The Effectiveness of Psychotherapeutic Interventions for Bereaved Persons: A Comprehensive Quantitative Review

Joseph M. Currier, Robert A. Neimeyer, and Jeffrey S. Berman
University of Memphis

Previous quantitative reviews of research on psychotherapeutic interventions for bereaved persons have yielded divergent findings and have not included many of the available controlled outcome studies. This meta-analysis summarizes results from 61 controlled studies to offer a more comprehensive integration of this literature. This review examined (a) the absolute effectiveness of bereavement interventions immediately following intervention and at follow-up assessments, (b) several of the clinically and theoretically relevant moderators of outcome, and (c) change over time among recipients of the interventions and individuals in no-intervention control groups. Overall, analyses showed that interventions had a small effect at posttreatment but no statistically significant benefit at follow-up. However, interventions that exclusively targeted grievers displaying marked difficulties adapting to loss had outcomes that compare favorably with psychotherapies for other difficulties. Other evidence suggested that the discouraging results for studies failing to screen for indications of distress could be attributed to a tendency among controls to improve naturally over time. The findings of the review underscore the importance of attending to the targeted population in the practice and study of psychotherapeutic interventions for bereaved persons.

Keywords: bereavement, grief, bereavement intervention, psychotherapy, meta-analysis

Losing a loved one to death is an inherent part of human life. Reactions to loss are as varied and multifaceted as the grievers themselves, although research indicates that the majority of bereaved people tend to experience strong emotions, a sense of cognitive disequilibrium, and impaired role functioning for at least a short period (see Bonanno & Kaltman, 2001, for a review). Despite the reality that almost every person will experience the pain of bereavement at some point, and that the majority will improve without formalized help, evidence suggests that many persons do not improve naturally (Bonanno et al., 2002; Bonanno, Wortman, & Nesse, 2004). For the 10% to 15% of bereaved persons who experience intense suffering, sometimes for years (Ott, 2003), the death of a loved one can precipitate a number of psychologically and medically debilitating symptoms, which can even prove fatal (Prigerson et al., 1997). Research on complicated or prolonged grief in particular also demonstrates that the griever’s distress is often not reducible to common psychological disorders (e.g., depression and posttraumatic stress disorder [PTSD]), for which people receive other psychotherapeutic interventions (see Lichtenholt, Cruess, & Prigerson, 2004, for a review). Moreover, recent studies have yielded evidence for the relative efficacy of interventions designed for complicated grievers as compared with more traditional treatments, such as interpersonal (Shear, Frank, Houch, & Reynolds, 2005) and supportive psychotherapies (Boelen, de Keijser, van den Hout, & van den Bout, 2007), thereby highlighting the relevance of interventions sensitive to the needs of those showing complicated grief symptoms and other clinically significant indications of poor bereavement adaptation. The purpose of this study is to provide a comprehensive quantitative summary of the effectiveness of bereavement interventions over a range of therapeutic outcomes for many different types of persons.

In general, years of outcome research leave little doubt that psychotherapy provides an effective means of help for the majority of distressed persons. Whatever theory drives the intervention and regardless of whether the therapy targets an individual, a family, or a group, studies usually conclude that treated clients are substantially better off than their untreated counterparts (see Lambert & Ogles, 2004; chap. 3 of Wampold, 2001, for reviews). However, attempts to evaluate the effectiveness of psychotherapies for bereaved persons have not yielded such conclusive results, and many important questions remain. Some researchers have claimed that no strong empirical basis exists for endorsing the effectiveness of these interventions (Jordan & Neimeyer, 2003; Schut & Stroebe, 2005); others have recently called for cautious optimism and recommended that the field suspend judgment as to whether bereavement interventions are less effective than other therapies (Larson & Hoyt, 2007). Even though narrative reviews have pinpointed studies that showed favorable outcomes, reviewers have reached different conclusions regarding the state of the evidence for the absolute efficacy of these interventions. These have included “no consistent pattern of results for well-designed studies”.
(Forte, Hill, Pazder, & Feudtner, 2004, p. 11), no empirical support for universal preventive efforts combined with cautious support for interventions with more distressed grievers (Schut, Stroebe, van den Bout, & Terheggen, 2001), “a small amount of quantitative evidence” for universal approaches with bereaved children (Curtis & Newman, 2001, p. 492), and mixed benefits for bereaved spouses (Potocky, 1993) and families (Schneiderman, Winders, Tallett, & Feldman, 1994).

Although researchers and clinicians increasingly have looked to meta-analysis as a way to integrate studies on the effectiveness of interventions, the existing quantitative reviews of bereavement interventions have failed to reach a clear consensus (Allumbaugh & Hoyt, 1999; Currier, Holland, & Neimeyer, 2007; Kato & Mann, 1999). As part of a larger qualitative review, Kato and Mann (1999) reviewed 11 studies that randomly assigned bereaved participants to intervention and no-intervention groups and examined modality and intervention type for differences in outcome. In comparison with the encouraging effects discussed in secondary reviews of outcomes in psychotherapy (Lambert & Ogles, 2004; chap. 3 of Wampold, 2001), Kato and Mann found a meager overall effect size (d = 0.11) and no clear moderators of outcome. Unfortunately, these authors departed from conventional meta-analytic procedures, thereby making it difficult to place full confidence in their results. First, they likely violated the assumption of statistical independence by treating different outcome measures within studies as separate observations in the overall analysis. Second, they adopted an overly conservative coding strategy and inferred effect sizes of d = 0 for four of the studies, each of which actually reported information permitting statistical derivation of non-zero effect sizes. Third, on the basis of their inclusion criteria, they failed to take into account nearly two thirds (n = 19) of the randomized controlled trials available for review before 1999.

In contrast to Kato and Mann (1999), Allumbaugh and Hoyt (1999) were clear about their methods and thorough in their exploration of moderators. Overall, Allumbaugh and Hoyt found a modest, but somewhat encouraging, aggregated effect size of 0.43 for the 35 bereavement interventions included in their review. Furthermore, they found that self-referred and/or clinically referred participants and those with a shorter length of bereavement showed greater benefit. The primary limitation of this review pertained to the use of statistical techniques to allow the inclusion of studies without a no-intervention control group. Supposing that each of the controlled studies reported sufficient information at each of the time points, Allumbaugh and Hoyt opted to use the average amount of change from baseline to posttreatment from the 19 no-intervention groups for the 16 studies that lacked a control group. For such a procedure to be viable, they had to assume that all of the control groups came from a homogeneous population, a possibility that they later rejected in favor of testing for systematic differences in outcome on the basis of study samples. Ironically, it appears that this same procedure that allowed the inclusion of studies without a control group also reduced the number of controlled studies by requiring baseline information and an inability to accommodate outcomes collected at different time points. Hence, Allumbaugh and Hoyt failed to include 24 of the more empirically sound studies that were available when they conducted their review more than 8 years ago.

A more recent meta-analytic review examined the effectiveness of bereavement interventions with children (Currier et al., 2007). We investigated 13 studies and, consistent with the overall finding of Kato and Mann (1999) for adults, found that the interventions failed to produce a statistically significant benefit for the bereaved children (d = 0.14). Even though the small sample restricted the number of questions we could explore, other results converged with Allumbaugh and Hoyt (1999) in that interventions occurring in a time-sensitive manner were shown to produce better outcomes. We also found support for Schut et al.’s (2001) conclusions that studies that either excluded distressed grievers altogether or failed to select children on the basis of their pre-intervention functioning yielded less favorable outcomes. In spite of the small overall effect size, this review therefore suggested that under certain circumstances bereavement interventions could possibly generate effect sizes comparable to those observed with therapies for other emotional and/or behavioral difficulties.

Unfortunately, the controversy over the supposed ineffectiveness of these interventions has arisen without the aid of a comprehensive and methodologically sound meta-analytic review. In addition, more than 8 years have now passed since the last quantitative reviews of the adult literature, and researchers have continued to study a diversity of psychotherapeutic interventions for bereaved individuals. As a way of exploring the issues of timing and selection further, along with other substantive factors that we were not able to examine in the child-oriented review (e.g., relation of outcome to client age), we expanded on Currier et al. (2007) to provide an up-to-date quantitative review of the entire controlled outcome literature on bereavement interventions.

Contrary to most psychotherapies, bereavement interventions are often practiced more as preventive aids than as treatments for clearly defined disorders or other specific problems in living. Whereas the overarching goal of preventive approaches for bereaved persons is to reduce the future probability of psychological or medical problems, treatment interventions aim for the immediate alleviation of these difficulties. Applying the framework issued by the Institute of Medicine (IOM; Mrazek & Haggerty, 1994), bereavement interventions can be grouped into three broad categories that vary in the amount of distress experienced by the targeted population. Universal interventions target anyone who suffers bereavement and do not distinguish on the basis of death-related risk factors or pre-intervention functioning (e.g., Scruby & Sloan, 1989). Although selective interventions do not focus on highly distressed grievers per se, they are geared toward particular groups of bereaved individuals who face a heightened risk of experiencing distress symptoms, such as those who lose a child to a violent death (e.g., Murphy et al., 1998). Lastly, there is a third group of indicated interventions (e.g., Wagner, Knaevelsrud, & Maercker, 2006) that restrict selection to those manifesting problems adapting to loss, which can include symptoms of an established psychiatric disorder (e.g., major depression; Reynolds et al., 1999) or other clinically significant difficulties (e.g., loss-related intrusions,
anxiety, and guilt feelings; Kleber & Brom, 1987). More aligned with the goals of treatment, these interventions aim for the direct amelioration of distress symptoms or resolution of the particular problem. Overall, qualitative (Jordan & Neimeyer, 2003; Schut & Stroebe, 2005; Schut et al., 2001) and quantitative (Currier et al., 2007) reviews have found that although there is some evidence for the usefulness of bereavement interventions with indicated griefers, little empirical support exists at this point for the effectiveness of universal and selective efforts.

In addition to evaluating the overall effectiveness of psychotherapeutic interventions for bereaved individuals and exploring the targeted population as a possible moderator, this review examined several other characteristics that can influence outcome. As researchers have argued that the timing (Currier et al., 2007; Larson & Hoyt, 2007) and method of recruitment (Larson & Hoyt, 2007; Schut & Stroebe, 2005; Schut et al., 2001) are critical considerations for evaluating the effectiveness of bereavement interventions, these two variables were explored. Because evidence suggests that males and females and persons of varying ages respond to loss differently (see Stroebe & Schut, 2001, for a review), the present review also explored differences for sex and age. The problematic implications of bereavement following violent modes of death (i.e., homicide, suicide, and fatal accident; Currier, Holland, & Neimeyer, 2006) and loss of a primary relationship, such as members of the nuclear family (i.e., spouse, child, parent, and sibling; Weiss, 1988), have been well established. Thus, this review explored the impact of violent loss and loss of a nuclear family member on outcome. Other substantive factors explored in this review included length of follow-up, dosage, modality, type of control group, attrition, and year of study report. As the measurement of outcome customarily plays a crucial role in evaluating an intervention (e.g., Smith, Glass, & Miller, 1980), the present review also examined within-study differences for the source of report and the domain of outcome used in the studies.

Even though the standardized mean difference between an intervention and a no-intervention control group at a single time point remains the best way to determine the effect of an intervention, this statistic does not directly address questions related to change over time for the same group of individuals. When it comes to intervening with bereaved persons, a substantial effect size for an intervention could be generated in a number of ways. For example, negative change in the control group and no improvement in the intervention group could yield an effect size identical to that produced by positive change in the intervention group and no improvement or even slight deterioration among controls. More importantly, discouraging results for bereavement interventions would raise several questions as well, including potential deterioration among intervention recipients or improved adjustment among those who did not receive any formalized help. In fact, the modest outcomes for universal and selective interventions (e.g., Currier et al., 2007; Kato & Mann, 1999; Schut et al., 2001) could be attributed to a pattern in which bereaved persons either worsen over time with intervention or display minimal distress without the aid of an expert helper (Bonanno et al., 2002, 2004), each of which would mitigate therapeutic gains for treated groups. Because the conventional standardized mean difference statistic cannot in itself address these questions, we also followed the procedures used by Horowitz and Garber (2006) to explore the magnitude and direction of change over time for intervention and control groups in studies of universal and selective interventions.

**Method**

**Studies**

The review was based on a total of 61 outcome studies reported in 64 papers, which included 48 published peer-reviewed articles and 16 unpublished dissertations. The studies were identified with two different search methods. First, we performed a comprehensive search of PsycINFO, PsycARTICLES, Medline, and Dissertation Abstracts International using groupings of relevant search terms, such as bereavement, grief, treatment outcome, intervention, control group, and evaluation. Second, reference sections of existing quantitative and narrative reviews and the identified studies themselves were consulted to locate other missing reports of outcome studies. The current review included 50 studies not included in Kato and Mann (1999) and 44 not covered in the meta-analytic review conducted by Allumbaugh and Hoyt (1999).

Several criteria were used to select the studies. First, the study had to evaluate a psychotherapeutic intervention aimed at promoting healthy adaptation to bereavement through sustained interaction with an identified expert helper. Self-help programs were included if the researchers specified that a leader possessing special skills and/or training facilitated the intervention. Only a few studies of pharmacological treatments meeting all of the inclusion criteria were located, and none of them tested drugs from the same class of medication, which makes aggregating these results for a meta-analytic review untenable at this time. A second criterion was that all of the study participants must have experienced the death of a loved one, regardless of study condition. Third, only controlled studies were included, meaning that all of the studies compared a group of bereaved persons who received grief therapy with a group of bereaved persons who did not receive any active type of intervention (e.g., wait-list control group, placebo). Although both randomized and nonrandom controlled studies were included, a fourth criterion was that none of the participants in the nonrandom studies were allowed to self-select to condition. Consequently, two studies included in Allumbaugh and Hoyt (1999)

---

1 It is important to note a slight discrepancy between our use of the term indicated in describing the third category of bereavement interventions in comparison with IOM’s definition (Mrazek & Haggerty, 1994). IOM defines an indicated intervention solely along preventive lines, as applying “to persons who, on examination, are found to manifest a risk factor, condition, or abnormality that identifies them, individually, as being at high risk for the future development of a disease” (p. 21). A proper distinction according to IOM would therefore classify cases of subclinical distress as indicated intervention and those with clinical levels of disorder as treatment. However, on the basis of the ways in which interventions for the bereaved are implemented, researchers and practitioners frequently rely on clinically significant indications of distress that do not necessarily align with or exceed a certain threshold for an established psychiatric disorder. In turn, such a procedure often makes indicated approaches to intervening with the bereaved indistinguishable from treatment or psychotherapy in general.

2 Because no-intervention controls were offered counseling in the studies that used a standard-of-care intervention for the comparison group (e.g., Kissane et al., 2006; Schut, de Keijser, van den Bout, & Stroebe, 1996; Sikkema et al., 2004), we excluded them from this review.
were excluded from this review (i.e., Cordsen, 1987; Lieberman & Videka-Sherman, 1986).
As meta-analysis requires the derivation of effect sizes, the final criterion was that the researchers provided sufficient statistical information or discussed not finding a statistically significant difference between the intervention and no-intervention control groups somewhere in the study report. Sufficient statistical information included situations in which the sample sizes, means, and standard deviations for the groups were given, or in cases where this information was not reported, studies needed to report the results of other statistical tests that allowed for the computation of effect sizes with other less exact estimation procedures (i.e., F test, p value, percentage improved). The formulas for estimating intervention effects in these instances are detailed elsewhere (for descriptions of these procedures, see chap. 5 of Glass, McGraw, & Smith, 1981; Miller & Berman, 1983; Appendix 7 of Smith et al, 1980).

The majority of the interventions used a group modality (63%), although individual (25%) and family (12%) approaches were represented as well. The researchers used some sort of treatment manual 56% of the time. The interventions included psychotherapy and counseling (63%), professionally organized support groups (17%), crisis intervention (11%), social activities groups (4%), writing therapy (3%), and a formal visiting service (1%) and helper training program (1%). The studies were conducted by researchers from psychology (34%), nursing (16%), psychiatry (13%), education (12%), counseling (8%), social work (5%), other medical specialties (5%), and other unspecified fields (7%). The studies were also coded for the targeted population (universal, selective, or indicated) and the method of recruitment (aggressive outreach, advertisements, or clinical referral processes). Other characteristics of the samples and the studies are presented in Table 1. These variables were usually coded on the basis of the mean, but the medians or midpoints of the ranges were used when authors failed to report this information. For example, in coding the relationship to the deceased for the study samples, we used the mean or another measure of central tendency that reflected the total proportion of individuals from the intervention and control groups who had lost a nuclear family member (i.e., spouse, parent, child, or sibling).

We assessed outcomes using three different reporting methods in the studies, with participants’ self-report (83%) being the method of choice compared with the reports of significant others (15%) and clinician-rated judgments (3%). Instruments that required the participants to respond to questions about themselves on a conventional rating scale were considered self-report. Rating scales that involved the evaluation of outward behaviors by other reporters were grouped as other report. Finally, there were instances in which outcomes were determined by a trained professional’s assessment of the participants’ functioning that was usually based on some accepted standard of distress (e.g., severity of depressive symptoms). Overall, researchers used the three reporting methods to assess outcomes in eight domains: general distress (23%), depression (17%), well-being (15%), grief (13%), relational functioning/social adjustment (11%), physical health symptoms (10%), anxiety (7%), and trauma symptoms (3%).

### Estimating Intervention Effects

The measures of the participants’ therapeutic outcomes were converted to a Cohen’s (1988) $d$, which is a standardized way of expressing the magnitude and direction of an intervention effect. The positive or negative valence of the effect size was calculated so that a positive $d$ always denoted the advantage of the bereavement intervention group. Aside from one instance in which significant attrition occurred between the first and second follow-up interviews (i.e., Black & Urbanowicz, 1987), follow-up effect sizes were always derived from the last point of outcome evaluation. We included all of the relevant outcomes from the studies. In Sandler et al.’s (1992, 2003) two evaluations of the Family Bereavement Program, in which they implemented a range of therapeutic activities with groups of parenally bereaved children of various ages and their surviving parents, effect sizes were only computed for the child participants. Even though the surviving parents’ “proximal” outcomes serve as a mediating process in the sophisticated underlying model of the Family Bereavement Program, this review was only concerned with evaluating “distal” outcomes, or the level of success with respect to the ultimate aims of the intervention. Finally, when researchers conducted analyses

<table>
<thead>
<tr>
<th>Sample characteristic</th>
<th>No. of studies</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participant average age (years)</td>
<td>59</td>
<td>40.1</td>
<td>21.2</td>
<td>8.00 to 71.00</td>
</tr>
<tr>
<td>Percentage female</td>
<td>57</td>
<td>71.2</td>
<td>24.8</td>
<td>0 to 100</td>
</tr>
<tr>
<td>Percentage loss of nuclear family member (i.e., spouse, parent, child, or sibling)</td>
<td>56</td>
<td>75.1</td>
<td>35.6</td>
<td>0 to 100</td>
</tr>
<tr>
<td>Percentage of violent loss (i.e., homicide, suicide, or fatal accident)</td>
<td>35</td>
<td>27.2</td>
<td>32.1</td>
<td>0 to 100</td>
</tr>
<tr>
<td>Percentage Caucasian</td>
<td>25</td>
<td>73.7</td>
<td>29.2</td>
<td>0 to 100</td>
</tr>
</tbody>
</table>

### Study Characteristic

- Year of study report
- Total participants at posttreatment
- Total participants at follow-up
- Length of follow-up (weeks)
- Number of sessions
- Weeks of intervention
- Total time of intervention (hours)
- Percentage total attrition
- Length of time since loss (months)

<table>
<thead>
<tr>
<th>No. of studies</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>61</td>
<td>1992</td>
<td>8.5</td>
<td>1975 to 2007</td>
</tr>
<tr>
<td>48</td>
<td>64.3</td>
<td>57.2</td>
<td>14 to 261</td>
</tr>
<tr>
<td>33</td>
<td>87.7</td>
<td>63.2</td>
<td>10 to 261</td>
</tr>
<tr>
<td>33</td>
<td>36.4</td>
<td>20.9</td>
<td>2 to 74</td>
</tr>
<tr>
<td>53</td>
<td>7.7</td>
<td>4.8</td>
<td>1 to 30</td>
</tr>
<tr>
<td>45</td>
<td>8.7</td>
<td>6.1</td>
<td>1 to 36</td>
</tr>
<tr>
<td>38</td>
<td>11.8</td>
<td>7.2</td>
<td>1 to 33</td>
</tr>
<tr>
<td>56</td>
<td>15.1</td>
<td>16.3</td>
<td>0 to 62</td>
</tr>
<tr>
<td>60</td>
<td>14.0</td>
<td>22.2</td>
<td>-3.0 to 111.7</td>
</tr>
</tbody>
</table>
on subsets of items from particular measures, effect sizes were always aggregated to yield one score for the total outcome measure.

For the primary analysis of intervention effects, 321 effect sizes were computed, including 190 at posttreatment and 131 at follow-up. On average, researchers used four measures to evaluate outcome at both time points. Whenever authors did not report sufficient statistics to allow for inclusion, this information was requested via e-mail, and additional data were provided so that an exact estimate could be calculated for four studies that otherwise would have been excluded from this review (Caserta & Lund, 1993; Dalton & Krout, 2005; Sandler et al., 2003; Tonkins & Lambert, 1996). Estimation procedures were not necessary for the majority of the computations, as 63% of the effect sizes were either calculated directly—from the sample sizes, group means, and standard deviations—or from t statistics. In 12% of cases, researchers gave group means and numbers of participants, but the pooled standard deviation had to be estimated from an F ratio with more than 1 degree of freedom in the numerator. In another 20% of cases, effect sizes were based on other, less exact estimation procedures from the available statistical information (see chap. 5 of Glass, McGaw, & Smith, 1981; Miller & Berman, 1983; Appendix 7 of Smith et al, 1980). Finally, the researchers failed to report any statistics for their measures 5% of the time but nonetheless discussed finding no statistically significant difference between groups somewhere in the paper. So as not to overlook the studies reporting null findings in narrative terms, the effect sizes were estimated to be zero in these cases. Overall, comparisons between intervention and control groups requiring one of the less exact estimation procedures yielded effect sizes that were 0.1 and 0.2 of a standard deviation larger than comparisons that did not require such a procedure at posttreatment and follow-up, respectively, which suggests that any assumptions did not introduce a high degree of bias in the derivation of effect sizes.

Because estimates of effect size based on smaller samples tend to overestimate the magnitude of effect sizes, Hedges’s (1984, Formula 4) correction formula was applied to all of the effect sizes reported in this review.

Estimating Change Over Time

As a way to further explore the results of universal and selective interventions, the standardized mean differences from baseline to posttreatment and/or baseline to follow-up were calculated for intervention and control groups. Specifically, we calculated change scores in terms of Cohen’s (1988) d by subtracting the mean at baseline from the mean at the later time point and dividing the difference by an estimate of the standard deviation of this difference. Rather than estimating differences between independent groups at a single time point, as we did for the intervention effects, these effect sizes represent change over time for the same individuals. Hence, a positive d always denoted positive change or improved functioning for the intervention recipients or control group. In 9 out of 10 cases, these effect sizes were based on the actual standard deviations from the two time points. We estimated the variability from the results of F tests for the remaining studies in this secondary analysis. On average, all of the effect sizes for the intervention and control groups included in the secondary analysis were based on three or four outcome measures.

Preliminary Analyses

Before conducting the formal analyses, we addressed a couple of potential statistical issues with the data. Given the comprehensive aims of this review, controlled studies were included regardless of how researchers assigned participants to conditions. In a model that included all of the studies, the mean effect size at posttreatment from randomized studies (d = 0.16) did not differ significantly from the mean effect size of nonrandom studies (d = 0.37), Qa(1, 46) = 2.23, p = .1. Similarly, the mean effect size at follow-up from randomized studies (d = 0.05) did not differ significantly from the mean effect size of nonrandom studies (d = 0.07), Qa(1, 31) = 0.01, p = .9. Nevertheless, consistent with other findings that quasi-experimental designs tend to be more variable in their outcomes than studies adhering to random assignment (Shadish & Ragsdale, 1996), studies that relied on a nonrandom method of assignment showed nearly five times as much variability in posttreatment effect sizes (variance = 0.86) than did the randomized controlled trials (variance = 0.18). So as not to violate the assumption of equality of variances for the two categories on the randomization factor, randomized and nonrandom designs were not combined in any of the analyses. Additionally, because of the limited number of nonrandom studies at posttreatment (n = 12) and follow-up (n = 6), we conducted tests of moderation only for studies that used random assignment. However, for the purpose of providing additional information, we performed a separate set of overall analyses on the nonrandom studies, and the results are reported for the interested reader.

A second issue involved the possibility of a publication bias in this literature (Begg, 1994). As a way of evaluating whether the present results may overestimate the helpfulness of bereavement interventions on account of a tendency for researchers not to report studies that generated results disconfirming their hypotheses, we conducted two sets of statistical tests. First, we investigated the association between sample size and effect size for the randomized studies. If studies with smaller samples that generated negative results were systematically missing from this literature, there would be a negative association between the total size of the sample and the degree of success shown for the interventions. Although there was some suggestion of such an association, analyses failed to achieve statistical significance at posttreatment, r(36) = −.23, p = .1. or follow-up, r(27) = −.20, p = .3. A second, more indirect reflection of publication bias would involve a tendency for published studies to yield higher effect sizes than

---

3 Because of the current controversy regarding the possible ineffectiveness of bereavement interventions, adopting this conservative procedure might appear suspect. Therefore, as a way of evaluating the tenability of this coding procedure, we ran the analyses without the zero study-level effect sizes included and found the same pattern of results as reported in the article.

4 The expected variance of a difference is based on the variance of each component along with the correlation between the two points (e.g., see Equation 2.20 of Winer, Brown, & Michels, 1991). However, the correlation between time points was rarely reported. In such cases, the variance of the difference was estimated by assuming the correlation between time points to be .5, thereby reducing the formula to the pooled standard deviation used to calculate Cohen’s d (1988) with independent groups at a single time point.
the unpublished dissertations included in this review. However, the mean effect size for published (posttreatment $d = 0.13$; follow-up $d = 0.06$) and unpublished (posttreatment $d = 0.22$; follow-up $d = -0.03$) studies failed to demonstrate significantly different outcomes at either time point, posttreatment. $Q_d(1, 34) = 0.62$, $p = .4$, follow-up $Q_d(1, 25) = 0.1$, $p = .8$. Taken together, results of these analyses failed to suggest that a strong publication bias exists in this area of research.5

Data Analysis

We used Lipsey and Wilson’s (2001) computational macros to estimate random effects models (see Hedges & Vevea, 1998; Shadish & Haddock, 1994, for a description of the random effects approach) for the primary analyses on the effectiveness of bereavement interventions.6 To ensure that the assumption of statistical independence would not be violated in the analyses, we conducted the analyses on effect sizes that were always aggregated to the study level so that each study contributed only one effect size to any particular analysis. For the second set of analyses, we used a weighted least squares regression procedure to compute the mean amount of change for intervention and control groups, with each study weighted by the size of the particular sample at the later time point such that groups with larger numbers of participants had a greater contribution to the overall average.

Results

Effectiveness of Bereavement Interventions

Overall effects. Effect sizes ranged from $-0.65$ to $2.54$ at posttreatment and from $-0.40$ to $0.76$ at follow-up for comparisons between intervention and no-intervention groups. The random effects weighted-average effect sizes at posttreatment and follow-up and the results of the homogeneity tests are presented in Table 2. Of the four overall analyses, bereavement interventions were shown to outperform no-intervention control conditions immediately following the intervention in the randomized and nonrandom studies. Both of the overall analyses for follow-up effect sizes failed to yield effect sizes that were significantly greater than zero, suggesting that bereavement interventions did not produce any general benefit at the point of follow-up. Although nonrandom studies appeared to present a more encouraging overall picture for bereavement interventions, the earlier preliminary analyses failed to demonstrate statistically significant differences between the outcomes of randomized and nonrandom studies at either assessment. Homogeneity tests revealed that variation in effect sizes remained to be accounted for across the randomized studies at each of the assessment points, which indicated the need to subdivide these studies on the basis of several clinically and theoretically relevant moderators.

Targeted population. To establish the reliability of classifying studies according to the targeted population, we arranged for independent coding to be done by the first author and two psychology students. Studies were coded as to whether researchers (a) intervened on a universal basis and did not attend to objective death-related risk factors (universal), (b) strictly focused on subsets of bereaved individuals at heightened risk of experiencing distress on the basis of the relationship to the deceased and/or violent types of death (selective), or (c) assessed for and only selected participants who were manifesting bereavement-related difficulties before the intervention began (indicated). Using the three categories, coders achieved a kappa value of .95. Altogether, the coders had one disparity on coding the targeted population, which was successfully resolved by consensus.

As shown in Table 3, there was a statistically significant overall effect for the targeted population at both time points. Weighting each study by the inverse of the conditional variance, Fisher’s least significant difference test revealed that the mean effect size at posttreatment from studies with grievers who showed indications of poor bereavement adaptation was significantly higher than that found for studies with a universal population ($p = .02$) or a selective population ($p = .05$). The results at follow-up demonstrated the same overall pattern; indicated interventions promoted

\[ Q = \frac{(b_1 - b_2)^2}{\text{var}(b_1) + \text{var}(b_2) - 2 \text{cov}(b_1, b_2)} \]

Note. SE = standard error; CI = confidence interval; $Q$ = degree of heterogeneity; $v$ = random effects variance component.

\[ ^{*} p > .05. \]

---

Table 2

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of studies</th>
<th>Mean effect size</th>
<th>SE</th>
<th>95% CI</th>
<th>$Q$</th>
<th>$v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttreatment</td>
<td>36</td>
<td>0.16$^*$</td>
<td>0.06</td>
<td>0.05 to 0.27</td>
<td>58.14$^*$</td>
<td>.04</td>
</tr>
<tr>
<td>Follow-up</td>
<td>27</td>
<td>0.05</td>
<td>0.05</td>
<td>-0.05 to 0.16</td>
<td>36.81</td>
<td>.02</td>
</tr>
<tr>
<td>Nonrandom studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posttreatment</td>
<td>12</td>
<td>0.51$^*$</td>
<td>0.20</td>
<td>0.13 to 0.90</td>
<td>35.84$^*$</td>
<td>.29</td>
</tr>
<tr>
<td>Follow-up</td>
<td>6</td>
<td>0.04</td>
<td>0.09</td>
<td>-0.13 to 0.22</td>
<td>8.25</td>
<td>.00</td>
</tr>
</tbody>
</table>

---

5 We also investigated publication bias using Egger et al.’s (1997) linear regression method and Duval and Tweedie’s (2000) trim and fill procedure; each of these alternative analytic approaches also failed to demonstrate any clear indication of publication bias in this literature.

6 Unlike a random effects model that relies on the modeling of variation associated with both random differences and sampling error, another method for analyzing meta-analytic data uses the variation from the actual differences in effect sizes between studies and weights these observations by the sample sizes of the studies (see Robinson, Berman, & Neimeyer, 1990, for example). As a way to check whether the pattern of results would change with this other method, all of the analyses were also performed using the empirical derivation of the variation with the weights defined by the harmonic mean for the sample sizes. This set of analyses yielded the same pattern of results as reported in this article.
significantly greater benefit than universal \((p = .01)\) or selective interventions \((p = .01)\).

**Timing of intervention.** Because of the differences in the average time elapsed since loss in the samples, the impact of timing was examined in three ways. First, analyses performed on all of the studies with average time since the loss as the predictor of effect sizes failed to suggest the influence of timing at posttreatment, \(r(35) = .20, p = .2,\) or follow-up, \(r(27) = -.27, p = .27\). Second, analyses were performed with only the universal and selective interventions, which presumably sampled greater proportions of grievers whose distress would abate naturally over time compared with the indicated grievers. These analyses also did not show that studies intervening earlier yielded significantly better outcomes at posttreatment, \(r(32) = .13, p = .4,\) or follow-up, \(r(25) = -.23, p = .3\). Finally, we explored a possible nonlinear association between timing and effect size; this analysis also failed to demonstrate clear support for the statistical superiority of studies that intervened after an intermediate period either at a posttreatment assessment, partial \(r(34) = -.03, p = .9,\) or at a point of follow-up, partial \(r(26) = -.21, p = .3\).

**Client recruitment.** On the basis of the researchers’ descriptions of their sampling procedures, strategies to initiate contact with bereaved persons were initially coded in one of three ways: (a) aggressive outreach (e.g., phone call or letter mailings), (b) media and community advertising, and (c) self-referrals and referrals from clinical professionals. The majority of the studies adhered to recruitment procedures that fell into a single category, although about a quarter (24%) of the remaining studies that clearly discussed their sampling procedures relied on a combination of advertising and clinical referrals. So as not to lose this information, we included these studies as a fourth category. Independent coders achieved agreement with the four-level coding scheme \((\kappa = 76;\) Fleiss, 1981). Overall, there were five discrepant ratings across the studies, each of which was again resolved by consensus.

Analysis of posttreatment effect sizes showed a statistically significant main effect for the method of recruitment, \(Q_b(3, 27) = 8.61, p = .03\). Weighting each study again by the inverse of the conditional variance, Fisher’s least significant difference test revealed that studies that strictly intervened with self- and clinically referred participants generated significantly better posttreatment outcomes (mean \(d = 0.40\)) than did studies using aggressive outreach procedures (mean \(d = 0.05\)), \(p = .03\). Notwithstanding some marginal differences, studies using community advertising (mean \(d = 0.05\)) or a combination of advertising and clinical referrals (mean \(d = 0.16\)) each yielded intermediate outcomes that could not be distinguished reliably from outcomes on the basis of other recruitment procedures (all \(p s > .1\)).

Although the four mean effect sizes showed a similar pattern at follow-up, each of the differences between the recruitment strategies failed to achieve statistical significance, \(Q_b(3, 20) = 1.55, p = .7\).  

**Modality of intervention.** As highlighted earlier, researchers tended to administer their interventions with groups of bereaved persons; however, interventions were also administered individually or in family formats. We first examined differences between group and individual interventions and found that these did not differ significantly from one another at posttreatment, \(Q_b(1, 31) = 0.18, p = .9,\) or follow-up, \(Q_b(1, 18) = 1.03, p = .3\). Second, we performed the same analyses with the family interventions included. These analyses also failed to show that bereavement interventions administered in a particular modality generated statistically superior outcomes at posttreatment, \(Q_b(2, 33) = 0.14, p = .9,\) or follow-up, \(Q_b(2, 24) = 1.16, p = .6\).

**Additional study features.** The remaining moderators we explored included (a) the relationship between the proportion of
female participants and outcome at posttreatment, \( r(34) = .18, p = .3 \), and follow-up, \( r(26) = .22, p = .27 \); (b) the linear effect of age with posttreatment outcomes, \( r(34) = -.16, p = .3 \), and follow-up outcomes, \( r(27) = -.25, p = .2 \), and the quadratic effect of age with posttreatment outcomes, partial \( r(33) = -.28, p = .1 \), and follow-up outcomes, partial \( r(26) = .22, p = .3 \); (c) total proportion of violent deaths in the sample and outcomes at posttreatment, \( r(19) = -.12, p = .6 \), and follow-up, \( r(26) = -.22, p = .4 \); (d) percentage loss of a nuclear family member with posttreatment outcomes, \( r(34) = -.08, p = .7 \), and follow-up outcomes, \( r(26) = .08, p = .7 \); (e) length of follow-up and follow-up outcomes, \( r(27) = -.06, p = .8 \); (f) number of sessions and outcome at posttreatment, \( r(30) = .02, p = .9 \), and follow-up, \( r(23) = .30, p = .2 \); (g) total time spent in intervention and outcomes at posttreatment \( r(26) = .07, p = .7 \), and follow-up, \( r(15) = -.23, p = .4 \); (h) type of control group and outcomes at posttreatment, \( Q(2, 35) = .08, p = .9 \), and follow-up, \( Q(1, 25) = 1.53, p = .2 \); (i) percentage of total attrition and outcomes at posttreatment, \( r(34) = .10, p = .5 \), and follow-up, \( r(26) = -.01, p = .9 \); and (j) year of publication and outcomes at posttreatment, \( r(36) = -.23, p = .1 \), and follow-up \( r(27) = -.17, p = .4 \). Thus, none of these additional factors was reliably related to outcome.

**Outcome measurement.** As a way of examining within-study variability among the different measures researchers used to assess outcome, we calculated one effect size for each specific outcome type by first averaging all relevant outcomes within each study and then aggregating them with a random effects model. Contrary to evidence that self-report increases the reactivity of an outcome measure (Smith et al., 1980), there was no clear tendency for measures that relied on participants’ self-report to produce significantly better outcomes than reports of a significant other or a clinician rating. Self-report was the only method to produce a statistically significant mean effect size at posttreatment (\( d = 0.16, SE = 0.06, p = .008 \)); however, the three 95% confidence intervals overlapped considerably, suggesting that the observed differences could be explained by chance. None of the methods of report generated mean effect sizes at follow-up that could be distinguished from zero (all \( p > .5 \)). Given the small percentage of studies using clinician ratings, we also collapsed them with ratings from significant others, and the 95% confidence intervals for results from self-report and those from this second category again overlapped considerably at both time points.

With regard to the eight domains that researchers used to gauge improvement, depression (mean \( d = 0.16, SE = 0.08, p = .04 \)) and relational functioning (mean \( d = 0.21, SE = 0.11, p = .05 \)) were found to be significantly greater than zero at posttreatment, and well-being (mean \( d = 0.25, SE = 0.08, p = .001 \)) was the only domain to generate a statistically significant mean effect size at follow-up. The 95% confidence intervals again overlapped across the eight domains at posttreatment, which did not support the statistical superiority of any particular domain of outcome immediately following bereavement intervention. In fact, the only non-overlapping 95% confidence intervals for any of the analyses on outcome measurement were between well-being and both depression and grief at follow-up, with measures of well-being generating a possible range of outcomes with values that were consistently higher than these other two domains.

**Change Over Time**

Similar to Horowitz and Garber’s (2006) review of the effectiveness of preventive interventions for depressive symptoms in children, none of the universal or selective interventions demonstrated clear evidence of a prevention effect: (a) worsening of distress symptoms in the no-intervention control group and (b) no decrease in distress symptoms in the intervention group. The weighted-average effect sizes for individuals who did not receive an intervention are presented in Table 4. At the level of the individual studies, although 29% of the control groups yielded negative study-level effect sizes at posttreatment, the majority of the effects were small in magnitude, and only one of the groups showed any evidence of deterioration at follow-up (\( d = -.05; Balk et al., 1998 \)).

Recipients of bereavement interventions also tended to show positive change, and all of the groups demonstrated further evidence of a treatment effect. In comparing the degree of improvement for intervention and control groups immediately following the intervention and then at follow-up, there was some indication that bereavement interventions may have accelerated the adjustment process but that no-intervention control groups achieved an equivalent amount of positive change over the course of time. It is also noteworthy that all of the intervention groups demonstrated improved functioning during the study and that only two groups generated a negative effect size at follow-up (\( d = -.28; Balk et al., 1993; d = -.05; Walls & Meyers, 1985 \)), which does not point to a deterioration in functioning among recipients of the interventions as an explanation for the small overall effects generated by universal and selective efforts. However, considering the decreases in distress shown by more than 70% of the control groups at posttreatment and almost 100% of the control groups at follow-up, it follows that interventions with nonindicated grievers would not add considerable benefit beyond the passage of time.

**Discussion**

The overall results from this review demonstrate that bereavement interventions have a small but statistically significant effect immediately following intervention but that therapeutic outcomes failed to differ reliably from zero at later follow-up assessments. These results contrast with meta-analytic reviews of general psychotherapy demonstrating that treatments help substantially to

<table>
<thead>
<tr>
<th>Table 4 Change Over Time in Control and Intervention Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>No-intervention controls</td>
</tr>
<tr>
<td>Posttreatment</td>
</tr>
<tr>
<td>Follow-up</td>
</tr>
<tr>
<td>Intervention participants</td>
</tr>
<tr>
<td>Posttreatment</td>
</tr>
<tr>
<td>Follow-up</td>
</tr>
</tbody>
</table>

*Note.* In calculating mean change over time, effect sizes were weighted by the size of the group at the last time point. SE = standard error; CI = confidence interval.

*\( p > .05.\)
ameliorate distress symptoms and to improve functioning (Lambert & Ogles, 2004; Smith et al., 1980; Wampold, 2001). Instead, the present results accord more closely with the majority of smaller-scale qualitative (Schut et al., 2001), quantitative (Currier et al., 2007; Kato & Mann, 1999), and secondary (Jordan & Neimeyer, 2003; Schut & Stroebe, 2005) reviews, each of which has concluded that, on average, recipients of bereavement interventions are not appreciably less distressed when compared with those who do not receive any formalized help. Beyond this general conclusion, the present comprehensive review documents the relevance of attending to the targeted population and reinforces the growing consensus that psychotherapeutic interventions for bereaved persons can be effective in instances when researchers and clinicians focus on persons who are genuinely in need of help.

Contrary to treatment interventions in which clients meet some predefined criteria for psychological disorder or present a specific need for help, bereavement interventions are frequently practiced in more of a preventive manner in which little attention is given to requisite manifestations of distress. The present findings for grief therapies are similar to those observed for interventions generically applied after other types of traumatic experiences (van Emmerik, Kamps-huis, Hulsbosch, & Emmelkamp, 2002). Such evidence challenges the common assumption in bereavement care that routine intervention should be provided on a universal basis or according to special objective circumstances surrounding the loss. Interventions targeting universal populations failed to produce better outcomes than would be expected by the passage of time, and although interventions with higher risk grievers showed a benefit at posttreatment, the gains were relatively small and the evidence failed to suggest that selective interventions yielded statistically significant outcomes at follow-up. Conversely, when the extra step was taken to assess for specific difficulties adapting to loss as a requirement for treatment, effect sizes compared favorably with the successes shown for psychotherapy in general. These results converge with the conclusions of others in the field (e.g., Jordan & Neimeyer, 2003; Schut & Stroebe, 2005; Schut et al., 2001) by pointing to an apparent relation between the level of bereavement-related distress and the likelihood of achieving successful therapeutic outcomes with bereaved clients. Viewed alongside the growing body of evidence that the passage of time frequently does not alleviate difficulties associated with maladaptive reactions to loss (Bonanno et al., 2002, 2004; Lichtenthal et al., 2004; Ott, 2003), it is encouraging that indicated grievers clearly were shown to benefit from intervention.

Aside from the targeted population and method of recruitment, no other factors were shown to relate systematically to outcome. Several of these analyses of moderators focused on person- and death-related risk factors and features of the interventions themselves. The present evidence also failed to show support for the timing of the intervention as a crucial moderator of outcome. In view of prior quantitative reviews that converged on a negative association between length of bereavement and outcome (Allumbaugh & Hoyt, 1999; Currier et al., 2007), it was surprising that even analyses of universal and selective samples did not yield differential levels of effectiveness based on the average amount of time elapsed since the loss. Insofar as it seemed possible that the limited effectiveness of bereavement interventions could be attributed to a reliance on aggressive outreach procedures (Larson & Hoyt, 2007; Schut & Stroebe, 2005; Schut et al., 2001), such as soliciting bereaved persons through obituaries or hospital records, it was also surprising that self- and/or clinically referred participants were not shown to achieve better outcomes across both assessments. Instead, the results suggest that the method of recruitment was only critical at posttreatment, at which time point studies intervening with referred clients generated better outcomes than those relying on aggressive outreach procedures. Although the current results highlight the need for greater focus on the subjective aspects of adaptation to loss and the identification of individuals presenting an immediate need of clinical intervention, these discrepancies with previous work highlight the importance of clarifying the issues of timing and recruitment with respect to the delivery and study of interventions for bereaved persons.

It is also noteworthy that effect sizes did not vary across the many different types of measures used to assess outcome in the studies. Overall, this finding contradicts a pattern in outcome research for more reactive measures—characterized in large part by client self-reporting of improvement and measuring changes that tap into what was actually done in the intervention—to show greater benefit (Smith et al., 1980). Even when researchers used self-report instruments or directly measured grief symptoms, the two instances in which one would anticipate finding the strongest benefit for bereavement interventions, effect sizes were still consistently weak in magnitude. Considering that four out of every five measures relied on the participants’ self-report of improvement, the lack of effectiveness shown for interventions with universal and selective samples becomes all the more striking, further underscoring the relevance of attending to the targeted population.

It remains possible that these results for outcome measurement could at least partially reflect an overreliance on generic measures of psychopathology or general functioning that are insensitive to the manifestations of bereavement adaptation (e.g., acute separation distress or attachment insecurity) most warranting attention in therapy (see Neimeyer & Hogan, 2001, for review). Furthermore, there was an absence of many of the potentially relevant nonpathologizing outcomes, such as constructing a subjective sense of understanding in loss (Currier et al., 2006; Keeseke, Currier, & Neimeyer, in press) and posttraumatic growth (Calhoun & Tedesci, 2006), each of which may provide useful information regarding underlying processes of improved outcomes via bereavement interventions. Therefore, even though issues of measurement were not shown to have a substantial influence on outcome, it seems possible that a lack of attention to crucial grief phenomena and elements of the restoration process may have limited chances to detect improvement across many of the studies.

Although the results thus far address questions of effectiveness, without also exploring the amount of change for intervention recipients and controls we cannot fully explain the reasons for the discouraging picture for bereavement interventions with nonindicated grievers. Similar to Horowitz and Garber’s (2006) findings in their review of preventive interventions for childhood depression, there was no clear evidence for a prevention effect among the universal and selective interventions. However, neither were there signs of deterioration for the typical intervention recipient. Instead, on average, all of the groups, both treated and untreated, displayed positive change at each assessment, and all of the positive intervention effects resulted from greater reductions in distress in intervention recipients relative to those who went without formalized help. In fact, it appears that bereavement interventions may...
have accelerated the adjustment process in many cases. Nonetheless, in view of the substantial improved adjustment of the control groups at posttreatment and their universal improvement at follow-up, it follows that the interventions would add little to no benefit beyond the participants’ existing resources and the passage of time. Given other evidence that the majority of bereaved individuals tend to regain pre-loss levels of functioning after a transitory period of distress without clinical intervention (e.g., 6 to 12 months; Bonanno et al., 2002, 2004), it seems likely that most of the control participants had successfully accommodated the experience of loss to varying degrees by the time of their involvement or over the course of the studies.

Quantitative reviews are only as sound as the studies on which they are based, and for this reason we chose to exclude all uncontrolled studies from this review. In addition, some of the current studies failed to report information on potential moderators, and these variables could not be coded or estimated in some cases, thereby leading to their exclusion from certain analyses and a reduction in statistical power. Ideally, all of the studies would have also adhered to the gold standard of random assignment (Shadish & Ragsdale, 1996), particularly in view of the fact that nonrandom studies showed a great deal more variability in their posttreatment outcomes than did randomized studies. Instead of excluding the information from nonrandom studies as Kato and Mann (1999) did, we reviewed these studies in separate analyses to gain a fuller picture of the absolute efficacy of bereavement interventions. However, we based our conclusions on the more methodologically sound studies that randomly assigned participants to treatment conditions.

Although we see distinct methodological advantages in analyzing the child-oriented studies alongside those reporting outcomes for adults, some might raise the possibility that the tests for moderation were flawed as a result of different therapeutic processes operating for children relative to adults. For this reason, we also performed the tests for moderation with only the adult studies included. As the pattern of results was the same in both instances, we did not exclude the child studies. Importantly, by retaining the child studies, we had the necessary range in ages to examine whether the outcomes of the interventions differed on the basis of age, which Allumbaugh and Hoyt (1999) had found to be an important predictor of treatment benefit in their meta-analysis but that we were unable to replicate in a larger sample of studies.

As in most intervention research, investigators in the reviewed studies reported treatment effects ignoring the fact that multiple clients were often treated by the same therapist or that the therapists were administered in a group format. Researchers have noted that such nesting can lead to the overestimation of treatment benefit because of therapist effects (Crits-Christoph, 1991; Martindale, 1978) or group effects (Baldwin, Murray, & Shadish, 2005; Rooney & Murray, 1996). Thus, the magnitude of intervention effects in our review may actually be an overestimate, and the actual effectiveness of bereavement interventions could be more modest than our results suggest. Unfortunately, the reviewed studies did not report the necessary statistical information that would permit accurate estimation of the degree of such nesting. However, overlooking potential nesting is common in primary outcome studies, and a similar inflation would also affect existing analyses of other types of therapeutic interventions. As such, the present estimates reported here likely represent the most appropriate values for comparing the effectiveness of bereavement interventions with current evidence for other therapies.

Another limitation of the existing literature that has implications for the interpretation of the current results relates to the tendency of researchers to evaluate outcome only on the basis of improved functioning of the individual client. However, particularly for preventive interventions with bereaved individuals, a more adequate interpretation of the effects would weigh the benefits to society alongside improvements in functioning at the individual level (Weisz, Sandler, Durlak, & Anton, 2005). For instance, selective interventions may have engendered practical benefits not reflected in the current set of effect sizes, such as improving educational performance and future employability among paren- tally bereaved children or reducing health-care utilization costs for those who lost a spouse later in life. We could not address such possibilities, as none of the studies assessed these more global outcomes. Additionally, in spite of the fact that we coded all of the relevant distal outcomes for each study, only a minority of the measures assessed outcomes that would be distinctly meaningful from a prevention point of view. Specifically, just over 1 out of 10 measured assessments indicated positive well-being, such as enhanced self-esteem and coping skills, each of which would be an important protective factor in diminishing the likelihood of bereavement-related distress symptoms emerging over time.

For all of these reasons, it is premature to dismiss the possible value of preventive interventions for bereaved persons. The designs of the majority of the available studies for universal and selective interventions do not provide sufficient evaluation of preventive goals. For example, of the 50% of universal and selective interventions that collected follow-up data, the average assessment was still less than a year from the end of the intervention. From a prevention point of view, researchers would want to assess these outcomes over several years rather than strictly focusing on shorter term changes. Whereas studies of indicated interventions require greater initial effort to identify distressed grievers, universal and selective interventions require greater effort well after the completion of the intervention to monitor the emergence of benefits over long periods of time. Unfortunately, most studies of preventive interventions have therefore relied on a methodological design best suited to evaluate the effectiveness of an indicated intervention. Because the primary goal of universal and selective interventions is not to reduce the severity of bereavement-related distress in the short-term but rather to provide the necessary resources to circumvent the onset of future difficulties, further research on preventive approaches to grief therapy needs to focus greater attention on measuring a range of therapeutic outcomes at meaningful follow-up periods.

Although studies of indicated interventions were more encouraging in their results, only five such studies could be identified that assessed outcomes at posttreatment (Denny & Lee, 1984; A. W. Forrest, 1990; Kleber & Brom, 1987; Reynolds et al., 1999; Wagner et al., 2006) and only two that did so at follow-up (Parkes, 1981; Raphael, 1977). Moreover, researchers used different definitions of distress to make decisions about who received the intervention in these studies, only one of which screened for complicated grief in the contemporary sense of the term (i.e., Wagner et al., 2006). Despite the important argument that bereavement is a multifaceted experience for which multiple outcomes and dimensions of distress can be relevant, such variability in identi-
fying distressed grievers suggests a need for greater focus concerning who is likely to benefit from clinical intervention. In particular, these differences support the value of selecting bereaved individuals with high levels of distress and further clarifying the thresholds for the various relevant outcomes (e.g., depression, anxiety, and PTSD) that best characterize individuals who need formalized help. Although the present results are largely mute concerning complicated or prolonged grief (see Lichtenhal et al., 2004, for review), assessing the utility of such an empirically derived diagnosis as a criterion for treatment deserves further consideration, especially in view of the growing evidence for the relative efficacy of specific treatments for this disorder (Boelen et al., 2007; Shear et al., 2005).

Another critical future task for researchers and clinicians is to flesh out the theoretical underpinnings and the operational implementation of their interventions in more detail. In the studies examined in this review, it was rare for researchers to describe how the interventions were developed, how they conceptualized the therapeutic process and its associated mechanisms of change, and how the interventions were linked to previous research on the needs of bereaved persons. Knowing that different types of individuals respond to loss differently (Bonanno & Kaltman, 2001; Stroebe & Schut, 2001), it is also improbable that a "one size fits all" approach will meet the needs of the majority of cases. Therefore, future intervention research should further identify approaches that work particularly well with subgroups of bereaved individuals and to report tests of moderation on the basis of clinically relevant factors (e.g., gender, pre-intervention level of functioning). If research further substantiates the benefit of indicated interventions, an understanding of the mechanisms of change would then be advanced more by implementing dismantling designs (Wampold, 2001) to identify what components of treatment are essential or expendable (e.g., psychoeducation, exposure to grief cues, encouragement of emotional expression, cognitive or meaning-oriented interventions, establishment of new life goals). As such outcome research proceeds, we would hope that clinicians might have available clearer guidelines regarding which bereaved individuals are likely to benefit from which types of intervention.

References

Asterisks denote controlled bereavement intervention studies examined in this review.


Received May 21, 2007
Revision received April 2, 2008
Accepted April 3, 2008

---

**AMERICAN PSYCHOLOGICAL ASSOCIATION**

**SUBSCRIPTION CLAIMS INFORMATION**

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do NOT duplicate claims through them and directly to us. PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.

PRINT FULL NAME OR KEY NAME OF INSTITUTION

ADDRESS__________________________________________________

CITY________________________________STATE/COUNTRY__ZIP__

YOUR NAME AND PHONE NUMBER____________________________

TITLE____________________________________________________

VOLUME OR YEAR__NUMBER OR MONTH____

(To be filled out by APA STAFF)

DATE RECEIVED:__________________________________________

ACTION TAKEN:___________________________________________

STAFF NAME:____________________________________________

DATE OF ACTION:________________________________________

INV. NO. & DATE:__________________________________________

LABEL NO. & DATE:________________________________________

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4–6 weeks.

Send this form to APA Subscription Claims, 750 First Street, NE, Washington, DC 20002-4242

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.