Journal of Counseling Psychology
1999, Vol. 46, No. 3, 370–380

0022-0167/99/$3.00

Effectiveness of Grief Therapy: A Meta-Analysis

Denise Litterer Allumbaugh and William T. Hoyt
Iowa State University

This meta-analysis addressed the question of how effective grief therapy is and for whom, using B. J. Becker's (1988) techniques for analyzing standardized mean-change scores. Analyses were based on 35 studies (N = 2,284), with a weighted mean effect size (ES) of Δ_+ = 0.43 (95% confidence interval = 0.33 to 0.52). Clients in no-treatment control groups showed little improvement (d_+ = 0.06), possibly because of the relatively long delay between loss and treatment in most studies (mean delay = 27 months). Moderators of treatment efficacy included time since loss and relationship to the deceased. Client selection procedures, a methodological factor not originally coded in this meta-analysis, appeared to contribute strongly to variability in ESs: A small number of studies involving self-selected clients produced relatively large ESs, whereas the majority of studies involving clients recruited by the investigators produced ESs in the small to moderate range.

Over 2 million people die in the United States every year (U.S. Bureau of the Census, 1995). For every person who dies, there are numerous people left behind to grieve, many of whom seek help from a therapist. Loss and grief are universal human experiences. The recently bereaved represent a large at-risk population, with higher overall death and suicide rates than age-matched control participants and with an increased incidence of depression, substance abuse, and certain medical disorders (Yalom & Vinogradov, 1988). So, although grief can be viewed as a normal developmental occurrence, and one from which relatively quick return to normal functioning is expected in most cases, the impact of bereavement varies widely, depending on characteristics of the griever and the nature of the loss (Schwartzberg & Halgin, 1991).

It is commonly assumed in the psychotherapeutic community that for at least some bereaved individuals, some form of psychotherapy is useful or even necessary to assist in recovering from loss, and numerous theoretical frameworks have been proposed for psychotherapists working with bereaved clients. The first intrapsychic theory of grief was proposed by Freud (1917/1957), who stated that the fixed amount of energy that was once invested in the deceased person must be retrieved before it can be reinvested in another person. His theory has developed into modern psychoanalytic approaches to grief therapy that allow the bereaved to attribute the unfinished business of mourning to the therapist, who becomes a temporary substitute for the deceased while the bereaved works through guilt and ambivalent feelings that are blocking the way to resolution (Sanders, 1989).

Another major therapeutic approach to bereavement, client-centered therapy, provides nurturance and empathic understanding for the client when safety needs are often unmet and raw feelings are overwhelming (Sanders, 1989). It also works to facilitate the expression of a variety of feelings and enhance the client's insight into the grief experience. Like client-centered therapy, Gestalt therapy focuses on feelings and is used especially to help bereaved people who are denying the enormity of their feelings of loss that have been shelved for a long time (Barbato & Irwin, 1992). In contrast to a focus on feelings, cognitive strategies aim to make the bereaved aware of destructive thought patterns, helping them to contest these irrational beliefs and replace them with rational beliefs. Finally, rather than focusing on grief itself, behavior therapy focuses on the specific behaviors that are impeding the grief resolution (Sobel, 1981). The therapist takes a problem-solving approach in helping the bereaved to operationalize their problems, set goals, and learn new skills that will replace self-defeating behavior (Barbato & Irwin, 1992).

Treatments for normal grief are typically brief, with a focus on working through normal developmental processes and on restoration of healthy functioning. Despite widespread agreement that grief therapy is sometimes desirable for individuals who have experienced a loss, there is as yet little consensus about the efficacy of grief therapy or about client characteristics that might enhance the effectiveness (and therefore the desirability) of grief interventions. Published outcome studies on grief therapy have yielded mixed results (Potocky, 1993), which may be attributable to one or more of three explanations. First, it may be that grief treatments are in general rather ineffective, relative to psychotherapy targeted at other types of psychological problems. Because grief reactions are expected to abate over time, treatment-control comparisons for grief interventions might be expected to yield smaller effect sizes relative to comparisons for treatments of conditions in which improvement in the no-treatment control group is unlikely. If the

Denise Litterer Allumbaugh and William T. Hoyt, Department of Psychology, Iowa State University.

This article is based in part on a master's thesis by Denise Litterer Allumbaugh, completed under the direction of William T. Hoyt.

Correspondence concerning this article should be addressed to Denise Litterer Allumbaugh or William T. Hoyt, Department of Psychology, Iowa State University, W112 Lagomarcino Hall, Ames, Iowa 50011-3180. Electronic mail may be sent to dlittere@iastate.edu or whoyt@iastate.edu.

overall effect size for grief therapy is relatively weak, it would not be unusual to obtain some significant and some nonsignificant results in grief outcome studies due to sampling error.

A second, and related, explanation posits that mixed results in studies of grief therapy are due to variation in statistical power among studies. Even if the overall effect size of treatment-control comparisons in grief therapy studies is moderate to large (as is the case for other psychosocial treatments; Lambert & Bergin, 1994; Lipsey & Wilson, 1993), studies with small sample sizes will have a relatively low probability of detecting statistically significant differences between treatment and control groups (Cohen, 1962, 1988). Indeed, a commonly cited problem with the statistical significance test as a summary of study findings is the likelihood of exaggerating the true variability of those findings (Schmidt, 1992, 1996). By using point estimates and confidence intervals as alternative summary statistics, one may find that apparent variability is no more than was expected due to sampling error.

Finally, it may be that variability in study findings is due not only to sampling error, but also to systematic differences among studies on one or more moderator variables that affect the effectiveness of grief treatment. For example, if, as Raphael (1977) asserted, grief therapy is more effective for clients experiencing some form of pathological bereavement than for those experiencing normative grief—and if the level of risk for abnormal grief varies by study—one would expect study findings to be mixed, with studies involving a high proportion of normal clients yielding lower effect sizes than those involving a high proportion of at-risk clients.

To assess the sources of apparent variation in studies of grief therapy, we conducted a meta-analysis of published and unpublished studies of grief interventions. Meta-analysis addresses questions about the overall effectiveness of grief therapy by synthesizing results across studies to derive omnibus point estimates and confidence intervals combining data for all included studies. By analyzing data from treatment and control groups separately (Becker, 1988), we were able to assess the impact of control group improvement on the overall effect estimate. Meta-analysis takes into account differences in statistical power among studies, because summary effect sizes are weighted aggregates in which each included effect size is weighted by the inverse of its sampling error variance. Thus, more reliable effect estimates are given greater weight in calculating aggregate estimates.

An additional goal of this study was to examine the impact of 12 potential moderator variables, derived from a review of the theoretical and empirical literature on grief, on the effectiveness of grief therapy. These potential moderators included characteristics of the intervention and the client, as well as design features of the studies themselves.

Method

This meta-analysis included published and unpublished studies on the effectiveness of grief therapy. The following methods were used to locate studies for inclusion: (a) computerized search of *PsycLIT* (1974–1996), *Medline* (1985–1996), and *Dissertation Abstracts* (1966–1981) databases using the keywords *bereavement, grief, mourning, psychotherapy, treatment, intervention,* and *outcome;* (b) manual search of *American Journal of Psychiatry,* (1966–1996) *British Journal of Psychiatry* (1966–1996), *Omega* (1970–1996), and *Psychiatry* (1966–1996); (c) manual search of numerous review articles, chapters, and books, as well as the reference lists of all located studies; and (d) contacting prominent grief researchers for unpublished studies.

We selected studies for inclusion in this meta-analysis if they examined the effectiveness of any type of grief therapy, used a quantitative measure of outcome (i.e., compared a treated group with a control group or compared pretreatment and posttreatment scores for treated participants), and reported sufficient information to allow for the computation of an effect size. An exception to this last criterion was made for studies that reported nonsignificant results, but did not report F or t statistics or means and standard deviations. Rather than exclude these effect sizes from the meta-analysis (which would tend to inflate the overall effect size estimate), we chose to use the conservative strategy of including them in all analyses with an effect size of $d = 0$. As a check on the impact of this conservative strategy, we also analyzed the data after eliminating all of the nonsignificant results that did not provide sufficient information to calculate effect sizes. With only a few exceptions (reported below), the pattern of results was unchanged when these effect sizes were omitted from analyses.

Finally, we excluded two studies (Horowitz, Weiss et al., 1984; Lieberman & Yalom, 1992) because they assessed treatment outcome 6 months or more after the end of therapy, whereas the included studies administered posttreatment assessments within 1 week of the end of therapy. Because few studies included both posttreatment and follow-up assessments, only outcome measures administered immediately after the end of treatment were considered in this meta-analysis.

Coding Procedure

We independently coded all studies for 12 potential moderator variables that were thought to be theoretically relevant to the outcome of grief therapy. Seven categorical moderators included characteristics of the treatment, study characteristics, and one client characteristic: for treatment characteristics, (a) practitioner training (professional, professional in training, nonprofessional) and (b) treatment modality (group vs. individual); for study characteristics, (c) type of control group (placebo vs. no treatment), (d) method of group assignment (random vs. systematic), (e) source of outcome measure (observer vs. self-report); and for the client characteristic, (f) level of risk (high vs. normal). In addition, (g) we considered publication status a categorical moderator, although it was not necessary to code this variable. Cohen's kappa (κ; Cohen, 1960), which indicates the proportion of agreement between the coders corrected for chance, ranged from .57 for source of outcome measure to .82 for method of group assignment.

Five continuous moderators included one additional treatment characteristic, (a) number of sessions, and four client characteristics: (b) mean age, (c) gender (percent female), (d) mean length of time since loss, and (e) relationship to the deceased (percent spouses). Intraclass correlations (Shrout & Fleiss, 1979) ranged from .83 for gender of client to .99 for length of time since loss. Differences in coding for both categorical and continuous moderators were resolved through discussion; all analyses were based on consensus codes.

Several variables of theoretical and practical interest were not coded because of lack of sufficient information included in the studies, including therapeutic and client characteristics such as

race, level of education, religiosity, and treatment modality. An additional variable that may affect the magnitude of the effect size in grief outcome studies is specificity of measurement. Although several measures of grief intensity are available, the vast majority of outcome measures used in these studies were generic measures of psychological distress. In fact, so few studies included grief-specific outcome measures that it was not possible to estimate the impact of this design choice on effect size.

Computation of Effect Sizes

Studies on the effectiveness of grief therapy have used either a one-group or a two-group design. The majority of studies ($k = 22$) included two (or more) groups, and results were reported as comparisons of treatment and control groups on the dependent measure. A minority of studies ($k = 13$) included only a single treatment group; results for these studies were reported as comparisons of pretreatment and posttreatment scores.

The ideal method for combining data from studies of these two types is the analysis of standardized mean-change scores recommended by Becker (1988). Standardized mean-change scores are computed for all groups, using the formula

$$g = (X_{post} - X_{pre})/SD_{pre}, \quad (1)$$

where g is the standardized mean-change score, X_{post} is the mean posttreatment score, X_{pre} is the mean pretreatment score, and SD_{pre} is the standard deviation of the pretreatment scores. These scores are then corrected for bias by using procedures developed by Hedges and Olkin (1985). Effect sizes for each study are computed as

$$\Delta = d_T - d_C, \quad (2)$$

where Δ represents the difference in outcome between treatment and control groups, and d_T and d_C are the bias-corrected standardized mean-change scores for treatment and control groups, respectively. Thus, Δ is a standardized estimate of the degree to which improvement in the treatment group exceeded that in the control group—it is a treatment-control comparison that takes pretreatment status in each group into account. For studies using the one-group (i.e., no-control) design, d_C is imputed from data on the control groups in all two-group studies (Becker, 1988).

Advantages of Standardized Mean-Change Scores

Becker's (1988) approach has not, to our knowledge, ever been used in a meta-analysis of psychological interventions. This is unfortunate, because this strategy has several advantages over the usual approach of computing effect sizes on the basis of within-group (pre–post) comparisons for one-group studies and between-group comparisons for two-group studies. The main virtue of the analysis of standardized mean-change scores is that this technique allows one-group and two-group studies to be compared on an equal basis. Assuming a sufficient quantity of two-group studies, a reasonable estimate of change in status in the absence of treatment can be obtained, and this imputed value is used as a basis for comparison for the one-group (treatment-only) studies.

A second benefit of the use of standardized mean-change scores is that this technique resolves the dilemma of what to do when significance tests in the original study reports (and therefore effect sizes in the meta-analysis) do not take pretreatment status into account. When primary studies that used the two-group design follow the common practice of reporting mean comparisons of posttreatment scores without controlling for pretreatment status, effect sizes based on these values are biased under conditions of nonrandom assignment. Because standardized mean-change scores make use of pretreatment as well as posttreatment data, they are not biased by pretreatment differences between groups. A third benefit of Becker's (1988) approach is that it yields an estimate of the average improvement in the condition under study when no treatment is available. This is particularly useful in the study of treatments for conditions (such as grief) that are expected to improve over time.

Exceptions

Reviewers can compute standardized mean-change scores when study authors report either (a) pretreatment and posttreatment means and standard deviations separately for each (treatment or control) group or (b) within-group (pre–post) t tests. Of the 35 independent treatment samples retrieved in this review, all one-group (pre–post) studies provided the necessary information, but 10 two-group studies did not provide the information necessary to compute standardized mean-change scores. Effect sizes for these 10 studies were based on comparisons of posttreatment scores for treatment and control groups. Two additional two-group studies provided alternative means of taking pretreatment scores into account, in the form of an analysis of covariance (ANCOVA) comparing posttreatment means after controlling for pretreatment scores. Because ANCOVA is the optimal approach to analysis of the two-group, pre–post design (Huck & McClean, 1975; cf. Cohen & Cohen, 1983, pp. 72–73), we computed effect sizes for these two studies directly from the ANCOVA results rather than using the Becker conversion.

In an exploratory analysis to examine possible differences in effect sizes based on the method of effect size computation, we compared the 10 studies that did not provide information to control for pretreatment scores with the 25 studies that did provide the necessary data. The weighted mean effect size for the former group ($d_+ = 0.37$, 95% confidence interval [CI] = 0.23 to 0.51) was somewhat smaller than that for studies controlling for pretreatment status ($\Delta_+ = 0.46$, 95% CI = 0.35 to 0.58). However, this difference was not statistically significant: $Q(1) = 0.99$, $p = .33$, indicating that any systematic differences in effect sizes attributable to different methods of computation were negligible relative to other sources of heterogeneity in these studies. Thus, we elected to retain the 10 studies that did not control for pretreatment status in all analyses. Although effect sizes from these 10 studies are technically estimates of Cohen's δ rather than Becker's Δ, we refer to weighted aggregate effect sizes derived from treatment-control comparisons (which are based primarily on estimates of Δ) as Δ_+, whereas weighted aggregate effect sizes derived from within-group (pre–post) comparisons are designated as d_+.

Multiple Effect Sizes From Individual Studies

Studies could contribute multiple effect sizes because of comparing multiple levels of a treatment against a common control or using multiple dependent measures. When multiple effect sizes occurred in the same study, we combined them before aggregating Δ values across studies. Because information was not available on the magnitude of dependency (e.g. correlations between dependent measures; Gleser & Olkin, 1994) between pairs of effect sizes, we weighted all effect sizes from the same study equally for purposes of aggregation.

We combined multiple effect sizes using a shifting unit of effect size (Cooper, 1989). For the omnibus analysis, we aggregated

multiple effect sizes from the same study, so that each study contributed one effect size to the overall estimate. However, when examining potential moderators, a study's results were aggregated only within levels of the moderator variable. For example, a study reporting separate effect sizes for observer and self-report ratings would contribute one effect size to the omnibus analysis but two effect sizes, one for each category (observer and self-report), to the examination of source of outcome measure as a moderator of therapy effectiveness. This shifting unit of effect size is recommended because it takes advantage of all information provided in each study but avoids problems with the inclusion of multiple dependent effect sizes in aggregate estimates (Cooper, 1989).

Results

Our comprehensive review of grief therapy outcome literature derived 274 effect sizes from 35 studies meeting inclusion criteria (a table listing the 35 studies with corresponding omnibus effect sizes is available from Denise Litterer Allumbaugh on request). Analyses of effect sizes for outliers did not reveal any extreme outliers (i.e., effect sizes 3 or more standard deviation units from the aggregate effect size), so we included all effect sizes in aggregation. The mean number of effect sizes per study was 7.8, with a range of 1 to 29 effect sizes per study, and the mean number of clients per study was 65 (N = 2,284). The mean percentage of female participants was 84%. The modal client was a woman age 52 who had lost a spouse 27 months before she started therapy and who participated in eight weekly treatment sessions.

Omnibus Analysis

The omnibus analysis examined the overall effect of grief therapy. The weighted aggregate effect size was $\Delta_+ = 0.43$ (95% CI = 0.33 to 0.52), with k = 35 effect sizes included. This effect size is moderate by Cohen's (1988) standards for research in the behavioral sciences, but it is small relative to the effect sizes derived from meta-analyses of other forms of psychotherapy (Lambert & Bergin, 1994). Although some improvement is expected among untreated bereaved individuals, very little improvement was observed among no-treatment control groups in these studies (weighted aggregate standardized mean-change score for the control group was $d_+ = 0.06$, 95% CI = −0.43 to 0.54). Observed variance among effect sizes was greater than that expected due to sampling error alone, $Q(34) = 118.31$, $p < .0001$, suggesting the presence of moderator effects.

Moderator Analyses

Categorical Moderators

Results of analyses of all categorical moderators are reported in Table 1. These analyses were conducted according to procedures recommended by Hedges & Olkin (1985). We used the program DSTAT (Johnson, 1993) to conduct categorical analyses, but we need to adjust Ns so that DSTAT would compute the correct variances (and therefore the correct aggregation weights) for studies whose effect sizes we estimated using Becker's (1988) procedures. A significant between-groups Q statistic indicates that grouping studies by levels of the moderator variable results in a significant reduction in heterogeneity among effect sizes—that is, the moderator explains at least some of the heterogeneity observed in the omnibus analyses. A significant within-groups Q statistic indicates that heterogeneity greater than expected by chance remains within studies at the designated level of the moderator variable (suggesting that studies within this group did not derive from a homogeneous population, and further moderators may yet be found). For significant moderators with more than two levels, we conducted post hoc contrasts between pairs of levels using $\alpha = .01$ to aid in the interpretation of the significant effect.

Table 1
Univariate Moderator Analyses for Categorical Variables

Variable	Δ_+	95% CI	Q	df
Practitioner training				
Within group				
Professional	0.66_a	0.49, 0.83	56.39***	12
Professional in training	$0.30_{a,b}$	−0.02, 0.61	8.62	8
Nonprofessional	0.11_b	−0.07, 0.30	5.61	4
Between group			18.66***	2
Modality of intervention				
Within group				
Group	0.40	0.30, 0.49	90.59***	24
Individual	0.65	0.38, 0.91	20.43**	7
Between group			3.09	1
Risk of complicated grief				
Within group				
High risk	0.38	0.06, 0.70	9.00	5
Normal	0.43	0.33, 0.52	109.24***	28
Between group			0.08	1
Type of control group				
Within groups				
Placebo treatment	0.74	0.27, 1.21	18.88***	5
No treatment	0.37	0.21, 0.42	46.16***	19
Between groups			2.98	1
Method of group assignment				
Within groups				
Random	0.40	0.27, 0.53	25.02*	13
Systematic	0.17	−0.03, 0.36	22.47**	8
Between groups			3.91*	1
Source of grief measure				
Within groups				
Observer	1.24	0.83, 1.64	12.53*	4
Self-report	0.43	0.33, 0.52	110.65***	33
Between groups			14.52***	1
Publication status				
Within groups				
Published	0.39	0.29, 0.49	79.86***	22
Unpublished	0.73	0.44, 1.02	33.64***	11
Between groups			4.82*	1

Note. Column entries with different subscripts differ at $p < .001$; Δ_+ = aggregate estimate of the difference in improvement between treatment and control groups; CI = confidence interval; Q = heterogeneity test statistic.
*$p < .05$. **$p < .01$. ***$p < .001$.

Treatment characteristics. Practitioner training was a significant moderator of treatment outcome, with licensed professionals (psychologists, social workers, and psychiatric nurses) producing significantly more favorable outcomes ($\Delta_+ = 0.66$) than nonprofessional therapists ($\Delta_+ = 0.11$) and marginally more favorable outcomes than professionals

in training ($\Delta_+ = .30$). Modality of intervention was marginally significant, $Q_B(1) = 3.09, p < .10$, indicating a trend for group treatments ($\Delta_+ = 0.40$) to be less effective than individual treatments ($\Delta_+ = 0.65$). This apparent difference in effectiveness between group and individual treatment modalities runs counter to the finding of comparability in the general psychotherapy outcome literature (Bednar & Kaul, 1994), and is likely due to the confounding of treatment modality and practitioner training: Of the studies that could be coded for both practitioner training and modality of treatment, three of the four studies using nonprofessional therapists involved group treatments, as did all nine of the studies using professionals in training.

Client characteristic. Contrary to expectation, level of client risk for complicated bereavement (as reported by study authors) was not related to outcome: Low-risk clients ($\Delta_+ = 0.43$) derived roughly equivalent benefit from grief therapy as did high-risk clients ($\Delta_+ = 0.38$). When we eliminated the nonsignificant results that had been set to zero because they did not provide sufficient information to calculate effect sizes, results were still nonsignificant but in the predicted direction, with high-risk clients ($\Delta_+ = 0.57$) gaining marginally higher benefit from grief therapy than low-risk clients ($\Delta_+ = 0.46$).

Although definitions of high-risk clients (or "complicated bereavement") differed from study to study, we coded clients as high-risk for the purpose of this analysis if the original authors characterized them as such. For example, some studies included only clients with *morbid grief,* defined as individuals with distress persisting longer than 1 year (Mawson, Marks, Ramm, & Stern, 1981; Sireling, Cohen, & Marks, 1988). Other authors operationalized high distress as a score of 5 or higher on the Goldberg General Health Questionnaire (Schut, Stroebe, Van den Bout, & de Keijser, 1997; Vachon, Lyall, Rogers, Freedman-Letofsky, & Freeman, 1980). Thus, the lack of significance for this moderator test may reflect variability among study authors in their risk criteria. We present additional information on the impact of clients' genuine distress on treatment outcome in the discussion of the moderator effect of clients' relationship to the deceased, below.

Study characteristics. Oddly, effect sizes based on comparisons with placebo treatments did not differ significantly from those based on comparisons with no-treatment control groups. In fact, the effect size based on placebo comparisons was marginally higher, $Q_B(1) = 2.98, p < .10$, than that for control comparisons. When we eliminated effect sizes conservatively set to zero from the data set, the difference was even more striking, $Q_B(1) = 10.44, p < .01$.

Further examination of the five studies that compared intervention with placebo treatment indicated that two studies were strongly influential in determining the large aggregate effect size for this category. The first, Mawson et al. (1981), with an overall effect size of $\Delta_+ = 3.43$ (excluding effect sizes set to zero), was designed so that a treatment group in which "guided mourning" exposed the bereaved to memories of the deceased was compared with a placebo group in which participants were told to actively avoid thoughts of the deceased through techniques such as thought stopping. These opposing conditions may have produced an artificially high effect size. The second study with an unusually high effect size for treatment–placebo comparison ($\Delta_+ = 2.93$) was Constantino (1981). However, when the treatment group in this study was compared with a nontreatment control group, the effect size was even higher ($\Delta_+ = 3.34$), thus actually supporting the expectation that a treatment–nontreatment comparison will produce a higher effect size than a treatment–placebo comparison. If the effect sizes from these two studies are eliminated, there are no significant differences between effect sizes based on a nontreatment control group and those based on a placebo control group, $Q_B(1) = 1.46, p = .22$.

Studies using random assignment ($\Delta_+ = 0.40$) yielded higher effect sizes than those using systematic assignment ($\Delta_+ = 0.17$) of clients to groups. At least some of the studies in this sample evidently used systematic assignment in a way that negatively impacted their estimates of treatment effectiveness. For example, Cordsen (1987) and Lieberman and Videka-Sherman (1986) both included control groups made up of clients who did not want to participate in treatment—and may therefore have been less distressed and more likely to improve even without treatment.

Most outcome measures used in the studies we reviewed were based on client self-reports, with a minority based on the judgments of trained observers (clinicians other than the therapist). Observer ratings ($\Delta_+ = 1.24$) were apparently much more sensitive to change than were client self-reports ($\Delta_+ = 0.43$). However, this apparent difference in sensitivity disappeared when we made within-study comparisons: The weighted mean effect size for self-report measures for the four studies that included both self-report and observer ratings was $\Delta_+ = 0.92$, and did not differ significantly from the weighted mean effect size for observer ratings.

Finally, the analysis of publication status as a moderator of treatment effectiveness showed that the conventional wisdom (e.g., Rosenthal, 1979) that unpublished studies will have smaller effect sizes than published studies does not always hold true. In fact, the weighted aggregate effect size for unpublished studies in this sample ($\Delta_+ = 0.73$) was significantly higher than that for published studies ($\Delta_+ = 0.39$). Results from similar analyses in which effect sizes set equal to zero were eliminated failed to achieve significance, but still showed an effect size for unpublished studies ($\Delta_+ = 0.71$) that was substantially higher than that for published studies ($\Delta_+ = 0.45$).

Continuous Moderators

We evaluated continuous moderators using weighted regression. To test for linear trends, we regressed effect sizes onto values of the moderator variable, with values weighted by the inverse of the variance of the corresponding effect size (Hedges & Olkin, 1985). Quadratic trends were also examined for all continuous moderators by entering the square of the predictor variable at a second step in the weighted regression (Cohen & Cohen, 1983), and we present these trends when results were significant.

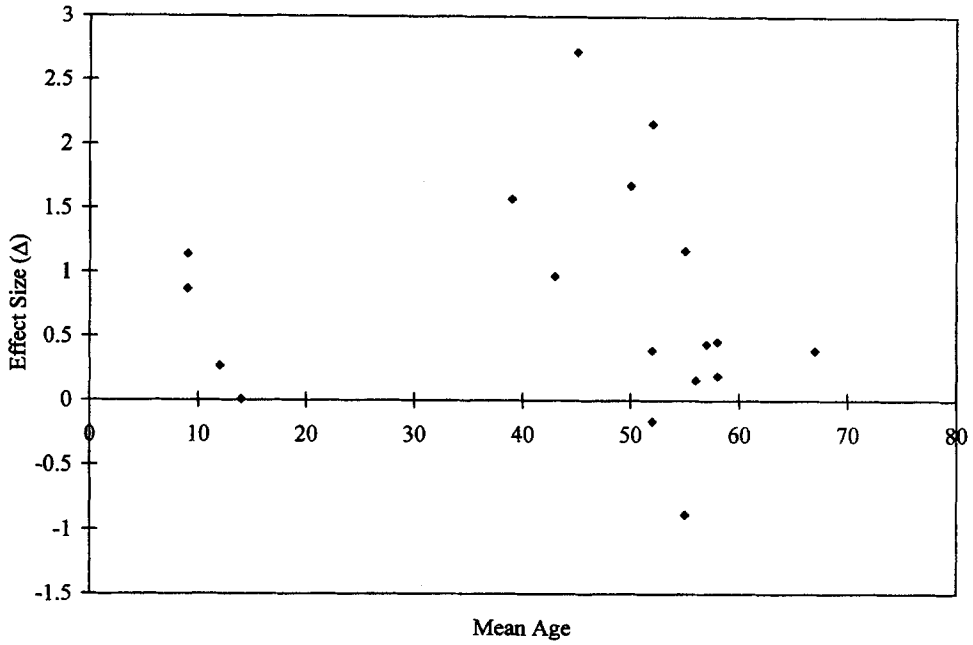

Figure 1. Scatter plot of mean age of clients versus effect size.

Treatment characteristic. Number of treatment sessions was found to be linearly related to effect size ($B = 0.07$, $p < .01$). On average, each additional session of therapy increased the expected effect size by approximately 0.07 units. Examination of the scatterplot indicated that one bivariate outlier had a strong influence on this estimate, however: Clients in one of the treatment conditions in Schut, de Keijser, Van den Bout, and Stroebe's (1996) study underwent 20 sessions of therapy (the longest treatment in the remaining studies was 14 sessions), with a large effect size of $\Delta = 2.20$. When we excluded this study from the regression, the linear relation was no longer significant ($B = 0.03, p > .10$). Quadratic components in both analyses were not significant ($\Delta R^2 = .06$ and .002, respectively; both $ps > .05$). Thus, although we found equivocal evidence of a linear relation between treatment length and effectiveness, there was no evidence of the curvilinear dose-response relation observed for psychotherapy for other problems (Howard, Kopta, Krause, & Orlinsky, 1986), probably because of the relative brevity of treatments included in these studies.

Client characteristics. Another continuous moderator that achieved significance was mean client age, which was related to effect size both linearly ($B = -0.014$, $p < .01$) and quadratically ($\Delta R^2 = .25, p < .05$). The scatter plot (see Figure 1) shows that clients of intermediate age experienced better outcomes, on average, in these studies than either younger or older clients. A related finding that may shed light on the importance of client age is Steele's (1992) study of 60 bereaved spouses. Steele found that the young (ages 25–35) and the elderly (ages 66–85) experienced more despair and hopelessness than the middle-aged group of clients. This finding suggests that older and younger bereaved individuals may be more difficult to treat because their grief reactions are more complex or intense than the middle-aged bereaved individuals' reactions. However, as most of the studies in this meta-analysis used clients between the ages of 50 and 60 and produced a wide range of effect sizes, the results are somewhat difficult to interpret. More studies examining clients outside of this narrow age range are needed to clarify the relationship between age of clients and effectiveness of grief therapy.

Gender of client (percent female) did not predict treatment effectiveness ($B = 0.001$, $p = .80$). Thus, there is no evidence that men and women benefit differentially from grief treatment.

Average time (in months) since loss was also a significant predictor of effect size ($B = -0.01, p < .01$), indicating that interventions beginning closer to the time of the loss were more effective than those interventions that were delayed. In quantitative terms, each month that intervenes between the loss and the onset of treatment decreases the expected standardized mean-change score (i.e., standardized improvement in treatment group relative to control group) by .011. Delaying the onset of treatment by a year, for example, decreases the expected efficacy of treatment by .13 *SD* units. Thus, the fact that the mean time since loss for bereaved clients in these studies was more than 2 years ($M = 27$ months) is one major problem in extrapolating from our findings to infer the effectiveness of grief interventions as typically implemented in the real world, a point to which we return in the Discussion section.

The final continuous moderator we examined was clients' relationship to the deceased. Studies in our sample differed in their heterogeneity with respect to this variable. A majority of these studies ($k = 22$) included spouses only,

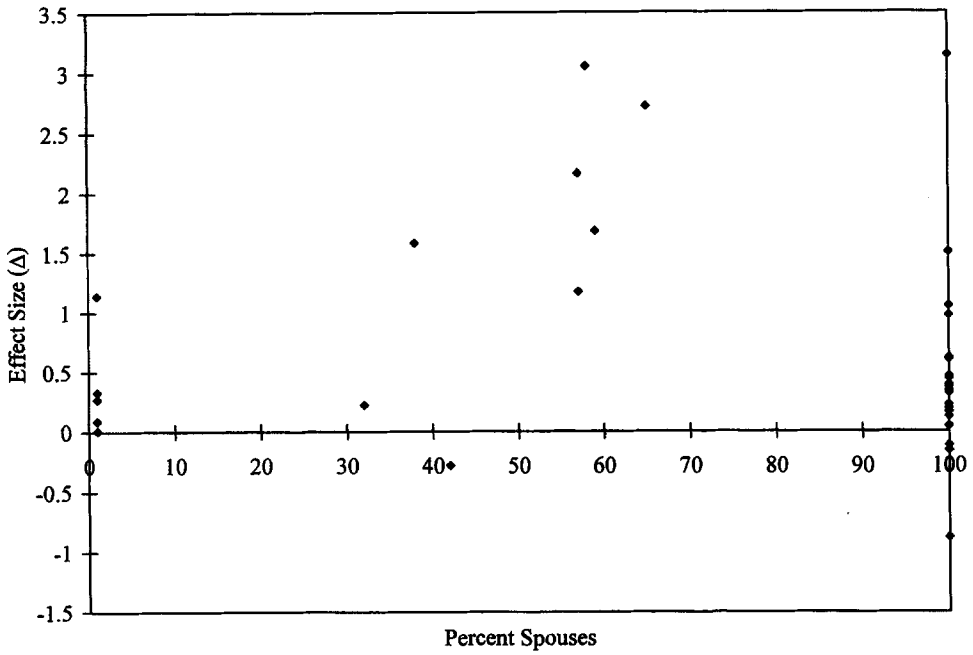

Figure 2. Scatter plot of percentage of spouses versus effect size.

and some of these examined treatments specifically tailored to the needs of widows and widowers. A small number of studies ($k = 5$) targeted bereaved children who had lost a parent or caregiver, sibling, or grandparent (Quarmby, 1993; Ryan, 1982; Schilling, Koh, Abramovitz, & Gilbert, 1992; Tonkins & Lambert, 1996; Wilson, 1994). Finally, a larger number of studies ($k = 8$) were heterogeneous on this variable, including clients who had lost spouses, children, parents, other family members, and sometimes close friends. Because of the special nature of spousal bereavement (Fulton & Owen, 1977; Stroebe, Stroebe, Abakoumin, & Schut, 1996) and because few studies included homogeneous samples of nonspouse clients, we elected to compare outcomes of spouses and nonspouses by computing the percentage of spouses in each sample.

Weighted regression analyses revealed a significant negative linear relation ($B = -0.06$, $p < .001$) between percentage of spouses and effect size, with effect size increasing as percentage of spouses decreased, suggesting that there may be something uniquely unresponsive about spousal bereavement. However, the quadratic relation, shown in the scatterplot (see Figure 2), was also significant ($\Delta R^2 = .33$, $p < .001$), indicating that treatments using a heterogeneous client sample were more effective than treatments targeting only bereaved spouses or only nonspouses.

Post hoc analyses. The finding that effect sizes were highest for those studies examining heterogeneous (i.e., mixed spouse and nonspouse) samples is puzzling. For example, Budman, Simeone, Reilly, and Demby (1994) discussed the importance of homogeneity and focus to the success of short-term group treatments. Even in studies examining individual treatments for bereavement, it is not clear why mixed samples should have better outcomes than homogeneous samples. In an attempt to detect confounding variables that may explain this finding, we examined each of the mixed-sample studies to see whether these differed systematically from studies using homogeneous (i.e., spouse or nonspouse) samples.

Of the eight studies that used heterogeneous samples, two studies (Cordsen, 1987; Sireling, Cohen, & Marks, 1988) produced small effect sizes (less than $\Delta = .25$), whereas the remaining six studies produced very large effect sizes (ranging from $\Delta = 1.17$ to $\Delta = 3.05$). Of the latter six studies, two studies (Schut et al. 1996, A and B) used inpatient samples, and a third (Horowitz, Marmar, Weiss, DeWitt, & Rosenbaum, 1984) examined an outpatient sample that voluntarily sought treatment for bereavement at a university research center. Two other study samples (Moran, 1988; Polinskey, 1992) consisted of clients who requested services from existing bereavement programs. The sixth study (Mawson et al., 1981) did not provide specific information about the recruitment of clients; however, the clients were described as suffering from "morbid grief," experiencing "persistent [bereavement-related] distress" of over 1 year's duration as well as two or more indications of risk for morbid grief (e.g. delayed onset of grief, increased drug use, excessive guilt). Thus, the six mixed-sample studies that produced very large effect sizes used participants who were likely similar to real-world clients because they voluntarily seek treatment, were at-risk for complicated bereavement, or both.

In contrast, most of the studies that examined either all spouses or no spouses used clients who were less similar to typical clients. The majority of these studies ($n = 22$) actively recruited participants for their study through obituaries, lists of individuals who had recently died in a hospital

or hospice, referrals from clergy, funeral directors, community directors, or advertisements for bereavement programs designed specifically for the study. Only one study (Sabatini, 1988; $\Delta = 0.44$) resembled the mixed-sample studies in that it used advertisements for an existing bereavement program, thus likely including primarily voluntary participants. (Recruitment information was unclear or unavailable for an additional five studies.) Thus, a possible explanation for the apparent superiority of the mixed-sample studies in this meta-analysis is that these studies included clients who were genuinely distressed and sought out treatment. By contrast, the vast majority of homogeneous samples in these studies appear to consist of recruited clients, perhaps resulting in client samples that were less distressed or less motivated to work in therapy. We return to the question of client recruitment procedures in our discussion of external validity of grief treatment outcome studies, below.

Discussion

The omnibus effect size of $\Delta_+ = 0.43$ produced by this meta-analysis, although moderate by Cohen's (1988) standards, is small relative to the 0.80 effect size of psychotherapy for a variety of problems found by previous meta-analyses (Lambert & Bergin, 1994). As the intensity of grief is generally expected to decline over time, one possible explanation for this finding is that clients in control groups improved significantly without treatment. However, this clearly does not account for the difference between the grief interventions studied here and other psychosocial treatments, because bereaved clients assigned to no-treatment control groups in these studies did not actually improve all that much ($d_+ = 0.06$). The results of the moderator analyses included in this meta-analysis suggest alternative explanations for the apparent difference in effectiveness between grief and other therapies.

Client Recruitment

Perhaps the most striking results of the moderator analyses was that grief interventions in studies examining groups of participants who varied in their relationship to the deceased produced much higher effect sizes than studies with participant pools that consisted of either all spouses or no spouses. Post hoc analyses revealed a potential confounding variable of type of recruitment method, indicating that studies with heterogeneous client samples were often those that involved self-selected clients rather than persons recruited to participate in a treatment study. This observation is important because studies that attempt to control extraneous variables by recruiting and selecting homogeneous groups of clients for laboratory grief programs (similar to what Seligman, 1995, called *efficacy* studies) do not accurately reflect the real-world therapy in the same way that studies of voluntary participants in existing programs (Seligman's *effectiveness* studies) do. Furthermore, the handful of studies involving self-selected clients produced robust effect sizes, suggesting that individuals who seek treatment for bereavement benefit much more than the small omnibus effect size produced by this meta-analysis indicates.

Why, then, do bereaved individuals who voluntarily seek therapy appear to benefit more than those who are recruited? Potential contributing factors include higher initial distress levels (and corresponding greater potential for improvement), higher levels of extratherapeutic resources, and greater motivation for therapy among self-selected clients. In their examination of 339 bereaved older adults who had experienced the loss of a spouse within the previous 3 months, Caserta and Lund (1992) found that those who sought professional help reported higher depression, lower coping ability, and poorer perceived health than those who did not seek help. Similarly, Rynearson (1995) found higher grief, trauma, and intrusive reenactment imagery among 52 adult family members of homicide victims who sought treatment within the first year of their loss, relative to family members who did not seek treatment. Thus, bereaved individuals who voluntarily seek treatment appear to have higher levels of distress than bereaved individuals who do not, suggesting that the former group is more in need of, and perhaps more likely to benefit from, psychotherapy.

Not only do individuals who actively seek treatment seem to be experiencing higher levels of distress than those who do not, there are additional positive factors that may aid them in their therapeutic progress, or at least reduce the incidence of further distress. Such positive factors include a more active support network (Reif, Patton, & Gold, 1995) and possibly higher education and income (Vessey & Howard, 1993). Thus, people who actively seek professional help may be those who, because of a variety of personality and demographic variables, have the most potential to benefit from it. Finally, it seems reasonable to believe that individuals seeking therapy on their own initiative are more highly motivated than those agreeing to treatment as a result of experimenter recruiting efforts. Although empirical findings are mixed, client motivation is a frequently cited predictor of both client continuation in therapy and of therapeutic outcome (Garfield, 1994).

Optimal Timing of Grief Interventions

The mean length of time since loss for clients beginning therapy in the studies included in this meta-analysis was over 2 years, an unusually long time relative to what might be thought optimal for grief treatment. Our findings suggest that the lag between loss and onset of therapy in these studies may well have contributed to a relatively low aggregate effect size. It is probable that clients in both the treatment and control groups had experienced a certain amount of normal recovery before starting treatment, leaving a smaller scope for improvement. Indeed, the weighted regression analyses reported above predict a substantially larger omnibus effect size (on the order of $\Delta_+ = 0.70$) for studies involving clients who are relatively recently bereaved (i.e., those who begin therapy within 3 or 4 months of their loss, or approximately 24 months earlier than the average client in our sample).

The relatively late onset of treatment in these studies also has implications for our estimate of improvement among untreated clients. Although there is no universal timetable for grief reactions, it seems reasonable to believe that grief recovery in most cases is relatively rapid within 3 to 6 months of the loss and tapers off over time. Thus, the relatively small spontaneous recovery rate in control groups in these samples is not unexpected, given that the modal study examined improvement over an 8-week period more than 2 years from the time of the loss. Thus, the mean control group improvement of .06 *SD*s over the course of treatment does not likely reflect the range of normal expected improvement among recently bereaved clients over a similar interval.

Conclusions

The moderate omnibus effect size for the studies reviewed may say more about the nature of the studies than about the effectiveness of grief treatment per se. Indeed, our analyses suggest that grief interventions with self-selected clients that begin within a few months of the loss are likely to be as effective or possibly even more effective than psychotherapy in general.

Limitations of Findings

It is important to note that confounds with client-recruitment procedures also affect the results of other moderator analyses. The apparent superiority of professional therapists over trainees and paraprofessionals is attributable to the following: All four of the studies that used self-selected clients that could be coded on this variable fell into the professional therapist category. When these studies were excluded from the moderator analyses shown in Table 1, the apparent advantage for professionals disappeared.

In addition, the curvilinear relation between age and treatment effectiveness disappeared when we excluded studies involving self-selected clients. As one might expect from the fact that these studies were heterogeneous with respect to clients' relationships to the deceased, the studies involving self-selected clients also tended to have mean client ages in the intermediate range, resulting in the appearance of enhanced therapeutic effectiveness for clients in this intermediate age range. When we excluded the studies using self-selected clients from this analysis, both linear and quadratic trends relating age to outcome were nonsignificant.

Recommendations for Future Research

The relevance of research to practice is dependent on boundary conditions that limit the generalizability of findings. Kazdin (1992) noted that these boundary conditions can be conceptualized in terms of statistical interactions or, in the terminology of meta-analysis, *moderator variables* that govern the effectiveness of treatment. When treatment or client characteristics in research investigations do not match those characteristics typical of therapy settings, the generalizability of research findings is questionable (Seligman, 1995). Our findings suggest that outcome studies of grief interventions have largely involved groups of clients who are dissimilar to those who seek treatment for grief in the real world, in that clients in research contexts have often been (a) recruited by researchers and (b) grieving over events that occurred more than 2 years earlier. Furthermore, our findings suggest that each of these variables is an important moderator of outcome and, therefore, that the relevance of most existing studies to practice is highly suspect. Thus, future researchers should make every effort to include self-selected clients rather than recruiting participants and should attend to the recency of the loss in deciding what groups to target for grief interventions.

An important question that has not been resolved by the present review concerns the benefits of grief interventions for high-risk individuals. Given that therapeutic resources are scarce, death is normal and inevitable, and many if not most bereaved individuals return to normal functioning in time, it behooves therapists to focus their attention on those at risk for extended or complicated grief reactions (Parkes & Weiss, 1983; Raphael, 1977). A problem we encountered in trying to evaluate the special benefits of grief interventions for at-risk populations was the variety of criteria original study authors used to identify at-risk individuals, and existing studies (using divergent methods to identify high-risk clients) provide no support for the theory that high-risk grievers benefit more from therapy than normally bereaved individuals. More research is needed on the characteristics of individuals and events that increase the probability of delayed or complicated course of grief. Comparison of existing studies suggests that one important predictor of treatment success may be motivation for treatment.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Barbato, A., & Irwin, H. J. (1992). Major therapeutic systems and the bereaved client. *Australian Psychologist, 27,* 22–27.

*Barrett, C. J. (1978). Effectiveness of widows' groups in facilitating change. *Journal of Consulting and Clinical Psychology, 46,* 20–31.

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

Bednar, R. L., & Kaul, T. J. (1994). Experiential group research: Can the canon fire? In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed, pp. 631–663). New York: Wiley.

Budman, S. H., Simeone, P. G., Reilly, R., & Demby, A. (1994). Progress in short-term and time-limited group psychotherapy: Evidence and implications. In A. Fuhriman & G. M. Burlingame (Eds.), *Handbook of group psychotherapy: An empirical and clinical synthesis* (pp. 319–339). New York: Wiley.

Caserta, M. S., & Lund, D. A. (1992). Bereaved older adults who seek early professional help. *Death Studies, 16,* 17–39.

*Caserta, M. S., & Lund, D. A. (1993). Intrapersonal resources and the effectiveness of self-help groups for bereaved older adults. *The Gerontologist, 33,* 619–629.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46.

Cohen, J. (1962). The statistical power of abnormal-social psychological research: A review. *Journal of Abnormal and Social Psychology, 65,* 145–153.

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

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

*Constantino, R. E. (1981). Bereavement crisis intervention for widows in grief and mourning. *Nursing Research, 30,* 351–353.

*Constantino, R. E., & Bricker, P. L. (1996). Nursing postvention for spousal survivors of suicide. *Issues in Mental Health Nursing, 17,* 131–152.

Cooper, H. M. (1989). *Integrating research: A guide for literature reviews* (2nd ed.). Newbury Park, CA: Sage.

*Cordsen, M. L. (1987). Changes in grief over time with a brief psychoanalytic group intervention (Doctoral dissertation, California State University, Long Beach, 1988). *Dissertation Abstracts International, 26,* 61.

*Duran, Alicia C. (1988). A grief intervention for adults experiencing conjugal bereavement (Doctoral dissertation, University of Utah, 1987). *Dissertation Abstracts International, 48,* 3677.

*Flatt, B. (1988). Some effects of grief therapy groups. *Journal of Psychology and Christianity, 7,* 43–49.

Freud, S. (1957). Mourning and melancholia. In J. Strachey (Ed. & Trans.), *Standard edition of the complete psychological works of Sigmund Freud.* (Vol. XIV, pp. 239–260). London: Hogarth Press. (Original work published 1917).

Fulton, R., & Owen, G. (1977). *Adjustment to loss through death: A sociological analysis.* Minneapolis, MN: University of Minneapolis, Center for Death Education and Research.

Garfield, S. L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* 4th ed., (pp. 190–228). New York: Wiley.

*Gerber, I., Weiner, A., Battin, D., & Arkin, A. (1975). Brief therapy to the aged bereaved. In B. Schoenberg & I. Gerber (Eds.), *Bereavement: Its psychosocial aspects* (pp. 310–333). New York: Columbia University Press.

Gleser, L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In J. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 339–355). New York: Russell Sage Foundation.

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

*Horowitz, M. J., Marmar, C., Weiss, D. S., DeWitt, K. N., & Rosenbaum, R. (1984). Brief psychotherapy of bereavement reactions: The relationship of process to outcome. *Archives of General Psychiatry, 41,* 438–448.

Horowitz, M. J., Weiss, D. S., Kaltrider, N., Krupnick, J., Marmar, C., Wilner, N., & DeWitt, K. (1984). Reactions to the death of a parent: Results from patients and field subjects. *Journal of Nervous and Mental Disease, 172,* 383–392.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41,* 159–164.

Huck, S. W., & McClean, R. A. (1975). Using a repeated measures ANOVA to analyze the data from a pretest–posttest design: A potentially confusing task. *Psychological Bulletin, 82,* 511–518.

Johnson, B. T. (1993). DSTAT (Version 1.10) [Computer software]. Hillsdale, NJ: Erlbaum.

*Jones, D. R. (1979). The Grief Therapy Project: The effects of group therapy with bereaved surviving spouses on successful coping with grief (Doctoral dissertation, University of Southern California, 1979). *Dissertation Abstracts International, 39,* 6121.

Kazdin, A. E. (1992). *Research design in clinical psychology* (2nd ed.). New York: Macmillan.

Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed, pp. 143–189). New York: Wiley.

*Lieberman, M. A., & Videka-Sherman, L. (1986). The impact of self-help groups on the mental health of widows and widowers. *American Journal of Orthopsychiatry, 56,* 435–449.

Lieberman, M. A., & Yalom, I. (1992). Brief group psychotherapy for the spousally bereaved: A controlled study. *International Journal of Group Psychotherapy, 42,* 117–132.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist, 48,* 1181–1209.

*Marmar, C. R., Horowitz, M. J., Weiss, D. S., Wilner, N. R., & Kaltrieder, N. B. (1988). A controlled trial of brief psychotherapy and mutual-help group treatment of conjugal bereavement. *American Journal of Psychiatry, 145,* 203–209.

*Mawson, D., Marks, I. M., Ramm, L., & Stern, R. S. (1981). Guided mourning for morbid grief: A controlled study. *British Journal of Psychiatry, 138,* 185–193.

*Moran, M. R. (1988). A directed grieving paradigm for facilitating the mourning process and preventing abnormal grief response (Doctoral dissertation, University of South Dakota, 1988). *Dissertation Abstracts International, 49,* 2416.

Parkes, C. M., & Weiss, R. S. (1983). *Recovery from bereavement.* New York: Basic Books.

*Polinskey, T. G. (1992). The effects of a group counseling grief recovery program for clients experiencing bereavement (Doctoral dissertation, University of Arkansas, Fayetteville, 1990). *Dissertation Abstracts International, 52,* 3830.

Potocky, M. (1993). Effective services for bereaved spouses: A content analysis of the empirical literature. *Health & Social Work, 18,* 288–301.

*Quarmby, D. (1993). Peer group counselling with bereaved adolescents. *British Journal of Guidance and Counselling, 21,* 196–211.

Raphael, B. (1977). Preventive intervention with the recently bereaved. *Archives of General Psychiatry, 34,* 1450–1454.

*Reich, J. W., & Zautra, A. J. (1989). A perceived control intervention for at-risk older adults. *Psychology and Aging, 4,* 415–424.

Reif, L. V., Patton, M. J., & Gold, P. B. (1995). Bereavement, stress, and social support in members of a self-help group. *Journal of Community Psychology, 23,* 292–306.

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

*Rowland, W. A. (1987). The effects of a cognitive affective peer support program on mood states in conjugal grief (Doctoral dissertation, United States International University, 1987). *Dissertation Abstracts International, 48,* 346.

*Ryan, S. M. (1982). Childhood bereavement: Psychological test findings of a post-death intervention program (Doctoral dissertation, University of Arizona, Tucson, 1982). *Dissertation Abstracts International, 43,* 885.

Rynearson, E. K. (1995). Bereavement after homicide: A comparison of treatment seekers and refusers. *British Journal of Psychiatry, 166,* 507–510.

*Sabatini, L. (1988). Evaluating a treatment program for newly widowed people. *Omega, 19,* 229–236.

Sanders, C. M. (1989). *Grief: The mourning after—dealing with adult bereavement.* New York: Wiley.

*Schilling, R. F., Koh, N., Abramovitz, R., & Gilbert, L. (1992). Bereavement groups for inner-city children. *Research on Social Work Practice, 2,* 405–419.

Schmidt, F. L. (1992). What do data really mean? Research findings, meta-analysis, and cumulative knowledge in psychology. *American Psychologist, 47,* 1173–1181.

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

*Schut, H. A. W., de Keijser, J., Van den Bout, J., & Stroebe, M. S. (1996). Cross-modality grief therapy: Description and assessment of a new program. *Journal of Clinical Psychology, 52,* 357–365.

*Schut, H. A. W., Stroebe, M. S., Van den Bout, J., & de Keijser, J. (1997). Intervention for the bereaved: Gender differences in the efficacy of two counselling programs. *British Journal of Clinical Psychology 36 (1),* 63–72.

Schwartzberg, S. S., & Halgin, R. P. (1991). Treating grieving clients: The importance of cognitive change. *Professional Psychology: Research and Practice, 22,* 240–246.

Seligman, M. E. P. (1995). The effectiveness of psychotherapy: The *Consumer Reports* study. *American Psychologist, 50,* 965–974.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

*Sireling, L., Cohen, D., & Marks, I. (1988). Guided mourning for morbid grief: A controlled replication. *Behavior Therapy, 19,* 121–132.

Sobel, H. J. (Ed). (1981). *Behavior therapy in terminal care: A humanistic approach.* Cambridge, MA: Ballinger.

Steele, L. (1992). Risk factor profile for bereaved spouses. *Death Studies, 16,* 387–399.

Stroebe, W., Stroebe, M., Abakoumkin, G., & Schut, H. (1996). The role of loneliness and social support in adjustment to loss: A test of attachment versus stress theory. *Journal of Personality and Social Psychology, 70,* 1241–1249.

*Tonkins, S. A. M., & Lambert, M. J. (1996). A treatment outcome study of bereavement groups for children. *Child and Adolescent Social Work Journal, 13,* 3–21.

U.S. Bureau of the Census. (1995). *Statistical abstract of the United States: 1995* (115th ed). Washington, DC: Author.

*Vachon, M. L. S., Lyall, W. A. L., Rogers, J., Freedman-Letofsky, K., & Freeman, S. J. J. (1980). A controlled study of self-help intervention for widows. *American Journal of Psychiatry, 137,* 1380–1384.

Vessey, J. T., & Howard, K. I. (1993). Who seeks psychotherapy? *Psychotherapy, 30,* 546–553.

*Walls, N., & Meyers, A. W. (1985). Outcome in group treatments for bereavement: Experimental results and recommendations for clinical practice. *International Journal of Mental Health, 13,* 126–147.

*Wilson, D. L. (1994). A outcome study of a time-limited group intervention program for bereaved children (Doctoral dissertation, Washington State University, 1994). *Dissertation Abstracts International, 55,* 3793.

*Wright, P. J. (1988). *Grief recovery: A comparison of group therapies for widows.* Unpublished doctoral dissertation, Memphis State University, Memphis.

Yalom, I. D., & Vinogradov, S. (1988). Bereavement groups: Techniques and themes. *International Journal of Group Psychotherapy, 38,* 419–446.

*Zonnebelt-Smeenge, S. J. (1988). The effect of attending a bereavement support seminar on the level of expressed physical symptoms of bereaved spouses within the first year after the spouses' death: An experimental study (Doctoral dissertation, Western Michigan University, 1987). *Dissertation Abstracts International, 49,* 1379.

Received September 15, 1998
Revision received February 3, 1999
Accepted February 9, 1999 ■