J Psychopathol Behav Assess (2009) 31:190–201
DOI 10.1007/s10862-008-9113-1

The Underlying Structure of Grief: A Taxometric Investigation of Prolonged and Normal Reactions to Loss

Jason M. Holland · Robert A. Neimeyer · Paul A. Boelen · Holly G. Prigerson

Published online: 13 December 2008
© Springer Science + Business Media, LLC 2008

Abstract Recent studies have supported the distinctiveness of complicated and prolonged forms of grief as a cluster of symptoms that is separate from other psychiatric disorders. The distinction between prolonged and normal reactions to loss remains unclear, however, with some believing that prolonged grief represents a qualitatively distinct clinical entity and others conceptualizing it as the extreme end of a continuum. Thus, in this study a taxometric methodology was used to examine the underlying structure of grief. Participants included 1,069 bereaved individuals who had lost a first-degree relative. Each participant completed the Dutch version of the Inventory of Complicated Grief–Revised, which was used to create indicators of prolonged grief. The mean above and mean below a cut (MAMBAC) and maximum eigenvalue (MAXEIG) tests supported a dimensional conceptualization, indicating that pathological reactions might be best defined by the severity of grief symptoms rather than the presence or absence of specific symptoms.

Keywords Complicated grief · Prolonged grief disorder · Bereavement · Death and dying · Taxometric method

J. M. Holland (✉)
Center for Health Care Evaluation,
VA Palo Alto Health Care System,
Stanford University Medical Center,
795 Willow Rd. (152-MPD),
Menlo Park, CA 94025, USA
e-mail: jmhollnd@gmail.com

R. A. Neimeyer
University of Memphis,
Memphis, TN, USA

P. A. Boelen
Utrecht University,
Utrecht, The Netherlands

H. G. Prigerson
Center for Psycho-Oncology and Palliative Care Research,
Dana-Farber Cancer Institute,
Boston, MA, USA

H. G. Prigerson
Department of Psychiatry, Brigham and Women's Hospital,
Harvard Medical School,
Boston, MA, USA

H. G. Prigerson
Harvard Medical School Center for Palliative Care,
Harvard Medical School,
Boston, MA, USA

In contrast to the Diagnostic and Statistical Manual of Mental Disorder's (DSM-IV) view of grief as a normative stressor (American Psychiatric Association 2000), clinicians and researchers historically have made categorical distinctions between "normal" and "pathological" reactions to loss (e.g., Bowlby 1980; Worden 1982). Although theorists have differed in the way they have classified grief experiences (Bonanno and Kaltman, 2001), the concept of *complicated grief* or *prolonged grief*[1] has gained momentum over the last 10 to 15 years as a distinct psychological disorder that differs from normal reactions to loss (Lichtenthal et al. 2004). In particular, prolonged grief disorder (PGD) refers to a state of chronic grieving that persists for 6 months or longer and is characterized by intense separation distress, intrusive and emotionally troubling thoughts about the deceased, a sense of meaninglessness, trouble accepting the loss, and functional impairment (Prigerson and Jacobs 2001).

[1] Prolonged grief disorder is currently being considered for inclusion in DSM-V and in the past has also been referred to as *complicated grief* or *traumatic grief*. To maintain consistency the term "prolonged grief disorder" or PGD is used throughout this paper.

Although PGD has been shown to co-occur with other psychiatric disorders such as major depressive disorder (MDD), posttraumatic stress disorder (PTSD), and generalized anxiety disorder (GAD; Barry et al. 2002; Melhem et al. 2001), the symptoms associated with PGD have been shown to be fundamentally distinct from these other syndromes in numerous investigations (Boelen and van den Bout 2005; Boelen et al. 2003a; Chen et al. 1999; Prigerson et al. 1995a, 1996). In particular, PGD is primarily defined by symptoms of intense separation distress (e.g., yearning for the deceased), which are not presently represented by any other disorder in DSM-IV (Prigerson et al. 2008). PGD has also demonstrated impressive utility as a unique predictor (controlling for symptoms of mood and anxiety disorders) of physical and psychological outcomes, such as suicidality (Latham and Prigerson 2004), high blood pressure and heart problems (Prigerson et al. 1997) as well as psychological adjustment (Bonanno et al. 2007).

Despite this wealth of evidence supporting the unique impact of PGD, it remains unclear whether this disorder represents a qualitatively distinct clinical phenomenon that is distinguishable from normal reactions to loss or if it is better conceptualized as the extreme end of a continuum. Up to this point, normal grief has been described as being characterized by some symptoms, such as sadness and missing the deceased, particularly in the early stages of loss. However, in contrast to prolonged grievers, these individuals experience a gradual decrease in grief symptoms and eventually reinvest in relationships and activities (Lichtenthal et al. 2004).

Crucially, greater understanding of the underlying structure of normal and prolonged reactions, as either discrete categories or parts of a larger continuum, carries important clinical implications. These implications are illustrated best when considering recent meta-analyses of grief therapy (Allumbaugh and Hoyt 1999; Currier et al. 2007; Currier et al. 2008b; Kato and Mann 1999), which on average have yielded effects substantially smaller than most psychological interventions (i.e., effect sizes ranging from .11 to .43 averaged across a variety of relevant outcomes). Some evidence suggests that these small treatment effects might be due to lax inclusion criteria that allowed many normal grievers with little symptomatology into the treatment conditions (Currier et al. 2007; Currier et al. 2008b; Schut et al. 2001). Thus, a more precise understanding of the boundary between PGD and normal grief could enhance the effectiveness of interventions. For example, if a natural dividing line were found that distinguished those with "pathological" symptoms of grief from those found to be "normal," then this could provide a basis for identifying individuals who might benefit the most from treatment. Given these potential implications, the purpose of the present investigation is to examine the issue of whether grief is better represented on a continuum or as two distinct categories—normal grief and PGD.

Because no other study has examined this research question explicitly, past research that bears on this issue has yielded ambiguous results. One of the most frequently cited studies in support of a categorical conceptualization of PGD and normal grief is a longitudinal investigation that used cluster analysis. In this study, 11 of 120 participants were identified who exhibited a severe grief response, comparable to PGD, across all assessments from 1 to 13 months post-loss (Middleton et al. 1996). In another longitudinal study, 32 of 185 widows and widowers were found to exhibit a similar trajectory over time, characterized by minimal distress before the loss occurred followed by elevated and chronic symptoms of grief and depression post-loss (Bonanno et al. 2002). Although both of these studies parsed out a group of individuals who exhibited symptoms of PGD, the methods used were specifically designed to find homogeneous clusters of participants, and hence did not test the alternate possibility that grief symptoms might be better represented as a continuum.

In the absence of strong evidence to the contrary, some grief researchers have conceptualized PGD as the extreme end of a distribution that includes normal reactions to loss (e.g., Goodkin et al. 2006; Prigerson and Maciejewski 2006). Support for this view has been offered by a study conducted by Hogan et al. (2004) that demonstrated a high degree of convergence between the subscales of the Hogan Grief Reaction Checklist (HGRC; Hogan et al. 2001), a measure that attempts to assess normal grief, and the Separation Distress and Traumatic Distress symptom clusters derived from the Inventory of Complicated Grief-Revised (ICG-R; Prigerson and Jacobs 2001), a measure of PGD. The authors of this study interpreted these findings as a failure to establish the uniqueness of PGD from normal grief, arguing that if the symptoms of PGD were truly unique then the two symptom clusters derived from the ICG-R should correlate more highly with each other than with the HGRC subscales.

It should be noted, however, that others have challenged these findings and pointed out that the HGRC also assesses many of the symptoms of PGD, such as despair, detachment, disorganization, and anger (Prigerson and Maciejewski 2006). Interestingly, investigations that compare the ICG-R to different measures of normal grief have indeed yielded results that diverge somewhat from Hogan et al. (2004) findings. In one study, the ICG-R was found to distinguish between those bereaved by violent and non-violent means more sensitively than the Core Bereavement Items (Burnett et al. 1997), suggesting that the ICG-R might in fact measure unique symptoms exhibited by at—risk grievers (Currier et al. 2008a). Additionally, a confirmatory factor analytic study found good fit for a model, in which items from the Texas Revised Inventory of Grief (TRIG; Faschingbauer 1981), another normal grief measure, and items from the ICG-R loaded on separate, but correlated, latent factors (Boelen and van den Bout 2008). In this study, and others, the ICG-R was also found to be better than the TRIG at predicting detrimental outcomes, such as quality of life impairments, concurrent psychopathology, and general health (Boelen and van den Bout 2008; Boelen et al. 2003b; Prigerson et al. 1995b).

Given these conflicting findings, the present investigation sought to examine the underlying structure of grief using a group of statistical procedures, collectively referred to as the taxometric method (Meehl 1995), that are specifically designed to distinguish between continuous and categorical structures. In the present investigation, two competing hypotheses will be tested: (1) that prolonged and normal grief reactions are categorically distinct or (2) alternately that normal and prolonged reactions fall along different ends of the same continuum.

Method

Participants and Recruitment

Data from 1,069 bereaved individuals were used to examine the proposed research question. This sample was a restricted subset of individuals from two groups of mourners who had participated in two earlier studies, one that evaluated the psychometric properties of the ICG-R (Boelen et al. 2001) and another that examined the role of cognitions in grief (Boelen and Lensvelt-Mulders 2005). Both studies for which data were originally collected were approved by the institutional review board of the faculty of social sciences of Utrecht University. The two groups were recruited in 1998-1999 and 2001-2002, respectively, through an advertisement on a Dutch Internet-site that provided general information about grief and loss. To obtain informed consent, participants were asked to enter their name and email address and then click the "send" button if they understood and agreed that their information would be used for research purposes and handled confidentially. The questionnaires were administered on the Internet, and participants' responses were subsequently coded (in numbers and letters) and entered into a secure data file to protect their personal information. Additional information about these recruitment procedures and the original projects for which these data were collected can be found in previous publications (Boelen et al. 2001; Boelen and Lensvelt-Mulders 2005).

Because the previous studies for which these samples were originally drawn required a certain degree of homogeneity, only respondents between 18 and 65 years of age who had lost a first-degree relative were retained. In addition, unreliable data and data from people who had likely filled in the questionnaires more than once were removed. Examples of data that were deemed unreliable were from respondents who indicated that they lost a child older than they were or a parent younger than they were. Additionally, demographic variables were examined to determine if the same respondents had possibly filled in the questionnaires more than once. Specifically, successive sets of responses in which at least five demographic variables had exactly the same values were judged to be filled in by the same person and were removed.

With respect to the first group of mourners (Boelen et al. 2001), 1564 people participated in the initial period of data collection. Data of 233 participants (15%) were removed because they were deemed unreliable, another 269 participants were excluded because they had lost someone other than a first-degree relative, and 92 participants were excluded because they were under 18 or over 65 years of age. In the second group (Boelen and Lensvelt-Mulders 2005), 671 individuals participated in the data collection. Data of 71 people (11%) were removed because they were deemed unreliable. Then, data were removed from those who lost someone other than a first-degree relative (n=71) and those under 18 or over 65 years of age (n=21). In the present study, 409 additional participants who were bereaved less than 6 months ago were excluded given that a diagnosis of PGD can only be made 6 months post-loss or later (Prigerson and Maciejewski 2006), bringing the total combined sample to 1,069. The criterion of 6 month duration stems from research that demonstrates that PGD symptoms measured prior to 6 months post-loss do not reliably predict later symptoms or other negative outcomes (Prigerson et al. 1997).

The two study-groups differed in number of women, $\chi(1)=46.39$, $p<.001$, with the larger sub-sample (n=691) being made up of 67.4% women and the smaller one (n=378) being made up of 86.5% women. However, these groups did not differ on any other background variables or in terms of PGD severity. Given the similarity of these groups with regard to their demographics, symptom severity, and method of recruitment, it was considered acceptable to combine the two groups for this study. Overall, the mean age of the participants was 38.5 (SD=11.0) years. Most (74.1%) were female, 29.0% had lost a spouse, 16.6% a child, 45.0% a parent, and 9.4% a sibling. Cause of death was non-violent in 85.3% of cases and violent (i.e., due to accident, homicide, or suicide) in 14.7% of cases. Losses occurred on average 44.4 (SD=45.7) months ago. Although preliminary findings suggest that symptoms of PGD are often observed up to 2 years but may decline by 4 years after a loss (Boerner et al. 2005;

Prigerson and Jacobs 2001), all participants who had lost a loved one 6 months ago or longer were included in this study. An inclusive approach of this kind that maximizes the raw number of prolonged grievers, while possibly inflating the overall ratio of normal to prolonged grievers, has few drawbacks in a taxometric analysis. Notably, a recent simulation study suggests that the inclusion of additional "complement" members (in this case normal grievers) does not typically impact the ability of a taxometric analysis to identify latent structure (Ruscio and Ruscio 2004).

Measure

The Dutch version of the Inventory of Complicated Grief-Revised (ICG-R; Boelen et al. 2001; Boelen et al. 2003b) was used to assess symptoms of PGD in this study. This measure is composed of 29 declarative statements to which responses are made on a five-point Likert-type scale describing the frequency of symptoms (e.g., from 1 = *never* to 5 = *always*). Items assess such symptoms as the bereaved person's preoccupation with thoughts of the deceased; hallucinations; disbelief about the death; feelings of being shocked and overwhelmed, numb, out of control, anxious and unsafe, and purposeless; avoidance behaviors; and disturbances in sleep. The Dutch ICG-R has been tested with a sample independent from the present one, where it displayed high internal consistency (α=.94), concurrent validity (r=.71) with scores from the TRIG (Faschingbauer 1981), and good test-retest reliability (r=.92) over a period ranging from 9 to 28 days (Boelen et al. 2003a, b). In the present sample, this scale was found to be internally consistent (α=.94) and had an average corrected item-total correlation of .57 (*SD*=.16). The ICG-R has also been shown to predict a range of serious long-term health and mental health consequences of bereavement, justifying its interpretation as a measure of PGD symptomatology (Ott 2003; Prigerson et al. 1997; Prigerson and Jacobs 2001). In addition, PGD diagnoses made using the ICG-R have been shown to have perfect concordance with diagnoses made using a structured diagnostic interview, which in turn has demonstrated acceptable levels of agreement (kappa=.71) with raters' global assessment of grief (Prigerson and Jacobs, 2001). The Dutch ICG-R can be found in a previous publication (Boelen et al. 2003b).

Taxometric Procedures

In this study, taxometric methodology (Meehl 1995; Waller and Meehl 1998) was used to examine the underlying structure of grief. These analytic procedures search for relations among variables that are unique to latent categories. These latent categories are technically referred to as the *taxon* and *complement*, and in this study, they would be PGD and normal grief, respectively. If a construct is found to have a continuous distribution, then the latent structure would be regarded as *dimensional*. In contrast, if a construct appears to be better represented as a set of discrete categories then it would be termed *taxonic*. Rather than relying on tests of statistical significance, taxometric methodology emphasizes the use of consistency tests, which involves analyzing one's data in multiple, non-redundant ways to evaluate the reliability of the findings. One important way in which this is accomplished is to use two or more mathematically distinct, statistical procedures for evaluating a construct's latent structure. In this study MAMBAC (mean above and mean below a cut; Meehl and Yonce 1994) and MAXEIG (maximum eigenvalue; Waller and Meehl 1998) procedures were used[2].

The MAMBAC procedure examines latent structure by using two or more non-redundant indicators that are presumed to distinguish members of the taxon from members of the complement. One indicator (or in some cases the sum of several indicators) serves as the *input* indicator, which is used to align participants in ascending order along the x axis of a graph. A series of cuts is then made along this x axis, and at each point the mean score on the remaining *output* indicator is calculated for participants falling above and below the cut. The absolute difference between these means is then plotted on the y-axis of the graph. This procedure is based on the notion that if grief represents a taxonic construct, it will possess an optimal cutting point that separates participants with PGD from normal grievers, which will then be observed as a distinct peak on the graph. In contrast, if the latent structure of grief is dimensional, then the graph will tend to resemble a U-shape that lacks a well-defined peak.

The MAXEIG procedure follows a somewhat different logic. In this approach an input indicator is used to divide participants into consecutive sub-samples referred to as *windows*, which can be adjacent or overlapping (i.e., participants appear in more than one window). The degree of covariation among two or more *output* indicators for each of these consecutive windows is then calculated as an eigenvalue (the multivariate equivalent of covariance) and plotted on the y-axis. In this approach, if the latent structure of grief were taxonic, one would expect the covariation among the output indicators to be relatively small when a given window consisted of a homogenous group of either prolonged or normal grievers. However, in this same

[2] Other taxometric procedures, such as L-Mode (latent mode factor analysis; Waller and Meehl 1998) and MAXSLOPE (maximum slope; Grove 2004), were not implemented in this study because when they were used to analyze simulated datasets that mimicked the properties of the research data (e.g., in terms of sample size, indicator skew), but had known latent structures, ambiguous results were obtained.

scenario, if a window contained an equal number of prolonged and normal grievers, the covariation among the output indicators would presumably reach a maximum value, resulting in a peak in the graphed curve. Conversely, if the latent structure of grief is dimensional, the resulting curve would tend to resemble a flat line, as the covariation among the output indicators should be similar across successive windows.

Preliminary Classification of Cases

Before running the taxometric analyses, participants were first classified as either normal or prolonged grievers using the diagnostic criteria put forward in a recent field trial of PGD in order to get an accurate base rate estimate of the putative taxon (Prigerson et al. 2008). Notably, Criterion A Event Criterion) and Criterion D (Duration) were not used to classify cases but were rather used as inclusion criteria for this study, in that all participants had to have lost a person of significance and be bereaved for 6 months or longer. To be considered a prolonged griever a participant had to report on the ICG-R that they "often" or "always" experience at least one symptom of separation distress (criterion B); at least five cognitive, emotional, and behavioral symptoms (criterion C); and at least one form of functional impairment (criterion E)[3]. Using this set of (criteria, 17.8% of participants (n=190) were classified as prolonged grievers, which mirrors past research that has estimated the base rate of PGD to be between 10% and 20% (Bonanno and Kaltman 2001; Prigerson et al. 2008), established mostly with samples of older bereaved spouses.

Selection of Indicators

Because individual ICG-R items are on a 5-point scale, it was necessary to sum two of these items when creating indicators of PGD to obtain a wider spread of scores, which typically yields more interpretable taxometric curves. Additionally, two non-redundant sets of these indicators were constructed as a consistency test to see if similar findings could be obtained across more than one configuration of PGD indicators (see Table 1).

ICG-R Paired Indicators The first set of indicators was constructed by using principal components analysis with oblique rotation to identify non-redundant item pairs that capture a substantial portion of the variability on the ICG-R. All factors with an eigenvalue greater than 1 were retained, yielding five factors. For each factor, the two items with the highest standardized regression coefficients were summed together to form an indicator.

Table 1 Sets of indicators used in taxometric analyses

ICG-R Paired Indicators	Dutch ICG-R Items
Indicator 1	21. Future holds no meaning or purpose
	23. Unable to imagine life being fulfilling
Indicator 2	15. Hearing the voice of deceased
	16. Seeing the deceased
Indicator 3	10. Hard to trust people
	11. Lost the ability to care about, or feel distant from people
Indicator 4	27. On edge, jumpy, or easily startled
	29. Difficulties with sleeping
Indicator 5	4. Trouble accepting the death
	9. Stunned, dazed, or shocked
PGD Criteria Indicators	**Dutch ICG-R Items**
Indicator 1 (Criterion B) Separation Distress	3. Upsetting memories of the deceased
	5. Longing and yearning for the deceased
Indicator 2 (Criterion E) Functional Impairment	28. Experiencing impairments in functioning
	29. Difficulties with sleeping
Indicator 3 (Criterion C) Disbelief	8. Disbelief over the death
	9. Stunned, dazed, or shocked
Indicator 4 (Criterion C) Interpersonal Problems	10. Hard to trust people
	11. Lost the ability to care about, or feel distant from people
Indicator 5 (Criterion C) Acceptance	4. Trouble accepting the death
	13. Avoiding reminders of the deceased

Criterion A (Event Criterion) and Criterion D (Duration) were not used to create indicators but were rather used as inclusion criteria for this study, in that all participants had to have lost a person of significance and be bereaved for 6 months or longer

Once these indicators were created, their suitability for a taxometric analysis was then evaluated. To be considered sufficiently valid for such a test, a set of indicators must display low nuisance covariance (i.e., $r \leq .30$), expressed as the average of the intercorrelations within the hypothesized taxon and complement, and must also be able to separate the proposed groups reliably (i.e., $d \geq 1.25$; Meehl 1995). Using these five indicators, the nuisance covariance within the group of prolonged grievers was found to be .14, and the nuisance covariance within the group of normal grievers was .27. On average, these five indicators separated the hypothesized taxon and complement by 1.25 standard deviation units. Given these parameters and a sample size of 1069, taxometric investigations have typically generated reliable findings (Meehl 1995). Thus, this set of indicators was deemed to be valid and retained for later analyses.

[3] The proposed diagnostic criteria for PGD can be found in several recent publications (e.g., Prigerson and Maciejewski 2006; Prigerson et al. 2008). In this study, Separation Distress was represented using items 3, 5, and 22. Cognitive, Emotional, and Behavioral Symptoms were represented by items 4, 7, 9, 10, 13, 14, 17, 19, 21, 23, and 24. Impairment was represented by items 28 and 29.

PGD Criteria Indicators The second set of indicators was created by summing item pairs on the ICG-R that map onto the latest diagnostic criteria for PGD (Prigerson et al. 2008). Criterion B (Separation Distress) was measured by summing two items relating to symptoms of "longing and yearning" as well as "upsetting memories of the deceased." Likewise, Criterion E (Impairment) was assessed by combining an item dealing with general "impairments in functioning" with another item that inquired about "difficulties with sleeping." Because Criterion C (Cognitive, Emotional, and Behavioral Symptoms) was composed of a multitude of diverse symptoms, this criterion was broken down into six item pairs, which were grouped together based on their similar item content. In particular, these six pairs of items gauged symptoms of anger, shock and disbelief, interpersonal problems, meaninglessness, numbness, and difficulties accepting the loss. Upon further inspection, the item pairs that measured anger, meaninglessness, and numbness were found to increase nuisance covariance and/or did not adequately separate normal and prolonged grievers. Hence, these three item pairs were removed from the analyses. Thus, five unique indicators were used that tapped into separation distress, functional impairment, shock and disbelief, interpersonal problems, and difficulties accepting the loss. Though there was some redundancy between this set of indicators and the ICG-R Paired Indicators, this set included five previously unused ICG-R items and also excluded five items used in the prior configuration.

Overall, these five indicators exhibited nuisance covariance of .16 and .33 within the groups of prolonged and normal grievers, respectively. As a whole, this set of indicators separated the hypothesized taxon and complement by 1.28 standard deviation units. Because this set of indicators met most, but not all, of Meehl's (1995) criteria for a taxometric analysis (i.e., the nuisance covariance in the putative taxon was slightly above .30), the validity of these indicators was examined further using simulated data sets.

Simulated Data

Before analyzing the research data with these two sets of indicators, simulated data sets were created by drawing samples from a unit normal distribution and using a bootstrap procedure to mimic many of the parameters of the actual data (e.g., in terms of sample size, number and distributions of the indicators, and correlation matrix of the indicators). For each set of indicators two simulated data sets were created—one with a clearly dimensional structure and another with an unambiguous taxonic structure that reproduced indicator correlations and distributions within the groups of normal and prolonged grievers, as classified earlier. These simulated data sets then provided a way of testing the adequacy of the data and chosen indicators as well as guiding the plan of analysis. Additionally, results obtained from simulated data were also used as a point of comparison when interpreting the shape of MAMBAC and MAXEIG curves generated from the research data. In this case, 10 simulated data sets were created for each latent structure in order to provide a sampling distribution of results for comparison. Additional information about how simulated data sets are constructed can be found in previous works (e.g., Meehl and Yonce 1994, 1996; Ruscio et al. 2006).

Analysis

A plan of analysis was arrived at by systematically manipulating a variety of factors and observing which strategy best identified the known latent structures of simulated data. Using four simulated data sets (a dimensional and a taxonic data set that mimicked the properties of each of the two sets of indicators), MAMBAC analyses were found to be most sensitive when: (1) one indicator served as the output indicator and the sum of all other indicators was used as the input, (2) 50 evenly spaced cuts were made along the x-axis, starting and ending at the 25 most extreme cases, and (3) participants were resorted along the input indicator 20 times and the averaged results across these replications were plotted on the y-axis, which reduced sampling error resulting from cuts made between equal scoring cases. Because this plan of analysis yielded distinct MAMBAC curves for the simulated data sets, regardless of which set of indicators were being mimicked, it was deemed a fair and valid test of latent structure.

For the MAXEIG analyses, the latent structure of the simulated data sets was most apparent when: (1) a single indicator was used as the input and the remaining indicators served as the output, (2) cases were resorted along the input indicator and divided into windows 20 times, and (3) windows shared 90% overlap with neighbouring windows, which tended to have a smoothing effect on the curves. For the simulated data sets that mimicked the properties of the ICG-R Paired Indicators, the clearest results could be found when 130 windows were used with 77 participants in each window. Conversely, 120 windows (83 participants per window) yielded the most interpretable results for simulated data sets based on the PGD Criteria Indicators. Since the latent structures of all of the simulated data sets could be clearly identified using this plan of analysis, this strategy was considered a fair taxometric test and both sets of indicators were deemed valid. All analyses in this investigation were conducted using John Ruscio's suite of taxometric programs, which can be found at www.taxometricmethod.com.

Results

ICG-R Paired Indicators

A MAMBAC analysis was first run using the ICG-R Paired Indicators. Because this analysis used the composite of four indicators as the input, this analysis yielded five MAMBAC curves, each of which used a different indicator as the output. All five of these curves appeared U-shaped, which is consistent with a dimensional structure. Furthermore, when these five curves were combined to form an averaged curve (shown in Fig. 1), it was found to match the shape of the simulated dimensional data more closely than the simulated taxonic data, lending additional support for a dimensional conceptualization of normal and prolonged grief.

Beyond a visual inspection, an objective measure of fit termed the comparison curve fit index (CCFI; Ruscio et al. 2006, Formula 7.4, p. 188) was also calculated. This index takes into account the residuals between the plotted points produced by the research data and those produced by the simulated dimensional data as well as the simulated taxonic data. This information is then combined to form the CCFI whose values range from zero to one, with a score of 0 signifying a dimensional structure, a score of 1 supporting a taxonic structure, and scores near .50 indicating an ambiguous structure. For the ICG-R Paired Indicators MAMBAC analysis, the CCFI was .26, which offered further evidence of dimensionality.

Finally, each of the five MAMBAC curves also provided an estimate of the taxon base rate, which is approximated using a formula (Meehl and Yonce 1994, Appendix A) based on the mean differences found at the first and last cuts of the curve. It has been argued that if a construct is taxonic it should provide stable base rate estimates across analyses. However, if it is dimensional then the base rate

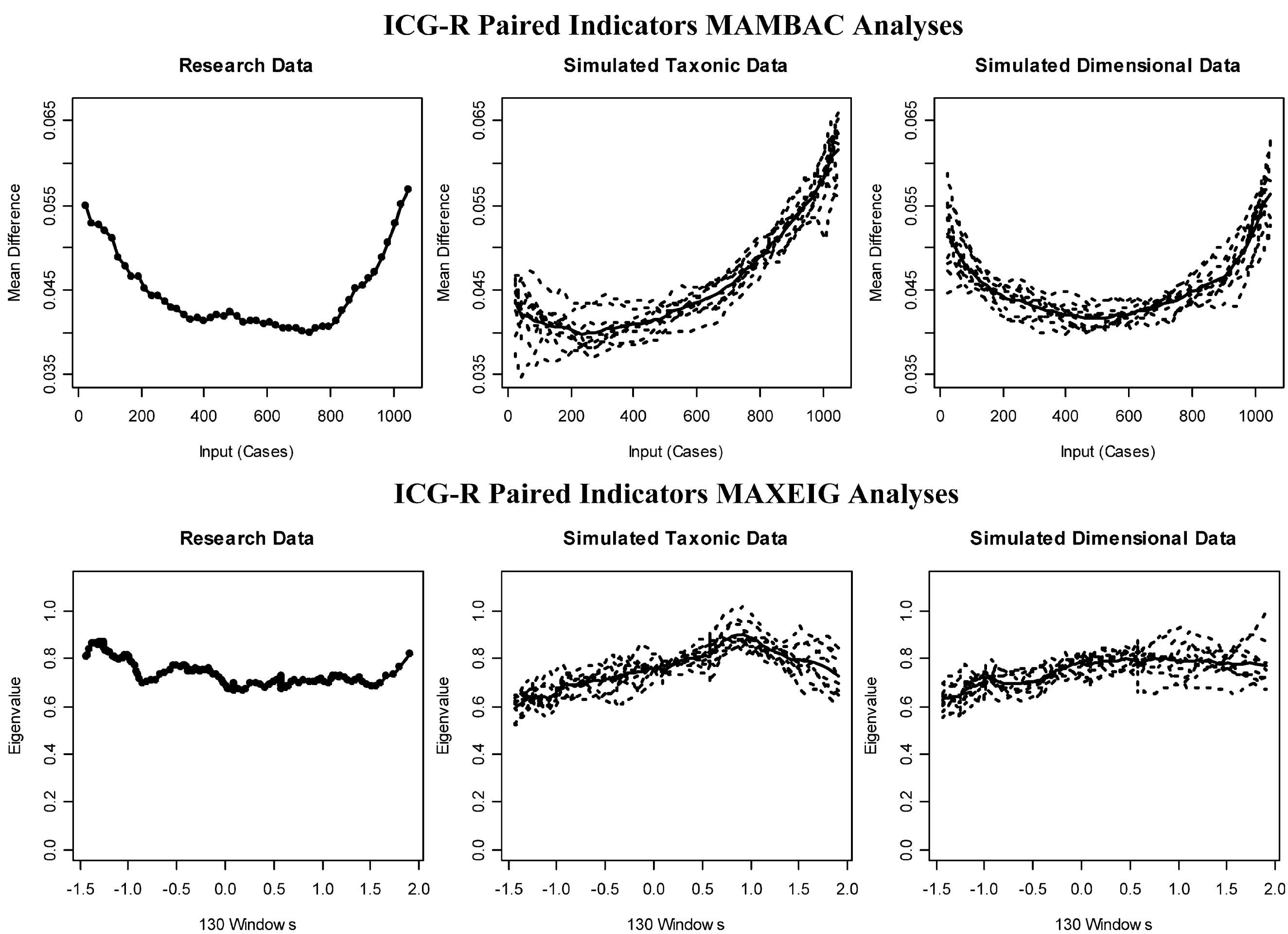

Fig. 1 In these panels, averaged curves are presented that summarize the results obtained using the ICG-R Paired Indicators in both the MAMBAC and MAXEIG analyses. The curves obtained using the simulated taxonic and dimensional data sets are also shown as a point of comparison. For each simulated latent structure, 10 samples were drawn in order to provide a sampling distribution of results. The broken lines represent the curves derived from each of the 10 simulated samples, and the solid line represents an average of these curves

estimates will tend to be more variable, indicating that no single entity is being reliably detected. It has been proposed that a standard deviation of the base rates that is less than .10 should be taken as evidence supporting a taxonic structure with one notable exception. Specifically, U-shaped MAMBAC curves with a consistent base rate of about 50% are considered to be indicative of a dimensional structure (Schmidt et al. 2004). It should be noted, though, that the notion of using base rate consistency to support conclusions about latent structure has been called into question (Ruscio 2007; Ruscio et al. 2006; Ruscio et al. 2007). Specifically, in a recent simulation study the CCFI was shown to be a more accurate way of discovering the latent structure of a construct compared to other techniques, including this base rate consistency method (Ruscio et al., 2007). In the present analyses the average base rate was found to be very close to 50% (*M*=.49) and was fairly consistent across each individual analysis (*SD*=.04), supporting a dimensional interpretation according to Schmidt et al. (2004) criterion.

In addition to these MAMBAC analyses, a MAXEIG procedure was also used with the ICG-R Paired Indicators. Five MAXEIG curves were produced, each of which used one indicator as the input and the remaining four as the output variables in an alternating fashion. Each of these five curves was relatively straight and did not possess a distinct peak, supporting a dimensional structure. Lines representing the average of these MAXEIG curves are also presented in Fig. 1 next to the curves resulting from the simulated dimensional and taxonic data. Generally, this average curve more closely resembled the simulated dimensional data, which was corroborated to some extent by a CCFI of .45. However, it is worth noting that this CCFI value of .45 is larger than the CCFI value found in the MAMBAC analyses. This more ambiguous CCFI value could be due to the peaked cusp at the far left side of the curve, which does not map onto either the simulated taxonic or dimensional data.

Base rates were also estimated for the MAXEIG analyses by calculating the proportion of individuals beyond the midpoint of the window with the largest eigenvalue (Waller and Meehl 1998). The average of these base rate estimates was .58 with a standard deviation of .15, which according to Schmidt et al. (2004) would be inconsistent with a taxonic structure.

PGD Criteria Indicators

Each of the five MAMBAC curves produced by the PGD Criteria Indicators appeared U-shaped, which is consistent with a dimensional structure. Likewise, the average of these MAMBAC curves more closely resembled the simulated dimensional data, as can be seen in Fig. 2. The dimensionality of the data is further supported by a CCFI of .32. In addition, the base rate estimates across these five curves were consistently very close to 50% (*M*=.52, *SD*=.04), which according to Schmidt et al. (2004) would support a dimensional structure.

The PGD Criteria Indicators were also used to create five MAXEIG curves, which used one indicator as the input and the remaining four as the output. The resulting curves were not shaped in a consistent fashion; hence, the full panel of curves is presented in Fig. 3. Three of these five graphs produced relatively straight lines consistent with a dimensional structure. Conversely, the curve that used Indicator 5 (which tapped into difficulties accepting the loss) as the input and the remaining four indicators as the output yielded two peaked areas—one on the far left side of the graph (suggestive of a rather large taxon) and another on the right half of the graph (suggestive of a small taxon). It should also be noted that an ambiguous structure was found for the curve that used Indicator 3 (which measured symptoms of shock and disbelief) as the input. In particular, a peak was observed on the far left side of the graph, which could be indicative of a taxonic structure. However, a peak in this area of the graph would be highly inconsistent with a relatively small latent category (e.g., 17.8%), as was predicted for PGD.

The averaged curve for these MAXEIG analyses appears in Fig. 2 next to the results obtained using the simulated taxonic and dimensional data. Interestingly, this averaged curve is peaked at the far left end of the graph, which does not map onto the graphs produced by either the simulated taxonic or dimensional data. It should be noted, however, that the right half of the averaged graph for the research data appears to be fairly straight and consistent with a latent dimension. Likewise, a CCFI of .41 was obtained, which also provides modest support for a dimensional structure. However, the average of the base rate estimates across all MAXEIG curves was .74 with a standard deviation of .06. According to Schmidt et al. (2004), base rate consistency at this level might suggest the presence of a large latent category.

Given this unexpected finding, six additional MAXEIG analyses were performed on the PGD Criteria Indicators using varying numbers of windows, ranging from 30 to 180 at intervals of 30 (e.g., 30, 60, 90), as another consistency test to see if a large latent taxon could still be observed across increasing numbers of windows. Interestingly, left-ended peaks were observed in the averaged curves regardless of the number of windows used. Furthermore, these analyses converged on a similar average taxon base rate (*M*=.71—.74), and in each analysis the standard deviation of the base rates were at or below .10 (*SD*=.04—.10), offering additional support for the presence of a large latent taxon.

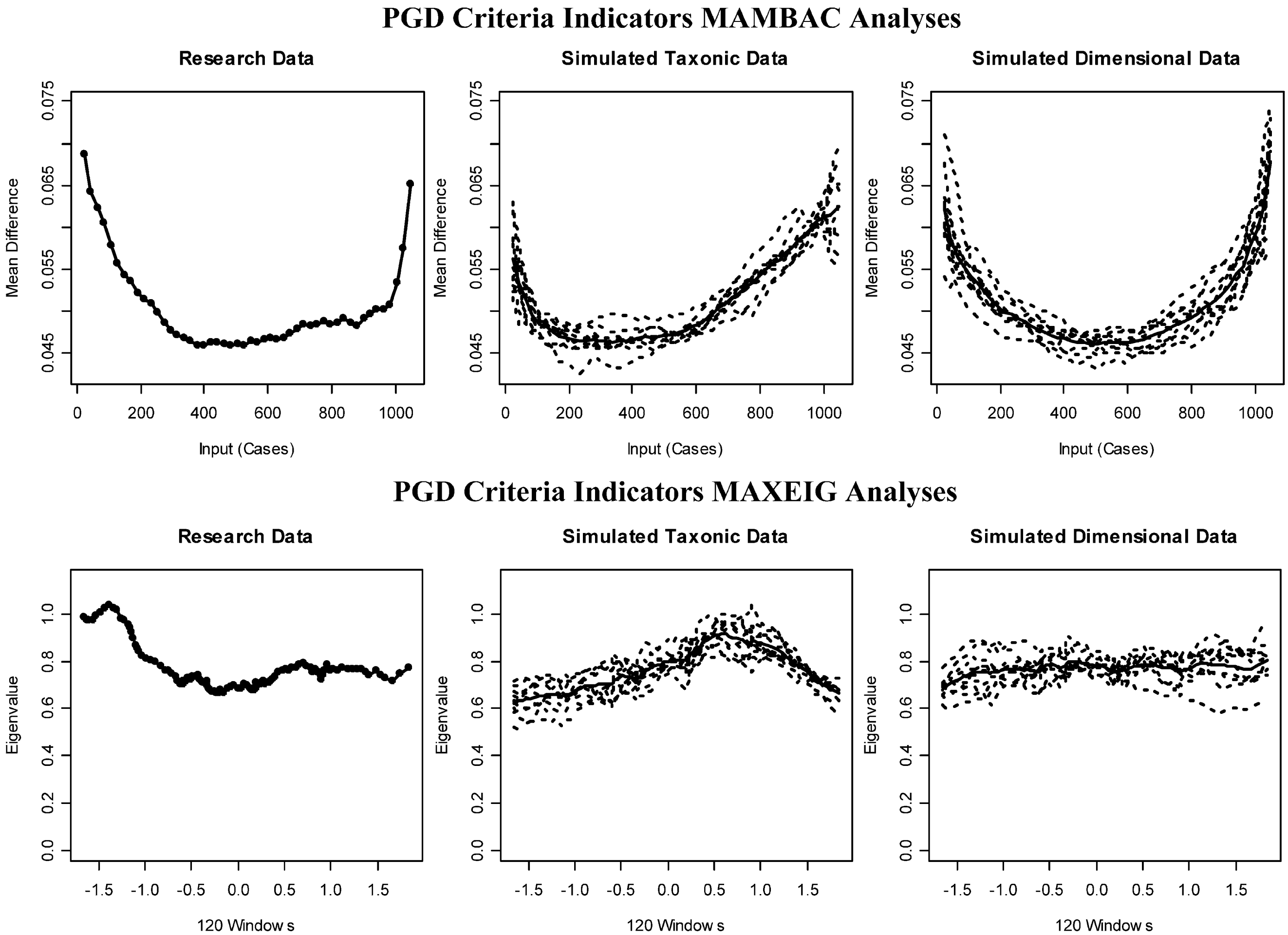

Fig. 2 In these panels, averaged curves are presented that summarize the results obtained using the PGD Criteria Indicators in both the MAMBAC and MAXEIG analyses. The curves obtained using the simulated taxonic and dimensional data sets are also shown as a point of comparison. For each simulated latent structure, 10 samples were drawn in order to provide a sampling distribution of results. The broken lines represent the curves derived from each of the 10 simulated samples, and the solid line represents an average of these curves

Discussion

As a whole, these taxometric analyses offered little support for a categorical conceptualization of normal grief and PGD. Specifically, 18 of the 20 taxometric curves were consistent with a dimensional structure. Likewise, all CCFI values were below .50, lending further support to the view that normal and prolonged grief simply fall along different ends of the same continuum. From a research perspective, this finding suggests that it would make sense for researchers to measure PGD using continuous instruments. By capturing the full spectrum of grief symptoms, researchers can measure grief more sensitively and thereby increase statistical power. In contrast, dichotomizing the construct of grief into normal reactions and severe and prolonged reactions would not be advised in most research situations, as this practice would likely discard meaningful variability. Indeed, if normal and pathological reactions to loss are not qualitatively distinct, then it would seem that the variability among normal grievers and/or sub-threshold cases of PGD might still offer valuable information about more severe grief reactions, as these cases simply represent a different point on the same continuum.

Beyond these implications for research, it should also be emphasized that these findings in no way vitiate the importance of PGD as a diagnostic descriptor in clinical practice. As DSM-IV states, "there is no assumption that each category of mental disorder is a completely discrete entity with absolute boundaries dividing it from other mental disorders or from no mental disorder" (American Psychiatric Association 2000, p. xxxi). Therefore, rather than challenging the merits of PGD as a clinical diagnosis, this research suggests that the clinical marker for PGD may be defined best by the severity of grief symptoms (persisting for 6 months or longer) rather than the presence or absence of a qualitatively distinct set of pathological symptoms. Given this conceptualization of PGD, greater understanding of this construct could be gained by

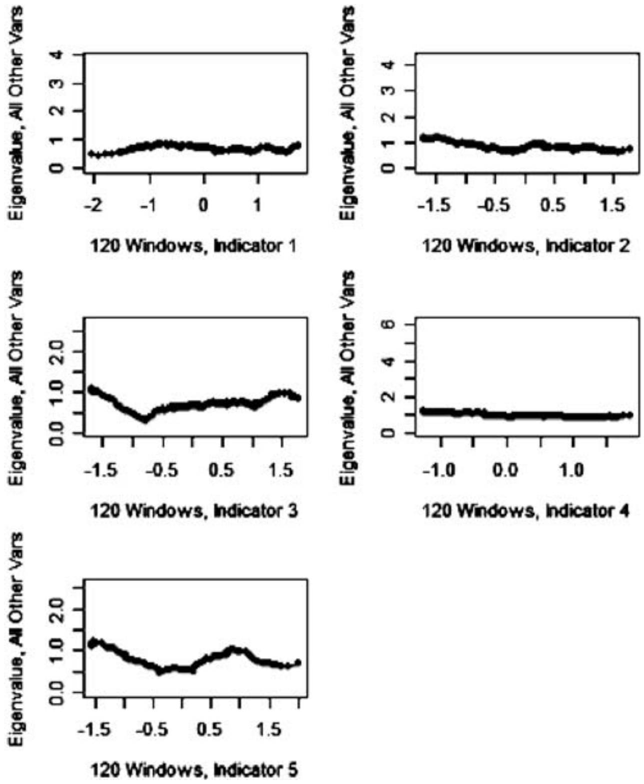

Fig. 3 These graphs show the MAXEIG analyses for the PDG Criteria Indicators. In these graphs cases were aligned along the x-axis according to their scores on a single input indicator, and then 120 consecutive subsamples (i.e., windows) were created that overlapped with neighbouring subsamples by 90%. For each of these windows, an eigenvalue was calculated for the remaining output indicators and was plotted on the y-axis

examining the boundary between normal and prolonged grief as a threshold at which a clinically relevant group of external criteria are met (Kessler 2002). Though future work would do well to continue to refine this boundary, past research suggests that the present threshold set by PGD diagnostic criteria has great practical utility as an indicator of problematic grieving (e.g., Latham and Prigerson 2004; Prigerson et al. 1997; Prigerson et al. 2008). For instance, Prigerson et al. (2008) found that those who met the current diagnostic cutoff for PGD at 6—12 months after a loss had a 2.6 fold increased risk of alcohol use, a 3.4 fold increased risk of smoking, and a 3.1 fold increased risk of sleep problems compared to bereaved individuals not meeting diagnostic criteria at 12—24 months after a loss.

In addition to this work, researchers would also do well to examine a PGD threshold in terms of treatment outcome. Reviews of the grief intervention literature suggest that studies that include more severely distressed grievers tend to produce larger effects (Currier et al. 2007, Currier et al. 2008b; Schut et al. 2001), but it remains unclear at what point clients benefit most from treatment. Thus, future research that locates a “region of significance” along the continuum of grief symptoms where clients are most likely to exhibit meaningful, treatment-induced changes could inform diagnostic decision-making and appropriately match clients to therapy.

Although these analyses did not support the existence of a taxonic boundary between PGD and normal grief, it is noteworthy that the MAXEIG analyses often produced peaks on the far left side of the graph. Indeed, in the initial 120-window analysis of the ICG-R Criteria Indicators, base rate estimates consistently clustered around 65—80%. This pattern of results was unexpected (i.e., PGD would presumably be observed as a small taxon) and suggests the possible existence of a different taxonic boundary outside of PGD and normal grief. Though this finding is subject to a variety of interpretations, one potential explanation is that these left-ended peaks represent a taxonic boundary between those participants who were actively grieving (i.e., roughly 65–80% of the sample) and those who were no longer in a state of grief (i.e., the remaining 20–35% of the sample). Interestingly, these estimated base rates are somewhat smaller than (yet still comparable to) those of another study that examined the longitudinal pathways through the grief process. In this study, it was found that 45.9% of bereaved spouses exhibited a “resilient” trajectory, in which minimal grief and depressive symptoms were observed 6 and 18 months after a loss among previously non-depressed individuals (Bonanno et al. 2002).

Of course, the cross-sectional design of the present study prohibits claims about long term trajectories. Additionally, if there is in fact a taxonic boundary between active grievers and inactive grievers, as possibly suggested by one set of MAXEIG analyses, future work that explicitly aims to address this somewhat different research question would provide an important expansion on the present study. Such a project would presumably use different base rate estimates and sets of indicators that are specifically tailored to identify this taxonic boundary. It is also possible that other categories of grief might exist that were not captured by the indicators used in this study. For example, future studies could examine proposed subtypes of grief such as *masked grief*, in which a bereaved individual experiences symptoms that lead to impairment but does not consciously identify these problems as resulting from the loss (Worden 1982).

More generally, it is noteworthy that the findings of this study were based on a self-report measure, and it is possible that different methodologies could yield different results. Indeed, a diagnosis of PGD is often made using a structured interview (Prigerson and Jacobs 2001), and this type of information could serve as a useful supplement to the self-reported questionnaire data used in this study. A multi-method approach of this kind could reduce the statistical "noise" (i.e., covariation among indicators due to shared methods) in a taxometric analysis, thereby increasing its power. Thus, the findings of this investigation should be tempered by this limitation until future studies are able to corroborate these results using a diversity of methods, including grief measures based on interview data—although accomplishing this for the number of participants required by a taxometric analysis would pose obvious practical challenges.

These results were also based on information provided by an exclusively Dutch sample, and any future replications would do well to examine the latent structure of grief among ethnically diverse samples. This sample was also younger (i.e., $M=38.5$ years, $SD=11.0$) than the sample of older bereaved spouses used in the initial empirical tests of PGD, and replication across different age groups is certainly warranted (Prigerson and Jacobs 2001). Likewise, the fact that this study surveyed a self-selected, Internet sample might have also influenced its external validity (e.g., Internet-users may tend to have higher educational levels compared to the general population), highlighting further the need for additional investigation of this issue using more representative samples. It is also possible that Internet recruitment/administration increased the likelihood of certain response biases. Although efforts were made to remove unreliable data, it was not possible to fully screen for factors such as inattentiveness, social desirability, or other barriers to thoughtful and honest responding. Thus, replication of these findings with data gathered using alternate recruitment and administration procedures (e.g., clinician administered measures collected in a clinic or laboratory setting) would help shed light on the potential impact of these limitations. Despite these limitations, this study represents the first taxometric investigation of grief and provides support for a continuous conceptualization of PGD and normal reactions to loss—a finding that has important implications for research and treatment.

References

Allumbaugh, D. L., & Hoyt, W. T. (1999). Effectiveness of grief therapy: a meta-analysis. *Journal of Counseling Psychology*, *46*, 370–380.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4thth ed.). Washington, DC: Author.

Barry, L. C., Kasl, S. V., & Prigerson, H. G. (2002). Psychiatric disorders among bereaved persons: the role of perceived circumstances of death and preparedness for death. *American Journal of Geriatric Psychiatry*, *10*, 447–457.

Boelen, P. A., & Lensvelt-Mulders, G. J. L. M. (2005). Psychometric properties of the grief cognitions questionnaire (GCQ). *Journal of Psychopathology and Behavioral Assessment*, *27*, 291–303.

Boelen, P. A., & van den Bout, J. (2005). Complicated grief, depression, and anxiety as distinct postloss syndromes: a confirmatory factor analysis study. *American Journal of Psychiatry*, *162*, 2175–2177.

Boelen, P. A., & van den Bout, J. (2008). Complicated grief and uncomplicated grief are distinguishable constructs. *Psychiatry Research*, *157*, 311–314.

Boelen, P. A., de Keijser, J., & van den Bout, J. (2001). Psychometrische eigenschappen van de Rouw VragenLijst (RVL) [psychometric properties of the inventory of traumatic grief]. *Gedrag & Gezondheid*, *29*, 172–186.

Boelen, P. A., van den Bout, J., & de Keijser, J. (2003a). Traumatic grief as a disorder distinct from bereavement-related depression and anxiety: a replication study with bereaved mental health care patients. *American Journal of Psychiatry*, *160*, 1339–1341.

Boelen, P. A., van den Bout, J., de Keijser, J., & Hoijtink, H. (2003b). Reliability and validity of the Dutch version of the inventory of traumatic grief (ITG). *Death Studies*, *27*, 227–247.

Boerner, K., Wortman, C. B., & Bonanno, G. A. (2005). Resilient or at risk? A 4-year study of older adults who initially showed high or low distress following conjugal loss. *Journal of Gerontology*, *60B*, P67–P73.

Bonanno, G. A., & Kaltman, S. (2001). The varieties of grief experience. *Clinical Psychology Review*, *21*, 705–734.

Bonanno, G. A., Wortman, C. B., Lehman, D. R., Tweed, R. G., Haring, M., Sonnega, J., et al. (2002). Resilience to loss and chronic grief. *Journal of Personality and Social Psychology*, *83*, 1150–1164.

Bonanno, G. A., Neria, Y., Mancini, A., Coifman, K. G., Litz, B., & Insel, B. (2007). Is there more to grief than depression and PTSD? A test of incremental validity. *Journal of Abnormal Psychology*, *116*, 342–351.

Bowlby, J. (1980). *Loss: Sadness and depression (attachment and loss)* (vol. Vol. 3). New York: Basic Books.

Burnett, P., Middleton, W., Raphael, B., & Martinek, N. (1997). Measuring core bereavement phenomena. *Psychological Medicine*, *27*, 49–57.

Chen, J. H., Bierhals, A. J., Prigerson, H. G., Kasl, S. V., Mazure, C. M., & Jacobs, S. (1999). Gender differences in the effects of bereavement-related distress in health outcomes. *Psychological Medicine*, *29*, 367–380.

Currier, J. M., Holland, J. M., & Neimeyer, 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*, 1–7.

Currier, J. M., Holland, J. M., Coleman, R. A., & Neimeyer, R. A. (2008a). Bereavement following violent death: An assault on life and meaning. In R. Stevenson, & G. Cox (Eds.), *Perspectives on violence and violent death* (pp. 177–202). Amityville, NY: Baywood.

Currier, J. M., Neimeyer, R. A., & Berman, J. S. (2008b). The effectiveness of psychotherapeutic interventions for bereaved individuals: a comprehensive quantitative review. *Psychological Bulletin*, *134*, 648–661.

Faschingbauer, T. R. (1981). *Texas Revised Inventory of Grief manual*. Houston, TX: Honeycomb.

Goodkin, K., Lee, D., Molina, R., Zheng, W., Frasca, A., O'Mellan, S., et al. (2006). Complicated bereavement: disease state or state of being? *Omega*, *52*, 21–36.

Grove, W. M. (2004). The MAXSLOPE taxometric procedure: mathematical derivation, parameter estimation, consistency tests. *Psychological Reports*, *95*, 517–550.

Hogan, N. S., Greenfield, D. B., & Schmidt, L. A. (2001). Development and validation of the Hogan Grief Reaction Checklist. *Death Studies*, *25*, 1–31.

Hogan, N. S., Worden, J. W., & Schmidt, L. A. (2004). An empirical study of the proposed complicated grief disorder criteria. *Omega*, *48*, 263–277.

Kato, P. M., & Mann, T. (1999). A synthesis of psychological interventions for the bereaved. *Clinical Psychology Review*, *19*, 275–296.

Kessler, R. C. (2002). Epidemiological perspectives for the development of future diagnostic systems. *Psychopathology*, *35*, 158–161.

Latham, A. E., & Prigerson, H. G. (2004). Suicidality and bereavement: complicated grief as psychiatric disorder presenting greatest risk for suicidality. *Suicide and Life Threatening Behavior*, *34*, 350–362.

Lichtenthal, W. G., Cruess, D. G., & Prigerson, H. G. (2004). A case for establishing complicated grief as a distinct mental disorder in DSM-V. *Clinical Psychology Review*, *24*, 637–662.

Meehl, P. E. (1995). Bootstraps taxometrics: solving the classification problem in psychopathology. *American Psychologist*, *50*, 266–274.

Meehl, P. E., & Yonce, L. J. (1994). Taxometric analysis: I. Detecting taxonicity with two quantitative indicators using means above and below a sliding cut (MAMBAC procedure). *Psychological Reports*, *74*, 1059–1274.

Melhem, N. M., Rosales, C., Karageorge, J., Reynolds, C. F., Frank, E., & Shear, M. K. (2001). Comorbidity of axis I disorders in patients with traumatic grief. *Journal of Clinical Psychiatry*, *62*, 884–887.

Middleton, W., Burnett, P., Raphael, B., & Martinek, N. (1996). The bereavement response: a cluster analysis. *British Journal of Psychiatry*, *169*, 167–171.

Ott, C. H. (2003). The impact of complicated grief on mental and physical health at various points in the bereavement process. *Death Studies*, *27*, 249–272.

Prigerson, H. G., & Jacobs, S. C. (2001). Diagnostic criteria for traumatic grief: 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: Consequences, coping, and care* (pp. 613–645). Washington, DC: American Psychological Association.

Prigerson, H. G., & Maciejewski, P. K. (2006). A call for sound empirical testing and evaluation of criteria for complicated grief proposed for DSM-V. *Omega*, *52*, 9–19.

Prigerson, H. G., Frank, E., Kasl, S. V., Reynolds, C. F., Anderson, B., Zubenko, G. S., Houck, P. R., et al. (1995a). Complicated grief and bereavement-related depression as distinct disorders: preliminary empirical validation in elderly bereaved spouses. *American Journal of Psychiatry*, *152*, 22–30.

Prigerson, H. G., Maciejewski, P., Reynolds, C., Bierhals, A., Newsom, J., Fasiczka, A., et al. (1995b). Inventory of complicated grief. *Psychiatry Research*, *59*, 65–79.

Prigerson, H. G., Bierhals, A. J., Kasl, S. V., Reynolds, C. F., Shear, M. K., Newsom, J. T., et al. (1996). Complicated grief as a disorder distinct from bereavement-related depression and anxiety: a replication study. *American Journal of Psychiatry*, *153*, 1484–1486.

Prigerson, H. G., Bierhals, A. J., Kasl, S. V., Reynolds, C. F., Shear, M. K., Day, N., et al. (1997). Traumatic grief as a risk factor for mental and physical morbidity. *American Journal of Psychiatry*, *154*, 616–623.

Prigerson, H. G., Vanderwerker, L. C., & Maciejewski, P. K. (2008). A case for inclusion of prolonged grief disorder in DSM-V. In M. Stroebe, R. Hansson, H. Schut, & W. Stroebe (Eds.), *Handbook of bereavement research and practice: Advances in theory and intervention* (pp. 165–186). Washington, DC: American Psychological Association.

Ruscio, J. (2007). Taxometric analysis: an empirically grounded approach to implementing the method. *Criminal Justice and Behavior*, *34*, 1588–1622.

Ruscio, J., & Ruscio, A. M. (2004). Clarifying boundary issues in psychopathology: the role of taxometrics in a comprehensive program of structural research. *Journal of Abnormal Psychology*, *113*, 24–38.

Ruscio, J., Haslam, N., & Ruscio, A. M. (2006). *Introduction to the taxometric method: A practical guide*. Mahwah, NJ: Erlbaum.

Ruscio, J., Ruscio, A. M., & Meron, M. (2007). Applying the bootstrap to taxometric analysis: generating empirical sampling distributions to help interpret results. *Multivariate Behavioral Research*, *42*, 349–386.

Schmidt, N. B., Kotov, R., & Joiner, T. E. (2004). *Taxometrics: Toward a new diagnostic scheme for psychopathology*. Washington, DC: American Psychological Association.

Schut, H., Stroebe, M. S., van den 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. 705–737). Washington, DC: American Psychological Association.

Waller, N. G., & Meehl, P. E. (1998). *Multivariate taxometric procedures: Distinguishing types from continua*. Thousand Oaks, CA: Sage.

Worden, J. W. (1982). *Grief counseling and grief therapy: A handbook for the mental health practitioner*. New York: Springer.