follow-up data for one-session interventions represents a significant limitation in the ability to draw firm conclusions about the importance of intervention length when devising an eating disorder prevention program. Although results from this study suggest that at posttest there were no significant differences in outcome based on duration of intervention, future studies that investigate "one-shot" interventions should follow the results over a length of time to determine whether these effects can be maintained. An additional area to target in future studies is the inconsistent or reduced ability that was found to influence self-esteem and negative affect. These two variables demonstrated relatively poor outcome as effects were either found to be inconsistent or statistically non-significant. These findings are not necessarily surprising given the reduced number of studies involved in these analyses

and the relatively low sample size per study within each outcome set. Moreover, negative affect and self-esteem are broadly defined constructs which can be difficult to measure.

A final recommendation offered for improving research in this area is to actually measure the incidence of eating disorders in study participants prior to and following an intervention. Although the findings presented here have implications for the prevention of eating disorders, it is argued that a clear link between intervention efficacy and a decreased incidence of eating disorders was not demonstrated. Researchers in this field tend to rely on changes in attitudinal and behavioral risk factors as indicators of intervention efficacy to the exclusion of assessing for the specific constellation of symptoms that signal the presence of an eating disorder. As such, accurate conclusions about whether or not these programs can effectively prevent eating disorders cannot be drawn at this time. Rather, information is only offered about the degree to which these interventions effectively influence problematic eating attitudes and behaviors that may be related to the development of an eating disorder

CONCLUSIONS

Eating disorder prevention programs were found to be effective in promoting knowledge and in reducing maladaptive eating behaviors and attitudes. These findings challenge previously drawn conclusions about the ineffectiveness of interventions in this area and alleviate concerns about the potential iatrogenic effects of including psychoeducational material on eating disorders in such programs. The value of focusing on population targeted was demonstrated and additional recommendations were offered to guide the development and implementation of future interventions in this area. Considerable emphasis was given to exploring the practical and clinical significance of the findings from this study, and it strongly recommended that quantitative methods continue to be applied when integrating the findings across studies in this field.

REFERENCES

Abelson, R.P. (1985). A variance explanation paradox: When a little is a lot. *Psychological Bulletin, 97*, 129-133.

Agras, W.S. (2001). The costs and consequences of eating disorders. *Psychiatric Clinics of North America, 24*, 371-379.

Andrewes, D., Say, S., & McLennan, J. (1995). A self-administered computer-based educational program about eating disorder risk factors. *Australian Psychologist, 30,* 210-212.

Austin, S.B. (2000). Prevention research in eating disorders: Theory and new directions. *Psychological Medicine, 30,* 1249-1262.

Barenholtz, D.E. (1995). The effects of an exercise program on the eating behavior, body image and self-esteem of adolescent girls. (Doctoral dissertation, California School of Professional Psychology – San Diego, 1995) *Dissertation Abstracts International, 56,* 2853.

Becker, B.J. (1988). Synthesizing mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41,* 257-278.

Borenstein, M., & Rothstein, H. (1999). *Comprehensive meta-analysis. A computer program for research synthesis* [computer software and manual]. Englewood, NJ: Biostat, Inc.

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

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

Cassell, D.K., & Gleaves, D.H. (2000). *The Encyclopedia of Obesity and Eating Disorders (2nd ed.).* New York: Facts On File, Inc.

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

Chase, A.K. (2001). Eating disorder prevention: An intervention for "at-risk" college women (Doctoral dissertation, The University of Texas at Austin, 2001) *Dissertation Abstracts International, 62*, 1568.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Coller, T.G., & Neumark-Sztainer, D. (1999). Taste of Food, Fun, and Fitness: A community-based program to teach young girls to feel better about their bodies. *Journal of Nutrition Education, 31*, 283E

Earnhardt, J.L., Martz, D.M., Ballard, M.E., & Curtin, L. (2002) A writing intervention for negative body image: Pennebaker fails to surpass the placebo. *Journal of College Student Psychotherapy, 17*, 19-35.

Fairburn, C.G. (1995). The prevention of eating disorders. In. K.D. Brownell & C.G. Fairburn (Eds.). *Eating disorders and obesity: A comprehensive handbook*. (pp. 289-293). New York: The Guilford Press.

Fingeret, M.C. (2002). Weighing in on the controversy: How effective are eating disorder prevention programs? Unpublished major area paper, Texas A&M University. College Station

Franko, D.L. & Orosan-Weine, P. (1998). The prevention of eating disorders: Empirical, methodological and conceptual considerations. *Clinical Psychology: Science and Practice, 5*, 459-477.

Garvin, V. & Streigel-Moore, R.H. (2001). Health services research for eating disorders in the United States: A status report and a call to action. In R.H. Streigel-Moore & L. Smolak (Ed). *Eating disorders: Innovative directions in science and practice*. (pp. 135-152). Washington DC: American Psychological Association.

Gary, A. (2001). The pathophysiology of eating disorders. In J. Robert-McComb (Ed). *Eating disorders in women and children: Prevention, stress management and treatment* (pp. 49-58). Boca Raton, FL: CRC Press, Inc.

Gordon, R.S. (1983). An operational classification system of disease prevention. *Public Health Reports, 98*, 107-109.

Grave, R.D., DeLuca, L., & Campello, G. (2001).Middle school primary prevention program for eating disorders: A controlled study with a twelve-month follow-up. *Eating Disorders: The Journal of Treatment and Prevention, 9*, 327-337.

Gurney, V.W. (1997). An eating disturbances intervention program for college students. (Doctoral dissertation, Ohio University, 1997). *Dissertation Abstracts International, 58*, 2675.

Huon, G.F., Roncolato, W.G., Ritchie, J.E., & Braganza, C. (1997). Prevention

of dieting-induced disorders: Findings and implications of a pilot study. *Eating Disorders: The Journal of Treatment and Prevention, 4*, 280-293.

Hutcuff A.I., & Arthur, W.A. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology, 80*, 327-334.

Irving, L.M. (2000). Promoting size acceptance in elementary school children: The EDAP Puppet Program. *Eating Disorders: The Journal of Treatment and Prevention, 8, 221-232.*

Irving, L.M., DuPen, J., & Berel, S. (1998). A media literacy program for high school females. *Eating Disorders: The Journal of Treatment and Prevention, 6*, 119-131.

Johnson, B.T. (1993). *DSTAT 1.10: Software for the meta-analytic review of research literatures*. Hillsdale, NJ: Erlbaum.

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

Killen, J.D., Taylor, C.B., Hammer, L.D., Litt, I., Wilson, D.M., Rich, T., Hayward, C., Simmonds, B., Kraemer, H., & Varady, A. (1993). An attempt to modify unhealthful eating attitudes and weight regulation practices of young adolescent girls. *International Journal of Eating Disorders, 13*, 369-384.

Kusel, A.B. (1999). Primary prevention of eating disorders through media literacy training of girls (Doctoral Dissertation, California School of Professional Psychology, 1999). *Dissertation Abstracts International, 61,* 524.

Landis J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.

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

Mann, T., Nolen-Hoeksema, S., & Burgard, D. (1997). Are two interventions worse than one? Joint primary and secondary prevention of eating disorders in college females. *Health Psychology, 16*, 215-225.

Martz, D.M. & Bazinni, D.G. (1999). Eating disorders prevention programming may be failing: Evaluation of 2 one-shot programs. *Journal of College Student Development, 40*, 32-42.

Mason, M.A. (1998). An educational intervention for the prevention of eating disorders in college female athletes: A risk factor approach. (Doctoral dissertation, Florida State University, 1997). *Dissertation Abstracts International, 58,* 2548.

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

Morris, S.B. & DeShon, R.P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*, 105-125.

Mussell, M.P., Binford, R.B., & Fulkerson J.A., (2000). Eating disorders: Summary of risk factors, prevention programming, and prevention research. *The Counseling Psychologist, 28,* 764-796.

Nelson, L.D. (1996). A comparison of gender differences in middle school adolescents

participating in an eating disorder prevention program (Doctoral Dissertation, State University of New York at Buffalo, 1999). *Dissertation Abstracts International, 57*, 2426.

Neumark-Sztainer, D., Sherwood, N.E., Coller, T., & Hannan, P.J. (2000). Primary prevention of disordered eating among preadolescent girls: Feasibility and short-term effect of a community based intervention. *Journal of the American Dietetic Association, 1000,* 1466-1473.

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

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

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

Phelps, L., Sapia, J., Nathanson, D., & Nelson, L. (2000). An empirically supported eating disorder prevention program. *Psychology in the Schools, 37*, 443-452.

Posavac, H.D., Posavac, S.S., & Weigel, R.G. (2001). Reducing the impact of media images on women at risk for body image disturbance: Three targeted interventions. *Journal of Social and Clinical Psychology, 20*, 324-340.

Prentice, D., & Miller, D.T. (1992). When small effects are impressive. *Psychological*

Bulletin, 112, 160-164.

Richman, R.D. (1998). Preventing disordered eating; promoting healthy attitudes and behaviors: A school-based program. (Doctoral dissertation, Simon Fraser University, Canada, 1997), *Dissertation Abstracts International, 59*, 425.

Rocco, P.L., Ciano, R.P., & Balestrieri, M. (2001). Psychoeducation in the prevention of eating disorders: An experimental approach in adolescent schoolgirls. *British Journal of Medical Psychology, 74, 351-358.*

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin, 86*, 638-641.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage Publications, Inc.

Rosenthal, R., & DiMatteo, M.R. (2001). Meta-Analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52,* 58-82.

Rosenthal, R., & Rubin, D.R. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*, 166-169.

Rosenvinge, J.H., & Borresen, R. (1999). Preventing eating disorders – Time to change programmes or paradigms? Current update and further recommendations. *European Eating Disorders Review, 7*, 5-16.

Ruffo, C.E. (1996). Resistance training added to an educational program for

dysfunctionally eating women reporting body-image dissatisfaction. (Doctoral dissertation, Northern Arizona University, 1996), *Dissertation Abstracts International 57*, 1505.

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

Sapia, J.L. (2000). Determining the efficacy of an eating disorder prevention program among college women (Doctoral Dissertation, State University New York at Buffalo, 2000). *Dissertation Abstracts International, 61*, 2780.

Schmidt, F.L. (1996). Statistical significance testing and cumulative knowledge in psychology: Implications for training of researchers. *Psychological Methods, 1,* 115-129.

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

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

Smolak, L., Levine, M.P., & Schermer, F. (1998b). Lessons from lessons: An evaluation of an elementary school prevention program. In W. Vandereycken & G. Noorbendos (Eds.). *The prevention of eating disorders*. New York: New York University Press.

Steiner-Adair, C., Sjostrorn, L., Franko, D.L., Pai, S., Tucker, R., Becker, A.E., & Herzog, D.B. (2002). Primary prevention of risk factors for eating disorders in adolescent girls: Learning from practice. *International Journal of Eating Disorders, 32*, 401-411.

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

Stice, E., Chase, A., Stormer, S., & Appel, A. (2001). A randomized trial of a dissonance-based eating disorder prevention program. *International Journal of Eating Disorders, 29*, 247-262.

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

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

Stormer, S.M. (1999). The cross-gender effects of an experimental media-focused

psychoeducation program. (Doctoral dissertation, University of South Florida, 1998). *Dissertation Abstracts International, 59,* 6080.

Varnado-Sullivan, P.J., Zucker, N., Williamson, D.A., Reas, D., Thaw, J., & Netemeyer, S.B. (2001). Development and implementation of the body logic program for adolescence. *Cognitive and Behavioral Practice, 8,* 248-259.

Waggoner, I.R.M. (1999). Cognitive-behavioral therapy and cognitive therapy for body image awareness in 6th grade females (Doctoral Dissertation, Auburn University, 1998. *Dissertation Abstracts International, 59*, 5060.

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

Withers, G.F., Twigg, K., Wertheim, E.H., & Paxton, S.J. (2002). A controlled evaluation of an eating disorders primary prevention videotape using the Elaboration Likelihood Model of Persuasion. *Journal of Psychosomatic Research, 53*, 1021-1027.

Wolfe, J.L. (1992). Applying cognitive dissonance within a personal health curriculum to reduce risk factors associated with the development of eating disorders in a university population (Doctoral Dissertation, University of Georgia, 1992). *Dissertation Abstracts International, 53*, 4221.

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

Wood, P. (1995). Meta-Analysis. In G.M. Breakwell, S. Hammond, & C. Fife-Schaw (Eds.), *Research methods in psychology (2nd ed.)* (pp. 414-425). London: Sage Publications.

Wood, W. & Christensen, P.N. (in press). Quantitative research synthesis: Examining study outcomes over setting, samples, and time. In C.C. Morf & A. T. Panter (Eds). *Handbook of methods in social psychology*. Thousand Oaks, CA: Sage.

Zabinski, M.F., Pung, M.A., Wilfley, D.E., Eppstein, D.L., Winzelberg, A.J., Celio, A., & Taylor, C.B. (2001). Reducing risk factors for eating disorders: Targeting at-risk women with a computerized psychoeducational program. *International Journal of Eating Disorders, 29,* 401-408.

Zabinski, M.F., Wilfley, D.E., Calfas, K.J., Winzelberg, A.J., & Taylor, C.B. (2003).An interactive psychoeducational intervention for women at-risk of developing an eating disorder. Unpublished manuscript.

Zanetti, T., Ferrara, S., Favaro, A., & Santonastaso, P. (1999, September). *Teaching teachers prevention interventions for eating disorders: A controlled study*. Paper presented at the meeting of the European Council on Eating Disorders, Stockholm.

APPENDIX A

CODING FORM

Reference:

Study ID:

Type of Publication:

1) journal article
2) book chapter
3) thesis or doctoral dissertation
4) conference paper
5) unpublished data provided by author
6) other (specify) : _____

STUDY DESCRIPTORS

Study Design:

1) uncontrolled
2) quasi-experimental
3) experimental

Study Setting:

1) primary
2) secondary school
3) college
4) broader community setting
5) other (specify): _____

Age Range of Participants: _____
Mean Age of Participants: _____ **SD**: _____

Gender of Study Participants:

1) females only
2) females and males

Population Targeted:

1) universal
2) selective
3) indicated

Overall Confidence of judgment on population targeted:

1) very low (little basis)
2) low (guess)
3) moderate (weak inference)
4) high (strong inference)
5) very high (explicitly stated)

Total Sample Size (Start of Study): _____

Treatment Group Sample Size (Start of Study): _____
Comparison Group Sample Size (Start of Study): _____

Type of Intervention Strategy:

1) purely psychoeducational
2) enhanced psychoeducational
3) purely interactive

Overall Confidence of judgment on intervention strategy

1) very low
2) low
3) moderate
4) high
5) very high

Was information related to eating disorders included in the intervention?

1) yes
2) no (specify a OR b)
    a. This information was *purposefully* excluded from the intervention due to concerns about iatrogenic effects.
    b. There was no explicit indication that the exclusion of this material was driven by concerns about iatrogenic effects.

Nature of Comparison Group:

1) no treatment control
2) delayed control (wait-list)
3) alternative treatment (specify: _____)

Duration of Intervention

1) single session (time interval: _____)
2) multiple discrete sessions (number of sessions _____, time interval: _____)
3) continuous/ongoing programming (time interval: _____)

Total Length of Study (including any follow-up periods): _____

Assessment Periods:

1) post-test ONLY
2) pre-test, post-test
3) pretest, post-test, and follow-up

EMPIRICAL FINDINGS/EFFECT SIZE INFORMATION

Total Number of Outcome Measures Employed: _____

Measure Employed	Purports to Assess (variable name)	Created for the study?	Psychometric properties reported

Was the equivalence of scores between groups at baseline tested?

1) no
2) yes: specify how:

   _____

   a. differences considered negligible and judged unimportant
   b. differences found were statistically significant and considered meaningful (explain further:

      _____
      _____)

Did analyses account for baseline scores?

1) no
2) yes
   a. with gains scores or change scores
   b. by using a covariate (e.g., ANCOVA)
   c. repeated measures analysis (e.g., MANOVA)
   d. other (specify: _____)

Multivariate Effects

Name of Measures used:

Univariate Effects

OUTCOME VARIABLE* = _____ *(variable name)
Name of Measure used: _____

Data Available for:
1) post-test
2) follow-up (length of follow-up period _____)

Type of Outcome Data Presented: (circle all that apply)
1) means and standard deviations
2) t value or F-value
3) p value
4) chi-squared
5) frequencies or proportions
6) effect size (specify what type: _____)

Page Number or Table where Raw Data Are Found : _____

Treatment Group Sample Size: _____ Comparison Group Sample Size: _____

If data are reported for only a subset of the initial sample – please describe here:
_____
_____

POST-TEST data:

FOLLOW-UP data:

Group X Time Interaction :

OUTCOME VARIABLE* = _____ *(variable name)
Name of Measure used: _____

Data Available for:
1) post-test
2) follow-up (length of follow-up period _____)

Type of Outcome Data Presented: (circle all that apply)

1) means and standard deviations
2) t value or F-value
3) p value
4) chi-squared
5) frequencies or proportions
6) effect size (specify what type: _____)

Page Number or Table where Raw Data Are Found : _____

Treatment Group Sample Size: _____Comparison Group Sample Size: _____

If data are reported for only a subset of the initial sample – please describe here:

_____

_____

POST-TEST data:

FOLLOW-UP data:

Group X Time Interaction:

APPENDIX B

CODING MANUAL

The first three items: **Reference, Study ID, Type of Publication** should already be filled out by the principal investigator.

Study Design:

1) An *uncontrolled study* design does not include a control/comparison group. If the study uses an uncontrolled study design DO NOT continue coding.
2) A *quasi-experimental* study design uses a control group but does not include the procedure of randomly assigning individuals to treatment and control conditions. Common examples of quasi-experimental designs include:
    - Taking the first 50 participants and assigning them to a treatment condition and placing remaining participants in a wait-list control condition
    - Randomly assigning the treatment and control conditions to different schools or classes
3) An *experimental* study design randomly assigns individuals to groups AND randomly assigns the groups to treatment or control status.

Study Setting: This particular item does not seem to require much elaboration. However, if there is something unique about the setting (e.g., residential ballet school) go ahead and make a note of that on the form.

Age Range of Participants: If not provided by authors – put N/A for not available.

Mean Age of Participants and SD: If not provided – put N/A.

Population Targeted: This classification system was initially proposed by Gordon (1983) and is based on a re-conceptualization of how to classify preventive interventions. In this model, a preventive intervention only refers to an action or intervention that occurs before the onset of a disorder. Thus, if any of the studies you are coding include individuals that have been diagnosed with an eating disorder – this would fall outside the realm of prevention literature (and into the realm of treatment). Please note whether any of the studies you are coding seem to involve treatment rather than prevention. Otherwise, the population being targeted in the study can be classified in the following manner:

1) *universal* : Efforts directed toward the general public or to a whole population group that has not been identified on the basis of increased risk for developing a mental disorder. The most common example includes the use of children and/or adolescents in a school or community setting (especially when both male and females are included in the sample) as participants.

2) *selective*: Efforts targeted to individuals or specific groups whose risk of developing a clinically diagnosable mental disorder is higher than average. This would include female high school students, female college students, gymnasts, ballet dancers, sorority members and other high-risk groups that have not been previously identified as having signs or symptoms of an eating disorder.
3) *indicated:* Efforts targeted to high-risk individuals who, although they do not meet DSM-IV criteria for a mental disorder, are identified as having minimal but detectable signs of a particular disorder. This would encompass a study in which participants were screened prior to inclusion for elevated scores on body image and/or weight control measures, previous history of dieting, preoccupation with weight or body image concerns, or other personal motivations for entering a study on body image. (*Note. Pre-screening and elevated scores are not necessary to qualify, can also be coded as indicated if recruitment was intended to select for people with perceived or self-reported symptoms)

*** Note. If a study initially targets a universal population and then selects out some segment of that population for further analysis (e.g., high risk participants) – code the study as targeting a universal population and record the data for the entire sample. If data are also reported for the high-risk subset, utilize the section at the end of the coding form to report these data.

** Most researchers still conceptualize prevention as primary, secondary, and tertiary – note that these terms do not directly map onto the ones presented above. However, primary prevention efforts are generally seen as being directed at a universal population. Secondary prevention can map onto either selective or indicated – so just read over the guidelines presented above when making a determination. Tertiary preventive efforts generally encompass treatment and thus do not map onto any of the ones presented above.

I have also included a rating scale for you to record your confidence level of rating this item. If the authors of the study explicitly state the population targeted within this framework and it is in accordance with the framework laid out above – that should increase the confidence of your rating. However, if you have to make an inference from the information presented – then use the rating scale accordingly.

Sample Size: Note that this question refers to the sample size at the start of the study. At a later time you are asked to code for the sample size relevant to the outcome data that are reported.

Type of Intervention Strategy: There is a great deal of variability regarding the types of intervention strategies employed across the studies in this area. I will first provide general descriptions of each category, and then give specific examples of common

intervention strategies used and how they would be coded. ** Keep in mind that there will often be multiple strategies used within a given study. This item refers to the collective use of intervention strategies, you are NOT coding each strategy separately.

1) *purely psychoeducational*: This refers to the provision of didactic/descriptive information to participants as the sole or predominant intervention strategy. Common subjects covered in psychoeducational interventions in this area include: pubertal development, determinants of body size and shape, genetic diversity of body shapes, social influences on body image, cultural determinants of attractiveness, and descriptive information about eating disorders. Studies coded in this category can also encompass group discussion, homework assignments, experiential games and activities, but there is a notable absence of explicit skill-building techniques.
2) *enhanced psychoeducational (or CBT psychoeducational)*: This refers to the provision of didactic/descriptive information in addition to the incorporation of skills-based intervention components such as those drawn from CBT therapies (e.g. progressive muscle relaxation, goal oriented problem solving, self-monitoring of eating habits, identifying and challenging cognitive distortions)
3) *purely interactive (or non-psychoeducational)*: This refers to the use of a purely interactive approach to intervention with a distinct lack of a psychoeducational component.

**Note that even information related to eating disorders is specifically eliminated (see also next question) – there may still be some didactic/descriptive information presented on body image or other related topics. If this were the case then you would code the psychoeducational component of the study.

Examples of common intervention strategies used and how they would typically be coded:

- undergraduate body image course = *purely psychoeducational* (However, be sure to check carefully that skills-based elements were not included)
- Experiential games and activities employed with children and adolescents often tend to be ways of presenting psychoeducational material (e.g., game shows about body image) However, some games and activities may NOT be ways of providing didactic information; for example, those involving art, dance, or music as ways of creative expression.
    - In the event that these types of activities are not paired with psychoeducational material they would be considered a *purely interactive* (or non-psychoeducational) technique.
    - In the event that these types of activities are included as one component of an intervention with other didactic information then code as enhanced psychoeducational.
- Skills-based intervention components such as those drawn from CBT treatment of eating disorders (e.g, self-monitoring of eating habits, identifying and challenging cognitive distortions, stress inoculation, goal-oriented problems

solving) are considered *enhanced psychoeducational (or CBT psychoeducational)* strategies because this type of approach provides skills and information extending beyond mere didactics.

- Media literacy training is generally considered to be a type of intervention strategy that combines didactic, interactive, and skills-based approaches to intervention – as such it would be coded *enhanced psychoeducational (or CBT psychoeducational)*. This intervention strategy tends to involve teaching participants about the dangers associated with the idealization of thinness, increasing awareness of how thinness is promoted in the media, and enhancing the knowledge of deceptive media tactics to create images of perfection. An additional component of media literacy training is the provision of tools for resisting pressures to conform to the images depicted and messages conveyed in the media (e.g., critical thinking skills).

Refer to item on **population targeted** for information about the confidence rating scale.

Inclusion of Eating Disorder Information: This item broadly refers to whether psychoeducational material about eating disorders or unhealthy weight regulation techniques (description, symptoms, causes, consequences, treatment) is incorporated into the intervention. If information about eating disorders is NOT included – I then want you to determine whether this information is being purposefully excluded from the intervention due to concerns about iatrogenic treatment effects. Inference or speculation should not be used to make a determination on this item – only answer A if author(s) explicitly state that the reason they are excluding eating-disorder related material is due to concerns about harmful treatment effects.

Nature of Comparison Group: This item refers to the type of comparison group employed by researchers.

1) *no treatment control*: This indicates that there was no treatment of any kind to the control group during the study or afterward
2) *delayed control (wait-list):* This indicates that those in the control group were placed on a wait-list and offered participation in the intervention following the duration of the initial study period
3) *alternative treatment* : This indicates that an alternative form of treatment (generally considered to be innocuous) was offered to participants in the control group. This can include providing psychoeducational material on alternative topics (e.g., health education, nutrition). It is important to specify the nature of the alternative treatment being provided to the control group if offered

Duration of Intervention: This item refers to the number of discrete sessions and/or time interval during which the intervention was presented.

1) *single session*: If the intervention was implemented in a single session please provide the duration (time interval) of the session in the unit of time reported by the author (i.e., minutes or hours).
2) *multiple discrete sessions*: Indicate the number of sessions used to implement the intervention as well as the time
   Examples:
   - number of sessions: 5 , time interval: 90 minutes each, 1 time per week for 5 weeks
   - number of sessions: 10 time interval: 60 minutes each, 2 times per week for 5 weeks
3) *continuous programming*: This refers to any study in which intervention efforts lasted continuously for a discrete period of time. Provide the duration (time interval) of the intervention period as reported by the author (i.e., weeks, months).

Total Length of Study : This item refers to the total length of the study including any follow-up periods. Report the study length in the unit of time reported by the author (weeks/months)

Assessment Periods: This item refers to the study design regarding data collection.
1) data collected at post-test ONLY
2) data collected at pre-test and post-test
3) data collected at pre-test, post-test, and follow-up

EMPIRICAL FINDINGS/EFFECT SIZE INFORMATION

Total Number of Outcome Measures Employed: This refers to the number of distinct outcome measures used in the study. You will want to carefully scan the methods and results section to determine the nature of each outcome measure employed and what variable the measure is purported to measure.

**For this item – an outcome measure can consist of either a composite scale score OR a subscale score as long as it is used a distinct entity. You should be able to determine whether a subscale score is serving as an outcome measure by examining the data presented in the results section. For example, you should be able to tell if EDI-2 composite scores are reported or whether the data are broken down by subscales (EDI-BD, EDI-DT, EDI-B). If the subscale scores are reported then each subscale would serve as an independent outcome measure. Note also that if both a composite score is reported AND the subscales scores are reported then you should report this information separately (that is report them as separate outcome measures)

- Example: If scores are reported for the EDI composite, EDI-BD, EDI-DT, and EDI-B – then 4 separate outcome measures were used. Although these are technically all part of the same instrument, they are being reported as 4 unique sets of data.

- If scores are only reported for the EDI-BD, EDI-DT, and EDI-B subscales and not the composite score then 3 separate outcome measures were used.
- Similarly, if data are reported for the SATAQ scale as a whole then only one outcome measure was employed; however, if data are reported separately for the SATAQ-A and SATAQ-I subscales and no data are presented regarding the composite score then 2 separate outcome measures were employed.

** When making determinations about what an outcome measure is purported to assess – report what the author states he/she/they are using the measure to assess.

Created for the Study?: Answer YES if the measure was created and/or designed specifically for the study. The most common example of this type of measure involves one assessing psychoeducational components of the program.

Psychometric Properties Reported: Simply answer YES or NO

Equivalence of Scores Between Groups at Baseline:
Note whether the first step in the data analysis was to test for the equivalence of scores between groups at baseline. F or t tests are usually conducted to test for equivalence. If differences were found – briefly describe the strategy used to deal with these discrepancies.

Accounting for Baseline Scores:
Report whether researchers utilized statistical techniques that accounted for baseline scores. Direct comparison of post-test scores (with t tests or F tests) is the most common type of analysis that does NOT account for baseline scores. The three most common ways to account for baseline scores are provided as potential responses:

- t tests of F tests on the gains scores or change scores (between pre and post test for example)
- ANCOVAs or MANCOVAs (covariate analysis)
- Repeated measures analysis (MANOVA)

Multivariate Effects
Although rarely used in the eating disorder prevention literature, there are some studies that conduct multivariate analyses. Because these analyses combine and utilize multiple outcome measures – report the values (Wilk's F or whatever stat used) and then which measures are being utilized in the analysis). If follow-up data are reported – go ahead and include that information in the box provided and make sure to indicate the direction of effect.
** When reporting the direction of effect – just indicate whether the effect is in the EXPECTED direction or UNEXPECTED direction. An effect in the expected direction generally means that the intervention group improved relative to the control group whereas an effect in the unexpected direction means the control group improved relative to the intervention group.

Univariate Effects
For this section, you will record the data for each outcome variable separately. (if more than one measure was used to assess a given outcome variable then you would report that separately as well. For example if two measures were used to assess internalization, you would name the first outcome variable internalization1 and the second outcome variable internalization2).

Start with the 1st outcome variable listed in the chart on p. 2 and record the variable name that you have listed on the right-hand side of the chart (e.g., frequency of dieting, body dissatisfaction, internalization) and the name of the measure used.

Scan the article to determine whether outcome data are presented for post-test only and/or for follow-up. If follow-up data are included, make sure to record the length of the follow-up period. Circle all that apply for the types of outcome data presented, and record the page number and/or table where the raw data can be found. *** If outcome data are not presented but the methods section indicated that these data were collected as part of the study make a notation and then move on to the next outcome variable. There will be instances where some of the outcome data are not presented.

Record the sample size for the treatment group and for the control group that the data are based on. Then provide the raw data for this outcome measure only. If data are reported for only a subset of the initial sample – provide descriptive information (females only, males only, high-risk sample (BSQ > 50) etc). You should fill out a separate outcome variable recording page for each subset of data reported. (EX: internalization 1 females only, internalization 1 males only).

Post-test data: Record all data that are available for a given measure

- *Means/SDs*: If raw means and standard deviations are presented, record the means and SDS for pretest and post-test. If only adjusted means are presented- then record these means and make a notation that they are adjusted – adjusted for what? – also report whether the effects are in the expected or unexpected direction.
- *T values OR F values*: Record the exact value provided and include the corresponding p value if available (note whether the p value is exact or approximate) – report whether the effect is in the expected or unexpected direction.
- *Chi squared* – record the exact value and whether the effect is in the expected or unexpected direction.
- *Frequencies/Proportion*: record the exact values, and include pretest and posttest data - and whether the effects are in the expected or unexpected direction.
- *Effect Size* – record this information if available and what type of effect size being reported (e.g., r, d)

Follow-up data: Use the same format as above, and make sure to record PRE-TEST and FOLLOW-UP values if applicable.

Repeat this process for each outcome variable and/or subset of sample. Please start with data based on the entire sample and then proceed to recording data based on subsets of the initial sample.

** when data are presented for individual items (most often this will occur with frequency data for measures generated specifically for the study) – do your best to record a composite score for a measure rather than individual items – but if this is all that is presented, go ahead and record the frequency data (both pre and post test) for each item – write down what the item is and the frequency of responses

APPENDIX C

PRIMARY STUDY FEATURES FOR EACH STUDY INCLUDED IN THE META ANALYSIS

Reference	Strategy	Population Targeted	Inclusion of Eating Disorder Information	Total Number of Sessions[a]	Number of Sessions per Week	Length of Follow-up[b]
Andrewes et al., (1995)	psychoeducational	selective	Yes	1	1	
Barenholtz (1995)	enhanced	selective	No	27	4.5	10 weeks
Buddeberg-Fischer et al., (1998)	enhanced	selective	Yes	3	.3	
Celio et al., (2000)	psychoeducational	indicated	Yes	8	1	16 weeks
Chase (2001)	enhanced	indicated	Yes	3	1.5	4 weeks
Earnhardt et al., (2002)	interactive	indicated	No	4	4	4 weeks
Grave et al., (2001)	enhanced	universal	Yes	6	1	52 weeks
Gurney (1997)	enhanced	selective	No	1	1	
Irving et al., (1998)	enhanced	selective	No	1	1	
Kaminski & McNamara (1996)	enhanced	indicated	Yes	8	1	5 weeks
Killen et al., (1993)	enhanced	selective	Yes	18	1	20 weeks
Kusel (1999)	enhanced	universal	No	2	2	12 weeks
Mann et al., (1997)	psychoeducational	selective	Yes	1	1	12 weeks
Martz & Bazinni (1999)	psychoeducational	selective	Yes	1	1	
Martz & Bazinni (1999)	enhanced	selective	Yes	1	1	
Mason (1998)	enhanced	selective	No	10	2	
McVey & Davis (2002)	enhanced	selective	No	6	1	52 weeks
Nelson (1996)	psychoeducational	universal	Yes	5	NA	
Neumark-Sztainer et al., (2000)	enhanced	universal	No	6	.5	12 weeks
Nicolino et al., (2001)	enhanced	selective	Yes	1	1	
Paxton (1993)	psychoeducational	selective	Yes	5	1	52 weeks
Phelps et al., (2000)	enhanced	selective	Yes	4	NA	
Posovac et al., (2001)	psychoeducational	indicated	No	1	1	
Richman (1998)	psychoeducational	universal	Yes	6	1	

Reference	Strategy	Population Targeted	Inclusion of Eating Disorder Information	Total Number of Sessions	Number of Sessions Per Week	Length of Follow-up
Rocco et al., (2001)	psychoeducational	selective	Yes	9	.3	
Ruffo (1996)	enhanced	indicated	Yes	12	2	
Santonastaso et al., (1999)	psychoeducational	selective	Yes	4	1	52 weeks
Sapia (2000)	enhanced	selective	Yes	4	1	
Shepard (2001)	psychoeducational	universal	Yes	5	5	8 weeks
Smolak, Levine & Schermer (1998[a])	psychoeducational	universal	No	10	NA	
Smolak, Levine & Schermer (1998[b])	psychoeducational	universal	No	10	NA	
Steiner-Adair et al., (2002)	enhanced	selective	No	8	1	24 weeks
Stewart et al., (2001)	enhanced	selective	Yes	6	1	24 weeks
Stice & Ragan (2002)	psychoeducational	selective	Yes	30	2	
Stice et a l., (2000)	psychoeducational	indicated	No	3	1	4 weeks
Stice et al., (2001)	enhanced	indicated	No	3	1	4 weeks
Stice et al., (2003)	enhanced	indicated	Yes	3	1	12 weeks
Stormer (1999)	psychoeducational	universal	No	1	1	3 weeks
Varnado-Sullivan et al., (2001)	enhanced	universal	No	3	1	11 weeks
Waggoner (1999)	enhanced	selective	No	7	1	
Winzelberg et al., (1998)	psychoeducational	selective	Yes	Multiple	NA	12 weeks
Withers et al., (2002)	psychoeducational	selective	Yes	1	1	4 weeks
Wolf-Bloom (1999)	enhanced	selective	No	1	1	
Zabinski et al., (2001)	psychoeducational	indicated	Yes	Multiple	NA	10 weeks
Zabinski et al., (2003)	enhanced	indicated	No	Multiple	NA	10 weeks
Zanetti et al., (1999)	psychoeducational	selective	Yes	6	1	

Note. [a]For total number of sessions, cells labeled as multiple refer to studies employing continuous or ongoing programming efforts over a specified period of time as opposed to offering discrete intervention sessions. [b]Cells left blank in the length of follow-up column either signified that there was no follow-up period or that follow-up data collected were not on outcome variables employed in the meta-analysis.

APPENDIX D

SUMMARY OF EFFECT SIZE DATA ACROSS ALL STUDIES INCLUDED IN THE META-ANALYSIS

Reference	Knowledge Posttest	Knowledge FU	Eating Pathology Posttest	Eating Pathology FU	Dieting Posttest	Dieting Fu	Internalization Posttest	Internalization FU	BD Posttest	BD FU	Negative Affect Posttest	Negative Affect FU	Self-esteem posttest	Self-esteem FU
Andrewes et al., (1995)	3.81*		0.62											
Barenholtz (1995)									0.11	0.26			0.06	
Buddeberg-Fischer et al., (1998)			0.10											
Celio et al., (2000)					0.16	0.35			0.28	0.52				
Chase (2001)	1.17	1.06			-0.16	0.18	0.27	0.12	0.27	0.12				
Earnhardt et al., (2002)			-0.11	-0.21	0.07	0.21			-0.01	-0.13	0.02	0.14		
Grave et al., (2001)	2.00	1.40	0.22	-0.21	0.01	-0.36							0.12	-0.73
Gurney (1997)					0.08		-0.04		-0.09					
Irving et al., (1998)							0.67		0.15		0.48			
Kaminski & McNamara (1996)							1.40	1.18	2.57*	2.75*			1.68	1.47
Killen et al., (1993)	1.29					0.17								
Kusel (1999)			0.24	0.15			0.39	0.25	0.24	0.06			0.36	0.09
Mann et al., (1997)			-0.09	0.06					0.16	-0.41			0.36	0.31
Martz & Bazinni (1999)					0.12				0.05					
Martz & Bazinni (1999)					0.10				-0.04					
Mason (1998)									0.09				1.02	
McVey & Davis (2002)			-0.03	-0.03					-0.15	-0.06				
Nelson (1996)					0.08				0.02				0.03	
Neumark-Sztainer et al., (2000)	0.31*	0.18					0.13	0.31	0.21	0.18				
Nicolino et al., (2001)					1.92*				1.79*					
Paxton (1993)										-0.23				
Phelps et al., (2000)					0.61				0.65				1.16	

Reference	Knowledge Posttest	Knowledge FU	Eating Pathology Posttest	Eating Pathology FU	Dieting Posttest	Dieting Fu	Internalization Posttest	Internalization FU	BD Posttest	BD FU	Negative Affect Posttest	Negative Affect FU	Self-esteem posttest	Self-esteem FU
Posovac et al., (2001)									0.73					
Richman (1998)	1.24		0.17						0.17					
Rocco et al., (2001)									-0.03					
Ruffo (1996)			0.15						-0.03					
Santonastaso et al., (1999)				0.13						0.13				
Sapia (2000)					0.78				0.63				0.42	
Shepard (2001)	1.00	1.05	0.02	0.03			-0.28	-0.05	-0.17	0.24				
Smolak, Levine & Schermer (1998[a])									0.02					
Smolak, Levine & Schermer (1998[b])									-0.30					
Steiner-Adair et al., (2002)	1.03	0.64					0.08	0.08	0.16	0.41			0.18	0.01
Stewart et al., (2001)	1.38	1.00	0.19	0.14	0.27	0.20							0.04	0.09
Stice & Ragan (2002)			0.50		0.47		1.14		0.62		0.29			
Stice et al., (2000)			0.44	0.50	0.22	0.22	0.49	0.65	0.59	0.73	0.86	0.14		
Stice et al., (2001)			0.14	0.05	0.24	0.24	0.63	0.92	0.32	0.00	0.17	0.10		
Stice et al., (2003)			0.45	0.32	0.03	-0.05	0.36	-0.38	0.15	-0.05	-0.09	0.69		
Stormer (1999)							0.34	0.13	0.19					
Varnado-Sullivan et al., (2001)			0.21	0.27	-0.11	0.20					-0.30	-0.06		
Waggoner (1999)									0.68				0.79	
Winzelberg et al., (1998)									0.61	0.52				
Withers et al., (2002)	1.40	0.39			0.35	0.24			0.13	0.07				
Wolf-Bloom (1999)			0.13				0.04		-0.04				-0.04	
Zabinski et al., (2001)			0.12	0.25	0.23	0.46			0.32	0.29				
Zabinski et al., (2003)			0.52	0.82	-0.02	0.19							-.017	-0.67
Zanetti et al., (1999)			0.10											

Note. Asterisks were used to mark the effect sizes identified as outliers.

VITA

Michelle Cororve Fingeret
5738 Cheena
Houston, TX 77096
mcfingeret@yahoo.com

Education

2003-2004	Clinical Internship, University of Texas Medical School at Houston
2001-2004	Doctor of Philosophy, Texas A&M University, Clinical Psychology
1998-2001	Master of Science, Texas A&M University, Clinical Psychology
1993-1997	Bachelor of Arts, Honors, Washington University, Psychology

Research Experience

2003-2004	University of Texas Medical School at Houston, Department of Psychiatry & Behavioral Sciences
1998-2003	Texas A&M University, Department of Psychology
1997-1998	Baylor College of Medicine, Alzheimer's Disease Research Center
1995-1996	Washington University, Department of Psychology

Clinical Experience

2003-2004	University of Texas Medical School at Houston, Department of Psychiatry & Behavioral Sciences
2003-2004	Harris County Psychiatric Hospital
2001-2003	Baylor College of Medicine, Neuropsychology Service
1999-2003	Texas A&M University, Psychology Department
1999-2001	Bryan Federal Prison Camp, Psychology Department
1997-1998	Baylor College of Medicine, Neuropsychology Service

Publications *(selected)*

Cororve, M.B., & Gleaves, D.H. (2001). Body dysmorphic disorder: A review of conceptualizations, assessments, and treatment strategies. *Clinical Psychology Review, 21*, 949-970.

Williams, T.L., Gleaves, D.H., Cepeda-Benito, A., Erath, S.A., & Cororve, M.B. (2001). The reliability and validity of a group administered version of the Body Image Assessment, *Assessment, 8,* 37-46.

Gleaves, D.H., Lowe, M.R., Green, B.A., Cororve, M.B., & Williams, T.L. (2000). Do anorexia and bulimia nervosa occur on a continuum? A taxometic analysis. *Behavior Therapy, 31*, 195-219.

Gleaves, D.H., Cepeda-Benito, A., Williams, T.L., & Cororve, M.B., Fernandez-Nieto, M.C., & Vila, J. (2000). Body image preferences of self and others: A comparison of Spanish and American male and female college students. *Eating Disorders: The Journal of Treatment and Prevention, 8,* 269-282.