

**Methodological issues in studies of the effectiveness of grief intervention: Can the type of control group make a difference in assessment?**

**Arjen van Schuppen (3056031)  
Master Thesis Clinical and Health Psychology  
Utrecht University, June 2010**

**Supervised by:  
M. Stroebe, PhD Utrecht University The Netherlands  
H. Schut, PhD Utrecht University The Netherlands  
S. Fleming, PhD York University Canada**

## **Abstract**

There is as yet little consensus about the efficacy of grief interventions and a substantial number of grief intervention efficacy studies suffer from methodological shortcomings. An important aspect in efficacy studies is inclusion of a control group. The purpose of the present study is to answer the following research question: What is the expected effect of the Bereaved Families of Ontario (BFO)-intervention, given the level of impact found in other studies of bereavement intervention using comparable control groups? The non-intervention control group of the (BFO)-projects was compared to nineteen grief intervention efficacy studies. The latter were categorized into subgroups according to inclusion of four types of control groups; non-intervention, waiting list, other treatment and other. Weighted mean effect sizes were calculated using grief and depressive symptoms scales for each category and the expected effect of the BFO-intervention was estimated. It was predicted that, although the BFO-project incorporated some methodological improvements, the effect size would be small to moderate. A mixed between-within subject analysis of variance was conducted to assess the impact of the two conditions (intervention N =52, control N= 50) on participants' scores on the HGRC (grief questionnaire) and the depression scale of the SCL-90, across three time periods (pre-intervention, post-intervention and 6 months follow-up). As was expected considering the effect sizes of the comparable grief intervention efficacy studies, the BFO-intervention effect size was small. Implications of these results are discussed.

## **Introduction.**

Losing a loved one through death is a universal human experience. Responses to bereavement are characterized by substantial variability, whereby some gradually adapt to the loss while others experience difficulties over a long time (Caserta & Lund, 1993). Over the past decades, an increasing number and range of intervention programs have been established for bereaved parents, spouses and children. These include group or individual psychotherapy, hypnotherapy, trauma desensitization, bereavement support groups and self-help groups. Some of the interventions are professionally led, others are led by individuals who have resolved their own bereavement reactions and are trained to be a group leader (Vachon, Lyall, Rogers, Freedman-Letofsky & Freeman, 1980). The basic idea behind offering help is to benefit the bereaved individual, to help him or her to deal with the emotional and practical problems following the loss of a loved one (Schut, Stroebe, van den Bout & Terheggen, 2001). Despite the widespread agreement that grief interventions are sometimes desirable for individuals who have experienced a loss, there is as yet little consensus about the efficacy of grief interventions or about the client characteristics that might enhance the effectiveness (and therefore the desirability) of grief interventions (Litterer Allumbaugh & Hoyt, 1999). A great number of grief intervention efficacy studies have resulted in disappointing results. Reviews concerning efficacy of bereavement research give multiple explanations for these findings. For example Kato and Mann (1999) state that it might be possible that interventions for the bereaved are, in fact, not effective in relieving the symptoms of bereavement. Another possibility Kato and Mann (1999) give in their review is that perhaps psychological interventions for the bereaved can be effective but the interventions evaluated in these effectiveness studies were not powerful enough to relieve symptoms as intense as those felt by bereaved people. Lastly they state that it is possible for psychological interventions to reduce the suffering associated with bereavement but the methodological problems in these studies have prevented effects from being detected. A common shortcoming in grief intervention efficacy studies is the exclusion of a non-intervention control group. Without a control group it is impossible to draw valid cause and effect conclusions (Borkovec & Castonguay, 1998). A control group in effectiveness studies is important because if two conditions (experimental and control condition) differ in outcome, several competing factors can be rejected (history, maturation, attrition, selection bias) as the causal explanation of the difference in degree of change between the two conditions (Borkovec et al., 1998). But the most important reason why a control group is imperative in bereavement effectiveness studies is because mental and physical health difficulties following bereavement diminish over time (Schut et al., 2001). In other words, without a control group, one does not know whether improvement is due to the passage of time or to the impact of the intervention.

Since a control group in bereavement effectiveness research is such an important aspect it will be the main focus of the current report. Characteristics and adequacy of different kinds of control groups in bereavement effective intervention are pointed out, thereby deriving a yardstick by which the control group, an important component of the Bereaved Families Ontario (BFO) project, will be evaluated. The BFO-project is a collaboration between the organization BFO-Toronto, BFO-Halton Peel and the Coping Centre and two Universities, which are the University of York in Canada and the University of Utrecht in the Netherlands. The projects focuses on the efficacy of a mutual help intervention. The BFO-intervention is a well-established, longstanding program for bereaved parents relying on trained volunteers under the continued guidance of skilled professionals specialized in bereavement intervention (Fleming, Schut & Stroebe, 2005). The BFO-project incorporated some

methodological improvements compared with other grief intervention efficacy studies. For example, with respect to the recruitment procedure and the exclusion of a non-intervention control group (Schut et al., 2001; Litterer Allumbaugh et al., 1999; Kato et al., 1999). The BFO-project uses an “inreaching” instead of an “outreaching” recruitment procedure which means selection of those participants who need and expect such a program to be appropriate for them. Most studies in bereavement research use an “outreaching” recruitment procedure but, almost without exception, these studies report less favorable results (Schut et al., 2001). This may be explained by the fact that people asking for help will probably be more motivated and have more trust in the counselor or therapist. Furthermore the BFO-project uses a non-intervention control group.

The purpose of the present study is to answer the following research question: What is the expected effect of the BFO-intervention, given the level of impact found in other studies of bereavement intervention using comparable control groups?

To come to an answer of this question, an overview of comparable bereavement research projects will be given and overall control group characteristics will be evaluated. Then, the BFO-project will be introduced and analyzed statistically with a focus on the non-intervention control group. Lastly, the results and the control group characteristics of the BFO-project will be evaluated in the light of the formerly-discussed contemporary research. Conclusions and implications for future research will be drawn.

### **Overview: contemporary bereavement intervention effectiveness research**

The purpose of the current section is to introduce general conclusions from contemporary bereavement research. For this five reviews were used.

Schut et al. (2001), present a narrative review and a critical assessment of efficacy studies, categorizing these into general or primary preventive interventions, preventive interventions for high risk groups or secondary prevention, and interventions aimed at the treatment of complicated grief (tertiary prevention). The general pattern from this review was that the more complicated the grief process appears to be or to become, the better the chances of interventions leading to positive and lasting results (Schut et al., 2001). These results, of grief interventions getting more effective in complicated grief, contrast with those of Litterer Allumbaugh et al. (1999), who claimed that intervention should take place as soon as possible after the loss to be efficacious. A likely explanation for this contrast is that Litterer Allumbaugh et al. systematically excluded follow-up data in their meta-analyses, which indicates that there is a reduction in the magnitude of the effectiveness (Schut et al, 2001). Kato and Mann (1999) also did not take long-term effects into account in their review as they only used measures taken directly after the intervention ended (and corresponding control measures) for each study. They concluded that psychological interventions for bereavement are not effective. This exclusion of follow-up data makes it impossible to assess enduring benefits of the intervention. It is possible that this apparent time effect also reflects the presence of greater complication, insofar as the most recent criteria for the diagnosis of traumatic grief require marked symptomatology that persist for six months or more following loss (Prigerson & Jacobs, 2001). Jordan and Niemeyer (2003) concluded in their review that in general it appears that the scientifically demonstrated efficacy of formal interventions for the bereaved is distressingly low, far below that of most other types of psychotherapeutic interventions. Focusing on child bereavement, Currier, Holland and Niemeyer (2007) found results that do not support the assumption that bereavement interventions with children have a significant influence on adjustment (mean weighted effect size .14). They also concluded that children at high risk or children showing already signs of difficulty were

the most beneficial among the children who were sampled. This result supports findings of Schut et al. (2001) in adult bereavement.

From these five reviews we can derive the conclusion that, although there are some promising results, not all grief interventions lead to positive outcomes. In all five reviews methodological issues in grief intervention efficacy studies were pointed out as an explanation of the sometimes disappointing results. That is why in the current report the focus lies on one important methodological issue: control groups.

#### **Four types of control groups evaluated**

In this section a closer look is taken at the use of control groups in grief intervention efficacy studies. As mentioned before, nineteen studies that are comparable with the BFO-project were included in order to categorize different kinds of control groups and calculate weighted effect sizes. In Table 1 the bereavement projects evaluating the efficacy of intervention are summarized. To come to this selection, studies that were used in the majority of the reviews have been picked out and a computerized search of *Omega (1990-2010)* database is performed using the keywords: *Grief efficacy studies* and *Bereavement Interventions*. Not only in grief interventions but in any other efficacy research the use of a control group is essential. As we have seen, valid conclusions about the effect of an intervention can only be drawn when comparing results of an experimental group with a control group (Borkovec et al., 1998). Grief itself is a process that is expected to change over time. This change can be taken into account when using a control group as both groups are likely to experience a natural change over time, but the experimental group is expected to show more change due to an intervention. As can be seen in Table 1, the nineteen studies can be categorized into subgroups according to inclusion of four types of control groups. Each category will be highlighted in the following section, whereby control group characteristics of the nineteen studies will be discussed briefly.

The **non-intervention control group** is the condition in which treatment is withheld (Goodwin, 2005). Ideally, the participants in the non-intervention control group are identical to those in the experimental group in all ways except that the control group participants do not get the experimental treatment. In essence, the non-intervention control group provides a base line measure against which the experimental group's behavior can be compared. Lund and Caserta (1992), Caserta et al., (1993), Vachon, Lyall, Rogers, Freedman-Letofsky and Freeman (1980) and Norris Huss and Ritchie (1999) investigated a mutual help group intervention, comparable with the intervention used in the BFO-project. The non-intervention control groups in these four studies were comparable on demographic variables. In these studies, subjects were randomly assigned to the intervention or the control condition except in Lund et al. (1992) where self-selection was used. Self-selection creates the risk of comparing two non-equivalent groups. Brief psychotherapy (Lieberman & Yalom, 1992), problem-focused and emotion focused support for bereaved parents (Murphy, Johnson, Cain, Das Gupta, Dimond and Lohan 1998) and therapy aimed at emotional support and expression of emotions (Raphael, 1977) were the interventions under investigation in the remaining non-intervention control group studies. Murphy et al. (1998) focus on bereaved parents as does the BFO-project. This study has a relatively large sample and also low drop-out rates in both condition (at six month follow up 79% of the control group still participated in the study), which makes it an

**Table 1:** Summary of the outcome of 19 bereavement intervention effectiveness studies categorized by control group type.

Bereavement effectiveness studies	Outcome
<b>Waiting list control</b>	
<i>Brom et al. (1989)</i>	Different interventions did improve functioning on feelings of inadequacy and trait anxiety.
<i>Morrison Tonkins et al. (1996)</i>	Results in this study were strongly in favor of the intervention.
<i>Barrett (1978)</i>	Hardly no differences between conditions.
<i>McCallum et al. (1990)</i>	Favorable results on self-esteem, neuroticism, depression and life satisfaction.
<b>Non-intervention control</b>	
<i>Norris-Huss et al. (1999)</i>	The intervention did not affect self-esteem, depression, beliefs, or behavioral problems.
<i>Murphy et al. (1998)</i>	Woman with high mental distress and grief symptomatology improved. Woman with low levels of grief and mental distress were worse after intervention.
<i>Lieberman et al. (1992)</i>	No difference on any of the variables having to do with grief symptomatology and overall mental health.
<i>Lund et al. (1992)</i>	The number of group meetings had no effect on depression or bereavement.
<i>Caserta et al. (1993)</i>	Group attendance predicted less depression and grief.
<i>Raphael (1977)</i>	Persons in the intervention group had fewer symptoms than the high-risk control group but more than those in the low risk control group.
<i>Vachon et al. (1980)</i>	More intervention group members reported relief in emotional distress than control group members.
<b>Other treatment control</b>	
<i>Reich et al. (1989)</i>	The intervention hardly resulted in any positive result on mastery, distress, well-being, positive and negative affect.
<i>Kang et al. (2007)</i>	The experimental group showed significantly greater decrements in grief levels and symptoms of stress.
<i>Sandler et al. (2003)</i>	The intervention led to improved parenting, coping and caregiver mental health and to reductions in stressful events at posttest.
<i>Marmar et al. (1988)</i>	Both treatment were found equally effective.
<b>Other</b>	
<i>Goodkin et al. (1999)</i>	A significant reduction of psychological distress and grief scores was found after the intervention.
<i>Murray et al. (2000)</i>	Slow decline and hardly any difference between conditions immediately and 6 months after.
<i>Levy et al. (1993)</i>	6, 13 and 18 months after the loss no difference was found in levels of depression, anger, anxiety or stress.
<i>Lieberman et al. (1986)</i>	Participations showed more improvement than the normative sample on most measures.

impressive and useful study. Taken together, these studies all created the ideal situation in which the two conditions should be identical. In all except one study random assignment was used to create equivalent groups. Results vary in outcome. Positive effects were found especially among high risk individuals, but also negative outcomes were found in these studies.

Besides the typical control group situation in which a group is left untreated, there are a few more types of control groups that are worth describing. **The waiting list control groups** are often used to assess the effectiveness of some program or in studies on the effect of psychotherapy (Goodwin, 2005). In the case of grief intervention research with a waiting list control group those assigned to the waiting list are assured help after the study has ended and will be given the same treatment as the experimental group has had (Barrett, 1978; Brom, Kleber & Defares, 1989; McCallum & Piper, 1990; Morrison Tonkins & Lambert, 1996). See Table 1 for the outcome of these studies. Giving the waiting list participants an opportunity to benefit from some therapy procedure provides an important protection for participants' welfare, but it also creates pressures on the researcher to use this control procedure only for therapies or intervention programs of relatively brief duration (Goodwin, 2005). The mean number of sessions in waiting list control group studies evaluated in this is twelve sessions, which corresponds with the assumption that intervention programs should be brief when using a waiting list control group. Three studies investigated brief psychotherapy (McCallum et al., 1990; Brom et al., 1989; Morrison Tonkins et al., 1996) and all three showed positive results. The groups used in these studies were equivalent, even though Morrison Tonkins et al. (1996) used non-randomization to assign their small sample into the conditions. McCallum et al. (1990) use an "inreaching" procedure to recruit their participants as does the BFO-project. Barrett, (1978) did not find significant differences between the two conditions. She hypothesizes that the promise of participating in an intervention group in the near future may have had a therapeutic effect on the waiting list control group. In short, waiting list control groups have the advantage that all participants are assured help after the study has ended. On the other hand intervention should be brief because of ethical reasons. This is a disadvantage with respect to long-term assessment of effects (Schut et al, 2001). The studies described above showed results that were in favor of the intervention, except in Barrett's study where hardly any difference was found between the two conditions.

Another type of control group used in effectiveness studies is the **other treatment control group**. An alternative to a waiting list or non-intervention control group could be to compare the efficacy of the intervention with that of a well-established program with known effect (Schut et al., 2005). Kang and Yoo (2007) compared mutual help group subjects with subjects who only received a health check, showing significant decrements in grief levels in the intervention group over the control group. Because the researchers assigned participants on their own preference to either the intervention or the control group, the problem of creating non-equivalent groups arises. This was not the case in this study. Reich and Zautra (1989) matched 29 recently bereaved individuals with 29 non bereaved individuals. In matching, participants are grouped together on some trait, such as bereaved versus non bereaved, and then distributed randomly to the different groups in the experiment (Goodwin, 2005). The intervention condition received cognitive control orientated therapy and the control condition received social visits. Hardly any positive results were found. Again, in a study of Marmer, Horowitz, Weis, Wilner and Kaltreider (1988) both interventions (brief psychotherapy and mutual help) were found equally effective. Unfortunately the mutual help group, which can be seen as the control group in this study, suffered from extremely high premature termination of treatment (77%).

Schut et al. (2001) doubt whether the mutual help condition was really put to the test because only seven participants completed treatment. Finally, positive results are found in a study by Sandler, Ayers, Wolchik, Tein, Kwok, Haine, Twohey-Jacobs, Suter, Lin, Padgett-Jones, Weyer, Cole, Kriege and Griffin (2003). Subjects were randomly assigned to an active and collaborative learning therapy to enhance group support and a sense of personal efficacy or a self-study control group. The authors assigned a greater proportion to the intervention condition to help ensure that the intervention group would be at adequate size to be clinically viable. No difference were found in demographic variables between the two groups. Results are in favor of the intervention. In summary, these type of control groups are used in designs where two interventions are compared. But as Schut et al. (2001) states, unfortunately thus far such well-established intervention programs (i.e. well-described and transferable, tested, replicated and found effective, and accompanied by indications and counter indications) are not available in the area of grief counseling and grief therapy. So the studies using other treatment control groups must be interpreted with caution.

In this section **other** types of control groups will be discussed, for example those control groups which receive “care as usual”, which means control group participants receive care that would have been given to them even if they had not participated in the effectiveness study. An example is routine community care after infant death (Murray, Terry, Vance, Battistutta & Connolly, 2000) or psychosocial and medical care that is given to HIV-1 seropositive or HIV-1 seronegative homosexual men who lost a friend to AIDS (Goodkin, Blaney, Feaster, Baldewicz, Burkhalter and Leeds, 1999). The other two studies included in this category assigned non-participants of a mutual help group to a control group (Lieberman & Videka-Sherman, 1986; Levy, Derby & Martinkowski 1993). In Levy et al. (1993) subjects were assigned posthoc to the intervention or control condition. Subject who did not show up at sessions or drop-outs were assigned to the control condition. Such a procedural weakness makes it difficult to evaluate the study. Taken together, this type of control group needs to be interpreted with caution as well. An advantage of using a “care as usual” control is that it simulates a real world situation where care is available for everyone. But not every participant uses this opportunity for care in the same way, so there is a risk of comparing non-equivalent groups.

### **Effect sizes**

What is the expected effect of the BFO-intervention, given the level of impact found in other studies of bereavement intervention using comparable control groups? To get an overview of the impact found in the other studies, Cohen’s standards for defining effect sizes were used. Effect sizes at posttest and follow-up (if available) were calculated for grief scores or depression scores in cases where grief was not measured. Unfortunately, seven studies did not provide the necessary data to calculate effect sizes. Therefore, effect sizes were calculated for twelve studies using the following formula:

$$d = \frac{X_1 - X_2}{\sqrt{\frac{SD_1^2 + SD_2^2}{2}}}$$

$d$  = Effect size

$X_i$  = Mean group i

$SD_i$  = Standard deviation group i

The next step was to calculate a weighted mean effect size for each type of control group described earlier. To calculate the weighted mean effect size, studies that had samples with more than fifty participants deserved to have more weight in the calculation. These studies were put twice in the calculation. The control group effect size (effect of the passage of time) was subtracted from the intervention group effect size (effect of the passage of time + the effect of the intervention) so only the effect of the intervention would remain. In Table 2 the weighted mean effect size for each type of control group are summarized.

**Table 2:** Weighted mean effect size calculated for each type of control group using 12 of the 19 bereavement effectiveness studies.

<b>Type of Control Group</b>	<b>Weighted mean effect size</b>
<b>Waiting list control</b>	.56
<i>Brom et al (1989)*</i>	
<i>Morrison Tonkins et al. (1996)</i>	
<b>Non-intervention control</b>	.17
<i>Norris-Huss et al (1999)</i>	
<i>Murphy et al. (1998)*</i>	
<i>Lieberman et al. (1992)*</i>	
<b>Other treatment control</b>	.43
<i>Reich et al. (1989)*</i>	
<i>Kang et al. (2007)</i>	
<i>Sandler et al. (2003)*</i>	
<i>Marmar et al. (1988)*</i>	
<b>Other</b>	.47
<i>Goodkin et al. (1999)</i>	
<i>Murray et al. (2000)</i>	
<i>Levy et al. (1993)</i>	

\* Studies with large samples.

Looking at Table 2 and using Cohen's standards for defining effect sizes, it can be concluded that effect sizes are moderate to small. The non-intervention control group category has a small effect. It is important to take a look at this category specifically because in the BFO-project, the study described in the current report uses a non-intervention control group. To hypothesize that the expected effect of the BFO-project would therefore be small as well would not be appropriate. However, looking at all weighted mean effect sizes in Table 2 it would be realistic to expect a moderate to small effect size in the BFO-study.

## **Method**

As discussed earlier in the introduction, data from the BFO-study in Toronto (Canada), that examines the long-term effects of the loss of a child and the effectiveness of support groups, were used for this report. Efficacy found in this study is being compared in terms of effect sizes calculated above (Table 2). Variables that were selected to measure efficacy were grief and depressive symptoms.

### *Respondents*

The 203 respondents in the sample were primarily female (66.4%), ranging in age from 21 to 81 with a mean age of 50. The respondents in this study were recruited from May 2004 until April 2010 using

an inreaching recruitment procedure. To examine the effectiveness of support groups, two separate groups of bereaved parents were recruited. The first group of parents were part of a support group during the study (BFO- intervention group) and the second group of parents had never participated, nor were they going to participate, in a support group (non-intervention group). Neither had the latter received any other form of professional help during the project. While participating in this longitudinal project, both groups of respondents received five questionnaires across five years. One hundred and eighty-nine participants were included in the longitudinal study of which 141 participants were part of the intervention group and 62 participants were part of the non-intervention group. Participants were recruited using an “inreaching” procedure which is considered to be a methodological advantage of the BFO-study. Respondents in the non-intervention group of the longitudinal study should not be involved in a mutual support group or any form of professional help at the time of the study, but should be equivalent in that they would have liked to have the intervention under study had it been available. The control group participants were recruited either via a newspaper advertisement, local library and community-center flyers, through a radio interview with one of the involved researchers or they were informed and recruited at a conference of Mothers Against Drunk Driving (MADD).

#### *Procedure*

Bereaved parents who sought help at BFO (Toronto or Halton-Peel) or at the Coping Centre showed interest in participating were sent an information package. This package included an introduction letter, an informed consent form, and a questionnaire. Participation was voluntary and respondents had the right to withdraw from the study at any time. The parents who were willing to participate gave written consent on the basis of thorough information about anonymity and confidentiality. If the questionnaire was not returned within one month, an email was sent to remind the respondent politely of the questionnaire. When the questionnaire was returned respondents were thanked for their participation via email.

#### *Measures*

The original questionnaire of the longitudinal study consisted of several instruments; however only the instruments used for the current study (grief symptoms and depressive symptoms) will be described next. An appendix shows the actual items used in this study.

#### *Personal history and background*

Socio-demographic data about the parents and the deceased child were gathered, such as gender and age. Further information also included date of loss, cause of death, whether the death was expected or unexpected and other life stressors in addition to the child’s death. Furthermore, respondents were asked about current attendance at support groups and current consultation of other professional help (e.g. psychiatrist, psychologist, social worker, etc.).

#### *The Symptom Checklist 90-Revised (SCL-90-R)*

The SCL-90-R (Derogatis & Melisaratos, 1977) assesses a range of psychological problems and symptoms of psychopathology. The SCL-90-R is a 90-item scale self-report symptom inventory, with the subscales; somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation and psychoticism. For the current study only the subscale depression (twelve items) was used. Respondents had to indicate how much they were distressed by certain problems or complaints during the past week including today. Answers ranged from ‘Not at

all' to 'Extremely' on a 5-point scale. Internal consistency measured by Cronbach's alpha was .92 for the depression scale in the BFO-study.

#### *The Hogan Grief Response Checklist (HGRC)*

The HGRC (Hogan, Greenfield & Schmidt, 2001) measures the multidimensional nature of the adult grief process as it changes and evolves over time. Content analysis identified six categories; despair, panic behavior, blame and anger, disorganization, detachment and personal growth. Respondents had to determine on a 5-point scale which box described the best the way they were feeling during the past two weeks. Choices were from 'does not describe me at all' to 'describes me very well'. Internal consistency estimates using Cronbach's alpha coefficient ranging from .79-.90. In the BFO-study, the Cronbach alpha coefficient was .89. To determine convergent and divergent validity the HGRC subscales were compared to subscale scores on the Texas Revised Inventory of Grief (TRIG), Grief Experience Inventory (GEI), and Impact of Event Scale (IES), they correlated appropriately (Kristjanson, Lobb, Aoun & Monterosso, 2006).

#### *Analyses*

The data was analyzed using the SPSS program, version 16.0. For measuring reliability and differences in demographic variables the entire data set was used. Only participants who completed the grief and depression scale of the first three time periods were included in a mixed between-within subjects ANOVA.

#### **Results**

The impact found in other studies of bereavement intervention using comparable control groups has been discussed in the introduction and summarized in Table 2. It is hypothesized that the expected effect of the BFO-intervention would be moderate to small considering the weighted mean effect sizes calculated for the different types of control groups in the previous described studies. To decide whether to accept or reject this hypothesis, effect sizes of the variables grief and depression must be calculated.

First, the background differences between the intervention (N= 137) and control group (N= 61) were assessed. As mentioned in the Method section, the intervention group is larger than the control group. It was difficult to get a fair amount of response from the latter group of bereaved parents, which could be due to the heaviness of the loss. An independent-samples t-test was conducted to compare the age of control group and intervention group participants. Results show a significant difference in years for control group participants ( $M=53$  years old,  $SD= 11.66$ ) and intervention group participants ( $M= 49$  years old,  $SD= 10.93$ ),  $t(196) = -2.31$ ,  $p < .05.$ , indicating that control group participants were significantly older at the beginning of their participation in the study than intervention group participants. Also, participants in the control group ( $M= 1998$ ,  $SD= 3681$ ) lost their child significantly longer ago than in the intervention group ( $M= 2006$ ,  $SD= 1208$ ),  $t(199) = 7.74$ ,  $p < .05$ . This significant difference in the 'time since loss' between intervention and control group creates a fundamental problem. As control group participant have had more time to adapt to the loss, non-equivalent groups were compared. This can be seen if an independent samples t-test is conducted for grief and depression scores at pretest. Results show a significant difference in grief scores for control group participants ( $M=60.11$ ,  $SD=16.74$ ) and intervention group participants ( $M=70.38$ ,  $SD=15.28$ ),  $t(187) = 4.15$ ,  $p < .05.$ , indicating that control group participants had significantly lower grief scores than the intervention group participants at pre-test. Similar results were found when

pretest depression scores for the control group ( $M=36.83$ ,  $SD=14.56$ ) and the intervention group ( $M=43.23$ ,  $SD=13.12$ ) were compared,  $t(188) = 3.05$ ,  $p < .05$ ., indicating that control group participants had significantly lower depression scores than intervention group participants at pre-test. No additional analyses were conducted to control for the variable time since loss (see discussion).

### Grief

Grief scores in the BFO-study were assessed with twenty-five items of the HGCR (Hogan et al., 2001). To calculate total scale scores, twelve positive worded items needed to be recoded. Mean scores and standard deviation for grief over three measure points can be found in Table 3.

**Table 3:** Mean grief scores and standard deviation (SD) at pre-intervention, post-intervention (after 10 weeks) and after 6 months.

<b>Grief</b>		<b>Pre-intervention</b>	<b>Post-intervention</b>	<b>After 6 months</b>
<b>Intervention</b>	<i>Mean</i>	70.31	67.11	66.59
	<i>SD</i>	15.58	14.90	17.18
<b>Control</b>	<i>Mean</i>	58.86	58.47	59.06
	<i>SD</i>	17.30	15.86	16.82

A mixed between-within subjects analysis of variance was conducted to assess the impact of the two conditions (intervention, control) on participants' scores on the HGRC, across three time periods (pre-intervention, post-intervention and 6 months follow-up). As the BFO-study is still in progress, not all participants already received 5 questionnaire over 5 time periods. Therefore it was decided only to include the first three time periods. Samples of both intervention ( $N=52$ ) and control group ( $N= 50$ ) decreased in number over time. There was no significant interaction between condition and time, Wilks Lambda = .98,  $F(2, 88) = 1.01$ ,  $p = .37$ , partial eta squared = .02. Suggesting that the effect of time does not depends on condition and vice versa. There was no significant main effect for time, Wilks Lambda = .97,  $F(2, 88) = 1.22$ ,  $p = .29$ , partial eta squared = .02., indicating that the BFO-intervention was not effective in decreasing grief responses. The main effect comparing the two conditions was significant,  $F(1, 89) = 8.92$ ,  $p < .05$ , partial eta squared = .09. Control group participants had significant lower grief scores than the intervention group participants, pre and post-intervention.

Visual binning was used to transform the continuous variable age into a categorical variable. Subjects were divided in two groups, namely older and younger than the median age of fifty. Then, a mixed between-within subjects analysis of variance was conducted to assess the impact of the two conditions and the two age groups (<50, >51) on the participants' scores on the HGRC, across the three time periods. Mean grief scores and standard deviations are presented in Table 4.

**Table 4:** Mean grief scores and standard deviation at pre-intervention, post-intervention (after 10 weeks) and after 6 months for two age categories .

<b>Grief</b>		<b>Pre-intervention</b>	<b>Post-intervention</b>	<b>After 6 months</b>
<b>Intervention</b>	<i>Age &lt;50 (N=26)</i>	71.80 (13.44)	69.03 (16.02)	68.73 (16.99)
	<i>Age &gt;51 (N=26)</i>	69.61 (16.15)	66.07 (12.45)	66.46 (16.63)
<b>Control</b>	<i>Age &lt;50 (N=24)</i>	60.29 (18.83)	59.33 (16.93)	57.87 (18.37)
	<i>Age &gt;51 (N=26)</i>	59.07 (15.52)	58.11 (13.83)	60.53 (14.04)

This analysis is conducted to see if differences in the results would occur when analyzing age differences. This was important because control group participants were found significantly older than intervention group participants. There was no significant interaction between age and time, Wilks Lambda = .98,  $F(2, 97) = .59$ ,  $p = .55$ , partial eta squared = .01. Suggesting that the effect of time does not depend on age and vice versa. The second-order interaction time\*group\*age was also not significant, Wilks Lambda = .99,  $F(2, 97) = .38$ ,  $p = .68$ , partial eta squared = .08. Suggesting none of the independent variables were interacting. The main effect comparing the two conditions was significant,  $F(1, 98) = 10.09$ ,  $p < .05$ , partial eta squared = .10, indicating that control group participants had significant lower grief scores than intervention group participants, pre and post-intervention. But no significant main effect was found comparing the two age groups.  $F(2, 98) = .17$ ,  $p = .68$ , partial eta squared = .00, suggesting that although there is an age difference between intervention and control group, this difference is not related to grief scores across three time periods.

### *Depression*

Depressive symptoms in the BFO-study were assessed with the depression scale of the SCL-90 R (Derogatis, 1977). Mean scores and standard deviation for grief over three measure points can be found in Table 5.

**Table 5:** Mean depression scores and standard deviation (SD) at pre-intervention, post-intervention (after 10 weeks) and after 6 months.

<b>Depression</b>		<b>Pre-intervention</b>	<b>Post-intervention</b>	<b>After 6 months</b>
<b>Intervention</b>	<i>Mean</i>	44.53	42.51	40.17
	<i>SD</i>	14.52	14.02	13.41
<b>Control</b>	<i>Mean</i>	35.70	34.70	32.95
	<i>SD</i>	13.64	13.05	12.03

Total scale scores were calculated and another mixed between-within subject analysis of variance was conducted to assess the impact of the two conditions (intervention, control) on participants' scores on the depression scale, across three time periods (pre-intervention, post-intervention and 6 months follow-up). Again, only the first three time periods of the BFO-study were used. Samples of both intervention (N=51) and control group (N=51) decreased in number over time. There was no significant interaction between condition and time, Wilks Lambda = .99,  $F(2, 89) = .34$ ,  $p = .72$ , partial eta squared = .01. Suggesting that the effect of time does not depend on condition and vice versa. There was a main effect for time, Wilks Lambda = .87,  $F(2, 89) = 6.38$ ,  $p < .05$ , partial eta squared = .13, with both conditions showing a reduction of depressive symptoms across the three time periods (see Table 4). The main effect comparing the two conditions was also significant,  $F(1, 90) = 9.39$ ,  $p < .05$ , partial eta squared = .09, suggesting a difference in depression scores before and after the intervention. Control group participants had significant less depressive symptoms than intervention group participants. As mentioned before the variable age was categorized into two groups so age differences could be analyzed. Another mixed between-within subjects analyses of variance was conducted to assess the impact of the two conditions (intervention, control) and the two age groups (<50 and >51) on participants' scores on the depression scale, across three time periods. Mean depression scores and standard deviations are presented in Table 6.

**Table 6:** Mean depression scores and standard deviation at pre-intervention, post-intervention (after 10 weeks) and after 6 months for two age categories .

<b>Grief</b>		<b>Pre-intervention</b>	<b>Post-intervention</b>	<b>After 6 months</b>
<b>Intervention</b>	<i>Age &lt;50 (N=26)</i>	45.07 (13.33)	42.30 (12.21)	40.26 (12.72)
	<i>Age &gt;51 (N=25)</i>	42.12 (14.67)	41.24 (15.37)	38.72 (14.40)
<b>Control</b>	<i>Age &lt;50 (N=24)</i>	36.33 (15.00)	34.00 (14.31)	30.66 (10.96)
	<i>Age &gt;51 (N=27)</i>	36.07 (13.66)	35.51 (13.01)	36.22 (13.46)

There was no significant interaction between age and time, Wilks Lambda = .97,  $F(2, 97) = 1.61$ ,  $p = .20$ , partial eta squared = .03. Suggesting that the effect of time does not depend on age and vice versa. The second-order interaction time\*group\*age was also not significant, Wilks Lambda = .98,  $F(2, 97) = .96$ ,  $p = .39$ , partial eta squared = .02. Suggesting none of the independent variables were interacting. As in the first analyses, the main effect for time was found significant, Wilks Lambda = .89,  $F(2, 97) = 5.85$ ,  $p < .05$ , partial eta squared = .11, with both conditions and age groups showing a reduction of depressive symptoms across the three time periods. A significant main effect was found for condition suggesting that control group participants had significant lower grief scores than intervention group participants, pre and post-intervention  $F(1, 98) = 7.52$ ,  $p < .05$ , partial eta squared = .07. But no significant main effect was found comparing the two age groups.  $F(1, 98) = .01$ ,  $p = .93$ , partial eta squared = .00. Suggesting that although there is an age difference between intervention and control group, this difference is not related to depression scores across three time periods.

As there does not seem to be a difference in the results when analyzing age differences, the formula for calculating Cohen's  $d$  could be applied on the mean and standard deviations presented in Table 3 and 5. Effect sizes were calculated for posttest and follow up and intervention group effect sizes were subtracted from control group effect sizes. Then a mean effect size was calculated, indicating the effect of the BFO-intervention on grief and depression scores. In Table 7 an overview is given of the weighted mean effect sizes found in comparable research together with the mean effect size found in the BFO-project for both grief and depression. These effect sizes were calculated for posttest and follow-up time periods and were subtracted from the control group effect sizes to get a mean effect size which measures the effect of the intervention.

**Table 7:** Weighted mean effect size calculated for each type of control group using 12 of the 19 bereavement effectiveness studies + mean effect sizes of the BFO project.

<b>Type of Control Group</b>	<b>Weighted mean effect size</b>
<b>Waiting list control</b>	.56
<b>Non-intervention control</b>	.17
<b>Other treatment control</b>	.43
<b>Other</b>	.47
<b>BFO-project (grief)</b>	.20
<b>BFO-Project (depression)</b>	.09

Using Cohen's standards for defining effect sizes it can be concluded that effect sizes found in the BFO-project for grief and depression are moderate to small.

## Discussion

What is the expected effect of the BFO-intervention, given the level of impact found in other studies of bereavement intervention using comparable control groups? Nineteen grief intervention effectiveness studies were categorized into subgroups according to inclusion of four types of control groups; non-intervention, waiting list, other treatment and other. Weighted mean effect sizes were calculated using grief or depression scores for each category, showing small to moderate effects. Considering these effect sizes it was hypothesized that the BFO-project, despite its methodological improvements compared with other grief intervention efficacy studies, would show a small to moderate effect as well. Results show that grief symptoms in both the intervention and the non-intervention control group of the BFO-project did not change significantly over three time periods (pre-intervention, post-intervention and 6 months follow-up). Depressive symptoms in both the intervention and the non-intervention control group declined over the three time periods. The effect sizes at posttest and follow-up were subtracted from the control group effect sizes and a mean effect size could be calculated for both grief (.20) and depression (.09). Control group effect sizes were subtracted from the intervention group effect size because by using this method it was possible to see the effect of the intervention regardless of the improvement over time that is expected after experiencing a loss. That symptoms of bereavement tend to diminish over time can be seen in the results of the depressive symptoms, where both intervention and control group participants show less depressive symptoms at each time period. It is expected that the improvements in the control group were due to the passage of time and that improvements made in the intervention group were due to both the passage of time and the effect of the intervention. By subtracting the control group effect size from the intervention effect size, the effect of the intervention remains. The effect sizes found in the BFO-project were small, as was hypothesized considering the effects found in studies using comparable control groups. But still some effect is found for the BFO-intervention, especially for grief symptoms. It was hoped that due to the methodological improvements of using an inreaching recruitment procedure and a non-intervention control group, the BFO-project would show larger effects than were found in the current study.

As been briefly discussed in the Results section, control group participants lost their child significantly longer ago and have had more time to adapt to the loss. This resulted in a significant lower grief score and less depressive symptoms for control group participants at pretest. This difference creates a fundamental problem for interpreting the results found in the BFO-project. In essence non-equivalent groups were compared and this forms a limitation for the study. No additional analysis is been conducted in order to decline this influence. Miller and Chapman (2001) stated that the problem of preexisting group differences arises very commonly in psychopathology research and make it extremely tempting for researchers to use analytic methods in an attempt to avoid the interpretative problems that arise when groups differ pretreatment. They also stated that unfortunately, in general case, no such analytic method is available. Using the variable "time since loss" as a covariate to control for the pretest group difference would be inappropriate and would give unreliable results. "Time since loss" is so intimately associated with group conditions in the BFO-project that removal of variance in grief scores and depressive symptoms associated with "time since loss" would remove considerable variance in grief scores and depressive symptoms associated with the two conditions. There was also a difference in age between the two conditions but in this case, it was possible to control for this by including an additional analysis, which showed that this variable did not seem to influence the outcome.

Despite the described limitation of the study, useful information can still be derived from the BFO-project. For example the finding that even though years have passed since the loss, control group participants keep recuperating and keep showing less depressive symptoms at each time period is important to know. Beyond the examination of the specific research question posed in this study, some other interesting information has emerged from Table 1 and 2, where results and effect sizes were presented for the four control group categories. Interventions compared with a waiting list control group were apparently most efficacious. When an intervention was compared with a non-intervention control group, like in the BFO-project, it resulted in a less favorable outcome. These results are surprising as waiting lists control groups, next to improvements over time, also lead to hope and thus to some more improvement (Barrett, 1978). In a non-intervention control group, like mentioned earlier, improvements are only due to the passage of time and more effect of the intervention under study would be expected if the control group effect size is subtracted from the intervention group effect size. Interventions compared with non-intervention control groups also seem to be less efficacious than interventions compared with an “other treatment” or a “care as usual” control groups. Again, this is not what was expected as participants in other treatment conditions or participants receiving care as usual would show more improvements other than only improvements due to the passage of time and therefore intervention effects are expected to be smaller. The results described above lead to a very interesting issue namely: on what basis should future grief intervention efficacy researchers assess the appropriateness of the various options for control groups? If one looks at the results described above it seems fair to say that a waiting list control group would be the best option to compare an intervention with, because it would lead to the most favorable results. But this is not the case, when choosing a control group for an efficacy study one must choose on methodological grounds. Advantages and disadvantages of the different types of control groups must be assessed and the purpose of the study must be taken into account. A waiting list control group may have the disadvantage that, for ethical reasons, the intervention under study should be relatively short. But if the purpose of the study is assessing the intervention effect over a brief period of time a waiting list control group can be a good option. Another advantage could be that attrition can be held low because participants have the perspective of receiving an intervention in the near future. As mentioned earlier a non-intervention control group has the main advantage of showing the least improvement due to other factors than only the passage of time and therefore creating a good baseline measure to which an intervention can be compared.

As can be seen in the BFO-project, it is hard to create this described ideal situation. But improvements can still help the BFO-project to create a better balance between the two conditions under study. The most important improvement it should therefore make in the coming years is its recruitment of control group participants. More specifically, effort must be made to overcome the “time since loss” limitation described. In this process more initiative must be asked from the cooperating organizations BFO-Toronto, BFO-Halton-Peel and the Coping Centre. They are in contact with bereaved parents and receive daily phone calls of parents asking about the BFO-intervention. In quite a few occasions these parents decide not to participate because it does not fit in their daily schedules, or the intervention is not given in their own region. In this case, the possibility of participating in the control group of the BFO-project must be evaluated by the researchers. Recruiting these parents could help to overcome the “time since loss” difference between conditions, as their “time since loss” is expected to be comparable to the intervention group participants. Another advantage is that these parents contacted the organizations themselves so an inreaching recruitment procedure is still in order. On the other hand there is a risk of creating a

selection bias as participants who decide not to participate in the intervention might differ on certain variables from participants who are currently included in the project. Still, the difference in “time since loss” between the two conditions in the BFO-project is significant and it creates such a fundamental problem for interpreting it’s result that inclusion of this group of parents in the non-intervention control group should be evaluated seriously.

## References

- Barrett, C.J. (1978). Effectiveness of widows' groups in facilitating change. *Journal of Consulting and Clinical Psychology, 46*, 20-31.
- Borkovec, T. D., & Castonguay, L. G. (1998). What is the scientific meaning of empirically supported therapy? *Journal of Consulting and Clinical Psychology, 66*, 136-142.
- Brom, D., Kleber, R. J., & Defares, P. B. (1989). Brief psychotherapy for posttraumatic stress disorders. *Journal of Consulting and Clinical Psychology, 57*, 607-612.
- 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.
- Derogatis, R.L., & Melisaratos, M. (1977). *The SCL-90 manual 1: Scoring, interpretations and procedures for the SCL-90*. Clinical Psychometric Research, Baltimore.
- Fleming, S., Schut, H. A. W., & Stroebe, M. S. (2004). Benefits of intervention for bereaved parents in the BFO program: An empirical investigation. Research proposal.
- Goodkin, K., Blaney, N. T., Feaster, D. J., Baldewicz, T., Burkhalter, J. E., & Leeds, B. (1999). *Archives of General Psychiatry, 56*, 52-59.
- Goodwin, C. J. (2005). *Research in psychology: methods and design. Fourth edition*. Crawfordsville: John Wiley & Sons, Inc.
- Hogan, N.S., Greenfield, D.B. & Schmidt, L.A. (2001). Development and validation of the Hogan Grief Reaction Checklist. *Death Studies, 25*, 1-32.
- Jordan, J. R., & Niemyer, R. A. (2003). Does grief counseling work? *Death Studies, 27*, 765-786.
- Currier, J. M., Holland, J. M., & Niemyer, R. A. (2007). The effectiveness of bereavement interventions with children: A meta-analytic review of controlled outcome research. *Journal of Clinical Child and Adolescent Psychology, 36*, 253-259.
- Kang, H. Y. & Yoo, Y. S. (2007). Effects of a bereavement intervention program in middle-aged widows in Korea. *Archives of Psychiatric Nursing, 21*, 132-140.
- Kato, P. M., & Mann, T. (1999). A synthesis of psychological interventions for the bereaved. *Clinical Psychology Review, 19*, 275-296.
- Levy, L. H., Derby, J. K., & Martinkowski, K. S. (1993). Effects of membership in bereavement support groups on adaptation to conjugal bereavement. *American Journal Of Community Psychology, 21*, 361-381.
- 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.

- Litterer Allumbaugh, D., & Hoyt, W. T. (1999). Effectiveness of grief therapy: A controlled study. *Journal of Counseling Psychology, 46*, 370-380.
- Lund, D. A., & Caserta, M. S. (1992). Older bereaved spouses' participation in self-help groups. *Omega, 25*, 47-61.
- Marmer, C. R., Horowitz, M. J., Weiss, D. S., Wilner, N. R., & Kaltreider, N. B. (1988). A controlled trial of brief psychotherapy and mutual-help Group treatment of conjugal bereavement. *American Journal of Psychiatry, 145*, 203-209.
- McCallum, M. & Piper, W. E. (1990). A controlled study of effectiveness and patient suitability for short-term psychotherapy. *International Journal of Group Psychotherapy, 40*, 431-452.
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 110*, 40-48.
- Morrison Tonkins, S. A., & Lambert, M. J. (1996). A treatment outcome study of bereavement groups for children. *Child & Adolescent Social Work Journal, 13*, 3-21.
- Murphy, S. A., Johnson, C., Cain, K. C., Das Gupta, A., Dimond, M., Lohan, J., & Baugher, R. (1998). Broad spectrum group treatment for parents bereaved by violent deaths of their 12 to 28 year-old children. *Death Studies, 22*, 209-236.
- Murray, J. A., Terry, D. J., Vance, J. C., Battistutta, D., & Connolly, Y. (2000). Effects of a program of intervention on parental distress following infant death. *Death Studies, 24*, 275-305.
- Norris Huss, S. & Ritchie, M. (1999). Effectiveness of a group for parentally bereaved children. *Journal for Specialists in Group Work, 24*, 186-196.
- Prigerson, H., & Jacobs, S. (2001). Traumatic grief as a distinct disorder: A rationale, consensus criteria, and a preliminary empirical test. In M. S. Stroebe, R. O. Hansson, W. Stroebe, & H. Schut (Eds.), *Handbook of bereavement research* (pp. 613-646). Washington, DC: American Psychological Association.
- Raphael, B. (1977). Preventive intervention with the recently bereaved. *Archives of General Psychiatry, 34*, 1450-1454.
- Reich, G. W., & Zautra, A. G. (1989). A perceived control intervention for at risk older adults. *Psychology and Aging, 4*, 415-424.
- Sandler, I. N., Ayers, T. S., Wolchik, S. A., Tein, J. Y., Kwok, O. M., Haine, R. A., Twohey-Jacobs, J., Suter, J., Lin, K., Padgett-Jones, S., & Weyer, J. L. (2003). The family bereavement program: Efficacy evaluation of a theory-based prevention program for parentally bereaved children and adolescents. *Journal of Counseling and Clinical Psychology, 71*, 587-600.
- Schut, H., Stroebe, M. S., van de Bout, J., & Terheggen, M. (2001). The efficacy of bereavement interventions: determining who benefits. In M.S. Stroebe, R.O Hansson, W. Stroebe & H. Schut (Eds). *Handbook of bereavement research: Consequences, coping, and care.* (pp. 375-404). Washington, DC, US: American Psychological Association.

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.