Investigating the Reactivity Effect of Experience Sampling Methodology on Prolonged Grief Disorder Symptom Levels in Recently Bereaved Adults: A Randomised Controlled Trial

Nicky M. E. L. Timmerman

Erasmus School of Social and Behavioural Sciences, Erasmus University Rotterdam

4.5 Master Thesis PSY

Dr. Minita Franzen

March 2024

Abstract

Recently included in diagnostic manuals, Prolonged Grief Disorder (PGD) can be described as elevated and persistent mental distress as a result of loss leading to significant functional impairment. The importance of early interventions has been highlighted and a potential reactivity effect of Experience Sampling Methodology (ESM) on symptomatology through self-monitoring has been proposed in post-traumatic stress disorder and PGD studies. However, there is a notable absence of Randomised Controlled Trials (RCTs) using ESM in PGD literature. The current study is part of a larger RCT and examined whether using ESM to record the severity of grief reactions five times daily for 2 weeks could lead to a reduction in PGD symptom levels. Additionally, this study explored if gender moderated the relationship between ESM and PGD symptoms. In total, 158 individuals who had lost a loved one 3-to-6 months prior were randomly allocated to a waitlist control, or ESM condition. PGD symptoms were assessed in telephone interviews before and after the 2 weeks using the Traumatic Grief Inventory-Clinical Administered. Results revealed that although participants demonstrated significant improvement over time, the ESM condition did not show significantly greater changes in PGD symptom levels than the control condition. Moreover, the effect of time and condition on PGD scores did not significantly differ depending on gender. Findings suggest that ESM alone may not be more effective in reducing PGD symptoms over time than being waitlisted. However, ESM can still serve as a valuable tool for treatment monitoring purposes and personalisation of treatment.

Keywords: Prolonged Grief Disorder (PGD), Experience Sampling Method (ESM), Randomised Controlled Trial (RCT), gender, recently bereaved individuals

Introduction

Losing a loved one has been found to be one of the greatest life stressors faced by humans (Holmes & Rahe, 1967). There is no universal response to loss, instead, it varies largely between individuals (Bonanno, 2004). Many people experience moderate to high levels of grief directly after the loss, which tends to decrease over time, and no treatment is required (Djelantik et al., 2017; Nielsen et al., 2019). However, approximately 7-10% of bereaved individuals experience high levels of grief over a longer period of time (Lundorff et al., 2017), leading to functional impairment (Jordan & Litz, 2014). This type of grief reaction can be classified as Prolonged Grief Disorder (PGD) (American Psychiatric Association, 2022; World Health Organization, 2022).

PGD can be described as elevated and persistent mental distress as a result of loss (Szuhany et al., 2021). This is characterised by an intense longing and yearning for, and/or a persistent preoccupation with thoughts or memories of the deceased. Other symptoms of PGD include identity disruption, disbelief about the death, avoidance of reminders, emotional numbness, intense emotional pain, difficulty re-engaging in relationships and activities, feeling a sense of meaninglessness, and severe loneliness (American Psychiatric Association, 2022). Although distinct constructs (Djelantik et al., 2017; Lenferink et al., 2020; Wen et al., 2022), PGD is related to post-traumatic stress disorder (PTSD) and depression as symptoms partially overlap (Komischke-Konnerup et al., 2021).

Various factors increasing the likelihood of PGD caseness have been identified such as having inadequate social support (Szuhany et al., 2021; Zisook et al., 2014) and shorter time spans since loss (Doering et al., 2022). For example, Boelen and Lenferink (2022) found that individuals who met probable PGD criteria 4 months after being bereaved were 32 times more likely to meet probable PGD criteria after 1 year compared to those who did not have an indication of PGD 4 months post-loss. Thus, as severe grief reactions that will not subside over time are detectable early on (Doering et al., 2022; Kersting et al., 2011; Prigerson et al., 2009), the first 3-to-6 months following the death of a loved one seem to be especially relevant for early interventions.

In terms of previous research, the majority of grief studies have chosen to asses PGD symptoms retrospectively (Treml et al., 2020). However, a different method of collecting data on psychological symptoms, which has been shown to have high ecological validity (van Os et al., 2017), is Experience Sampling Methodology (ESM) (Hektner et al., 2007). ESM is a momentary assessment technique designed for intensive data collection in daily life (Hektner et al., 2007). Generally, participants are required to complete short questionnaires at several time points throughout the day regarding current mood, cognitions, behaviours, and contextual information in response to beep prompts for a set period of time (Verhagen et al., 2016).

While ESM has been employed for data collection in a variety of contexts and populations within mental health literature (Myin-Germeys et al., 2018), with evidence supporting its effectiveness when incorporated into depression interventions (Kauer et al., 2012; Kramer et al., 2014; Suen et al., 2022), limited research has explored the potential reactivity effect of repeatedly assessing symptoms through ESM on symptom levels. However, preliminary findings indicate a reactivity effect, as studies have demonstrated improvement in assessed variables following a period of ESM (McDevitt-Murphy et al., 2018). For instance, an ESM study with combat veterans assessing PTSD symptoms, negative affect, coping, and alcohol use 4 times a day found that the severity of PTSD was significantly lower post-ESM compared to PTSD symptoms before the ESM period (Possemato et al., 2015). On the other hand, an exploratory ESM study using single items to assess affective and cognitive states related to depression did not find clinical improvement in symptomatology amongst individuals diagnosed with Major Depressive Disorder (Vachon et al., 2016). Yet, it is important to point out that participants in this study were required to self-assess their depression symptoms two times a day over a period of 5 months, whereas the majority of ESM research requires participants to complete assessments more than twice a day over a shorter period of time. Therefore, the difference in study design potentially influenced the results which may explain the contrasting findings.

However, despite preliminary findings from PTSD research that assessing psychological symptoms through ESM may contribute to the improvement of symptomatology, there has been limited research exploring the effect of ESM on PGD symptom levels. To my knowledge, the only ESM study within grief literature was conducted recently by Lenferink, van Eersel, and Franzen (2022). In this pilot study, participants were asked to respond to items assessing PGD symptoms through a smartphone application five times a day during a 2-week period. Although not the main aim of the study, results revealed that the bereaved individuals reported significantly lower levels of PGD following the 2-week ESM phase, compared to 2 weeks prior. The authors proposed that ESM may serve as a way of self-monitoring symptoms, which in turn might have encouraged change (van Os et al., 2017). Similarly to self-monitoring, ESM encourages individuals to record their thoughts, feelings, and behaviours in relation to PGD whilst going about their day-to-day lives. In doing so, ESM promotes participants' awareness of cognitive and emotional processes (Telford et al., 2012) and contributes to increased awareness of behavioural patterns (Myin-Germeys et al., 2018; Snippe et al., 2016) much like self-monitoring would (van Os et al., 2017). However, it is crucial to point out that this study was not a controlled trial, therefore when examining potential reactivity effects it cannot be ruled out that changes in PGD levels occurred due to the passing of time (Lenferink, van Eersel, & Franzen, 2022).

Since the body of research that uses ESM to assess PGD symptoms is extremely limited and no controlled trials have been conducted to date, it is imperative that a Randomised Controlled Trial (RCT) is carried out to establish if the reactivity effect found by Lenferink, van Eersel, and Franzen (2022) can be attributed to participating in an ESM phase repeatedly assessing PGD symptoms. Overall, considering Lenferink, van Eersel, and Franzen's (2022) findings and because evidence from PTSD literature indicates that ESM potentially improves symptom levels (Possemato et al., 2015), there might also be a reactivity effect when using ESM to assess PGD symptoms in daily life. This is why, in the present study, I expected to find that participating in an ESM phase would lead to a significant decrease in PGD symptom levels compared to being waitlisted.

The Potential Moderating Role of Gender on the Reactivity Effect of ESM on PGD Symptoms

The effect of ESM on grief however might not be straightforward but rather quantified by other variables such as gender[1]. Previous research tends to view gender as a binary construct (woman and man) and indicates that gender influences grief reactions (Burke et al., 2019). Studies suggest that gender-specific processes and social norms can influence how people express grief symptoms (Creighton et al., 2013; Stelzer et al., 2019) therefore, women and men may also respond differently to assessing PGD symptoms through ESM. If the potential reactivity effect of ESM on PGD symptom levels varies between women and men, distinct forms of support may be needed for each gender. Consequently, ESM interventions should be tailored accordingly to enhance effectiveness. However, to my knowledge, there have not been any studies investigating whether gender moderates the relationship between ESM and PGD symptoms.

During the ESM phase in the current study, participants were instructed to self-report 11 grief symptoms five times a day in response to smartphone beeps. The ESM items prompted individuals to reflect on their current emotional state when completing the assessments, therefore the ESM phase in this study could be considered emotion-focused. Interestingly, an older bereavement intervention study found that bereaved men benefitted more from an emotion-focused coping intervention whereas an action-focused coping intervention was more beneficial for women who were grieving (Schut et al., 1997). Even though generally speaking, men tend to use more action-focused coping whilst women are more likely to use emotion-focused coping strategies (Jones et al., 2019; Ptacek et al., 1992), the authors argued that the men and women in their study may have benefitted from less familiar coping strategies since their go-to coping strategies did not seem to be effective. Because research suggests that individuals employ different coping strategies for grief based on gender, and considering that emotion-focused interventions may yield greater benefits for

[1] Gender has been defined as a socially constructed phenomenon that is not the same as biological sex (World Health Organization, n.d.). Gender expression can vary between societies and can change over time. Sex, on the other hand, interacts with gender but refers to the biological and physiological characteristics of females, males, and intersex persons such as reproductive organs, hormones, and chromosomes (World Health Organization, n.d.). Throughout this text, gender will be referred to as self-identified gender.

men than women, it is anticipated that ESM will lead to a more pronounced reduction in PGD symptom levels among men compared to women.

Present Study

The current RCT aimed to investigate whether participating in a 2-week ESM phase influenced recently bereaved individuals' PGD symptoms. In doing so, participants were randomly allocated to an ESM or control condition and PGD symptoms were assessed in a baseline telephone interview (T1). Subsequently, participants in the ESM condition commenced with a 2-week ESM phase whereas those in the control condition were waitlisted for 2 weeks. During the ESM period, participants responded to ESM items assessing PGD symptoms and contextual information. After 2 weeks, symptoms were assessed again amongst all participants (T1b for the control condition and T2 for the ESM condition).

Based on findings from PTSD and PGD literature, I expected to find that taking part in an ESM phase assessing PGD symptoms multiple times a day led to a significant decrease in PGD symptoms compared to being waitlisted (H1). Furthermore, although no previous studies have explored moderators of the potential reactivity effect of ESM on PGD symptom levels, research findings on emotion-focused interventions led me to predict that gender moderated the relationship between ESM and PGD symptoms and that PGD symptoms decreased more for men than for women within the ESM condition (H2).

Methods

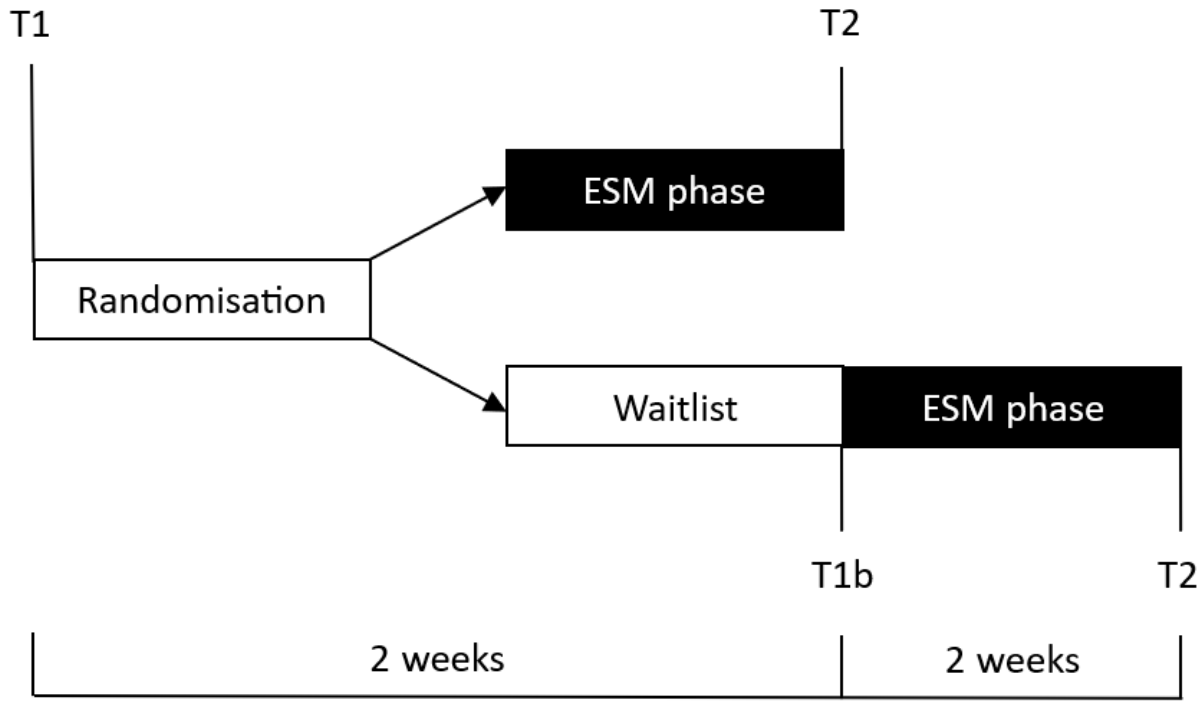
Design

The current RCT is part of the Grief in Daily Life project that intends to examine PGD symptoms in everyday life. Upon enrolling in the study, participants were allocated to the ESM condition or the control condition using computer-generated randomisation. All participants were interviewed at the beginning of the study (T1) and upon completion of the ESM phase (T2), as shown in Figure 1. After the baseline interview, participants in the ESM condition immediately commenced with the 2-week ESM phase, whereas individuals in the control condition were waitlisted for 2 weeks before starting

the ESM phase. After the 2 weeks of ESM or waitlist, a second interview was scheduled with all participants (T2/T1b). Subsequently, the participants in the waitlist condition had a final interview upon completion of the ESM phase (T2).

Figure 1

Study Design



Participants

Adults who lost a partner, family member, or friend in the past 3-to-6 months were able to enrol in the study. The bereaved individuals needed to speak Dutch fluently and have access to a smartphone to be eligible to participate. In line with previous research (Lenferink, van Eersel, & Franzen, 2022), participants who had been diagnosed with a psychotic disorder or were acutely suicidal were excluded as assessed with single items at baseline. Participants were recruited through convenience and snowball sampling using an online grief monitoring tool available at https://rouwbehandeling.nl/. This tool was created by the principal investigator to help individuals

assess whether they required professional grief support. Data collection took place from February 2023 until August 2023 and there was no reward for participating.

In total, 623 potentially eligible participants were invited to take part of whom 187 were interviewed at baseline. Of those, four were excluded because they met the exclusion criteria and a further 25 participants in the ESM condition dropped out during the ESM phase. Thus, the total sample consisted of 158 participants, with 70 people in the ESM condition and 88 individuals in the control condition. Table 1 contains an overview of the background and loss-related characteristics. In sum, 84.8% of participants were women (*n* = 134) and 15.2% were men (*n* = 24). Although gender originally included three categories (woman, man, other), none of the participants selected 'other' therefore this category was discarded. The mean age was 53.5 (*SD* = 12.2) ranging from 22 to 84. The majority were college or university-educated (63.9%). In almost half of the cases, participants reported the death of a spouse/partner (46.8%) and the average age of the deceased was 62.6 (*SD* = 20.1) years old. Most participants lost their loved ones due to a natural cause (77.2%) and indicated that they experienced the loss as completely unexpected (33.5%). Besides this, just over half of participants had a history of support (53.2%) yet the majority did not receive support specifically related to grief (59.5%). At T1, 94 out of 158 participants (59.5%) met the criteria for probable PGD based on a cut-off score of 54 or higher. This score is based on findings from previous research using the TGI-SR (Boelen et al., 2019), as studies are still being undertaken to determine a cut-off score for the TGI-CA (Lenferink et al., 2019).

Procedure

Individuals who had indicated in the online grief monitoring tool that they were willing to participate in grief research, were invited to take part through email. A standardised format was used explaining what the research was about, what was required from participants, the expected benefits of participating, the inclusion criteria, the contact details of the principal investigator, and a link to an information sheet and consent form. After 1 week, a reminder email was sent out to increase engagement.

Once individuals signed the informed consent form, they were contacted by a member of the trained research team (all master-level psychology students) to schedule a baseline phone interview (T1) and were randomly allocated to the ESM condition or the control condition. The T1 interview took between 30 and 45 minutes in which PGD symptoms, as well as other variables not relevant to the current study (i.e., depression, PTSD, self-reflection and self-insight), were assessed using questionnaires. Participants' responses were given verbally and inputted in Qualtrics by the interviewers. The order of the questionnaires and their questions remained the same for each participant. At the end of the baseline interview, the next interview was scheduled (T1b or T2). Participants in the ESM condition were sent an email directly after the baseline interview containing instructions on how to download and use the smartphone application for the ESM phase. To collect ESM data the Ethica app was used (www.ethicadata.com). Alternatively, participants in the control condition were waitlisted for 2 weeks, after which a second interview was conducted (T1b), app instructions were sent, and the ESM phase was started. Approximately 1 to 2 days before all scheduled telephone interviews, standardised reminder emails were sent out to all participants to increase adherence.

During the ESM phase, participants were prompted to complete 11 ESM items and six contextual questions in the smartphone app by beeps at semi-random time intervals approximately every 3 hours (i.e., between 08.30 – 09.30, 11.30 – 12.30, 14.30 – 15.30, 17.30 – 18.30, and 20.30 – 21.30). Participants had 60 minutes to complete the questions and reminder notifications were sent 10 and 20 minutes after the window opened. To increase compliance, encouragement emails were sent out if participants missed more than 50% of the measurements on one day (i.e., $\geq$3 surveys). Moreover, upon completing each assessment participants were provided with a link (www.rouwbehandeling.nl) containing additional information about formal and informal grief support in order to mitigate the possible harms of participating in the study.

Within 7 days after completing the ESM phase, the final phone interview was conducted (T2). This interview took approximately 20 minutes and consisted of questions about participants'

experience using the app as well as the PGD, depression, PTSD, and self-reflection and self-insight questionnaires used in the baseline interview.

Measures

Background and Loss-Related Characteristics

During the baseline interview, various questions were asked regarding the background and loss circumstances of participants. Concerning background characteristics, participants were asked about their gender (0 = male, 1 = female, 2 = other), date and country of birth, level of education (0 = primary school, 1 = high school, 2 = vocational education, 3 = college/university), history of psychological support (0 = yes, 1 = no), grief support (0 = yes, 1 = no), and current support (0 = yes, 1 = no). In terms of loss circumstances, individuals were asked about the date of death, age of the deceased, relationship to the deceased (0 = partner, 1 = child, 2 = parent, 3 = sibling, 4 = grandparent, 5 = grandchild, 6 = friend, 7 = other), cause of death (0 = physical illness, 1 = accident, 2 = suicide, 3 = homicide/manslaughter, 4 = other, namely), and unexpectedness of death (1 = completely expected to 5 = completely unexpected).

Traumatic Grief Inventory-Clinician Administered (TGI-CA)

The TGI-CA (Lenferink et al., 2019) was used to assess PGD symptoms at T1, T1b, and T2. This questionnaire has been developed in Dutch and German language to measure PGD criteria based on the DSM 5-TR, as well as the ICD-11 (Lenferink et al., 2023). The TGI-CA consists of 22 interview items, adapted from the self-report version of the measure: The Traumatic Grief Inventory – Self Report Plus (TGI-SR+) (Lenferink, Eisma, et al., 2022). Each item enquires about the extent to which participants experienced specific grief reactions during the specified time period. In 14 out of 22 questions, the interviewer can incorporate the deceased individual's name or the relationship to the deceased person (e.g., "In the past 2 weeks, did you have intrusive thoughts or images related to the death of John/your brother?"). Moreover, the original version of the TGI-CA enquires about "the past month", which was replaced with "the past 2 weeks" to match the time period of the ESM phase. For the TGI-CA, answers are given using 5-point Likert scales with anchors 1 = never, 2 =

rarely, 3 = sometimes, 4 = often, 5 = always. The tool yielded a total score between 22 and 110 with higher scores indicating a higher probability of PGD symptoms being present. Therefore a new variable was created by summing the scores of each participant across all 22 items to represent total scores on the TGI-CA at T1 and T1b/T2 to assess participants' PGD symptom increase or decrease. Importantly, the measure has been found to have good internal consistency based on previous research (Lenferink, Franzen, et al., 2023). At T1, Cronbach's alpha for the TGI-CA in this sample was .890 in the ESM group and .883 in the control group. At T1b/T2, Cronbach's alpha for the TGI-CA in this sample was .871 in the ESM group and .911 in the control group. Therefore, the TGI-CA seemed to be a reliable and valid interview tool to assess PGD symptomatology.

ESM-Items

During the 2-week ESM phase, participants rated 17 items five times a day. Eleven of these ESM items represented PGD symptoms and were adapted to fit the ESM timeframe (Lenferink, van Eersel, & Franzen, 2022). Examples of such items were: "In the past 3 hours, I felt lonely" and "In the past 3 hours, it was difficult for me to do something (e.g., studying, working, social activities, sports, hobbies". Besides this, six contextual items were presented at each interval (e.g., "being alone vs. in company", "being at work vs. being at home").

Planned Data Processing and Statistical Analyses

Data analysis was conducted using SPSS V.29.0.1.0. Before analysing the data, all participants who dropped out between T1 and T1b/T2 were removed from the data set (*n* = 25). Although PGD scores in the control condition were not normally distributed, as assessed by a Shapiro-Wilk test, normality was assumed due to the relatively large sample size.

Subsequently, baseline comparisons between the ESM (*n* = 70) and control condition (*n* = 88) with respect to background and loss-related characteristics were performed to ensure an equivalent distribution of scores across the conditions (see Table 1). Chi-square tests were employed to assess dichotomised variables (gender (0 = man, 1 = woman), educational level (0 = non-university education, 1 = university), kinship (0 = other than child/partner, 1 = child/partner), cause of death (0

= natural, 1 = unnatural), history of support (0 = yes, 1 = no), and grief support (0 = yes, 1 = no). A Mann-Whitney U test was used for the ordinal variable expectedness of death. T-tests were utilised for continuous variables (i.e., age in years, age of the deceased, time since loss in days, and PGD scores). Current support was not included in Table 1 due to the amount of missing values in the data. Finally, mean PGD scores and standard deviations at baseline and T1b/T2 were calculated as well as indication of probable PGD (see Table 2).

To compare the effect of condition (ESM and control) and measurement timepoints (T1 and T1b/T2) on PGD scores, a two-factorial mixed ANOVA was conducted with condition and time as the independent variables and PGD scores as the dependent variable (model 1). Before running the analysis, all assumptions were checked. The data was normally distributed, as assessed by Shapiro-Wilk's test of normality ($p > .05$), except for T1b/T2 PGD scores in the control condition ($p = .016$). However, because of the relatively large sample size and robustness of ANOVAs, the analysis was run as usual. Furthermore, there was homogeneity of variances ($p > .05$) as assessed by a Levene's test. Because the within-subjects factor consisted of two categories (T1 and T1b/T2), the F-value did not need to be corrected for violations of sphericity. The interaction effect of condition x time revealed if PGD scores at T1 and T1b/T2 differed significantly depending on the condition that participants were assigned to, testing hypothesis 1.

For the second hypothesis, the model was expanded to include a potential moderator (gender) which was tested and observed through a three-way interaction. To test the potential moderating effect of gender, a three-factorial mixed ANOVA was run with condition, time, and gender as the independent variables and PGD scores as the dependent variable (model 2). PGD scores were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$), except for T1b/T2 PGD scores in woman-control ($p = .016$) and T1 PGD scores in man-control ($p = .047$). Again, because of the sample size and robustness of the test, the mixed ANOVA was run regardless. There was homogeneity of variances for PGD scores at T1 ($p = .418$) and PGD scores at T1b/T2 ($p = .501$), as assessed by Levene's test for equality of variances. The three-way interaction of condition x time x gender

revealed whether the relationship of condition and time on PGD scores was the same or different for women as for men, testing H2.

For both analyses, a significance level of $p < .05$ was used to determine statistical significance.

Table 1

Background and Loss-Related Characteristics of Participants at Baseline

Characteristic	Total sample (*N* = 158)		Condition				Analysis		
			ESM (*n* = 70)		Control (*n* = 88)		Test	*df*	*P*
	N	%	*N*	%	*N*	%			
Gender							χ^2 = .372	1	.542
Women	134	84.8	58	82.9	76	86.4			
Men	24	15.2	12	17.1	12	13.6			
Other	0	0.0	0	0.0	0	0.0			
Level of education							χ^2 = 1.182	2	.554
Primary school	0	0.0	0	0.0	0	0.0			
High school	12	7.6	5	7.1	7	8.0			
Vocational	45	28.5	23	32.9	22	25.0			
College/university	101	63.9	42	60.0	59	67.0			
Kinship							χ^2 = .411	1	.522
Partner	74	46.8	30	42.9	44	50.0			
Child	14	8.9	7	10.0	7	8.0			
Parent	54	34.2	26	37.1	28	31.8			
Sibling	7	4.4	5	7.1	2	2.3			
Grandparent	1	0.6	1	1.4	0	0.0			

Characteristic	Total sample (*N* = 158)		Condition				Analysis		
			ESM (*n* = 70)		Control (*n* = 88)		Test	*df*	*P*
Grandchild	1	0.6	0	0.0	1	1.1			
Friend	3	1.9	0	0.0	3	3.4			
Other	4	2.5	1	1.4	3	3.4			
Cause of death							χ^2 = .161	1	.688
Natural cause	122	77.2	53	75.7	69	78.4			
Accident	8	5.1	6	8.6	2	2.3			
Suicide	9	5.7	5	7.1	4	4.5			
Homicide	0	0.0	0	0.0	0	0.0			
Other	29	12.0	6	8.6	12	14.8			
Expectedness of death							*U* = 3102.500, *z* = .081		.935
Completely expected	36	22.8	13	18.6	23	26.1			
Slightly unexpected	23	14.6	14	20.0	9	10.2			
Somewhat unexpected	22	13.9	12	17.1	10	11.4			
Very unexpected	24	15.2	8	11.4	16	18.2			

			Condition				Analysis		
Characteristic	**Total sample (*N* = 158)**		**ESM (*n* = 70)**		**Control (*n* = 88)**		**Test**	***df***	***P***
Completely unexpected	53	33.5	23	32.9	30	34.1			
History of support [a]	84	53.2	38	54.3	46	52.3	χ^2 = .063	1	.801
Grief support [a]	64	40.5	27	38.6	37	42.0	χ^2 = .195	1	.659
	M	*SD*	*M*	*SD*	*M*	*SD*			
Age (years)	53.5	12.2	52.3	13.3	54.5	11.1	*t* = -1.108	133.887	.270
Age of deceased (years)	62.6	20.1	62.4	18.8	62.7	21.2	*t* = -.114	157	.909
Time since loss (days)	144.8	36.1	145.4	34.8	144.3	37.3	*t* = .191	156	.849
PGD score	56.4	14.0	56.9	14.2	55.9	14.0	*t* = .440	156	.661

Note. The following variables were dichotomised as was done in previous research (Lenferink et al., 2023): educational level (0 = other than college/university, 1 = college/university), relationship to the deceased (0 = other than child/partner, 1 = child/partner), and cause of death (0 = natural cause (i.e., physical illness), 1 = unnatural cause (i.e., accident, suicide, homicide, or other)).

[a] Reflects the number and percentage of participants answering “yes” to this question.

Results

Descriptive Statistics

Baseline background and loss-related characteristics are presented in Table 1. There were no statistically significant differences at baseline between the ESM and control condition on the background and loss-related characteristics. Therefore, both conditions appear to have been well balanced.

At T1, 60% (*n* = 42) of participants in the ESM condition showed indications of clinical levels of PGD, whereas 59.1% (*n* = 52) of participants in the control condition demonstrated probable PGD. At T1b/T2, 41.4% (*n* = 29) of participants in the ESM condition exhibited potentially clinical levels of PGD, and 44.3% (*n* = 39) of participants in the control condition displayed symptom levels indicative of PGD (see Table 2).

Table 2

Descriptive Statistics of PGD Scores and Probable PGD per Condition

	Total	ESM condition	Control condition
PGD scores, *M* (*SD*)			
T1 measurement	56.4 (14.0)	56.9 (14.2)	55.9 (14.0)
T1b/T2 measurement	50.5 (14.6)	49.5 (14.2)	51.3 (14.9)
Probable PGD, *n* (%)			
T1 measurement [a]	94 (59.5)	42 (60.0)	52 (59.1)
T1b/T2 measurement [a]	68 (43.0)	29 (41.4)	39 (44.3)

Note. *M* = mean, *SD* = standard deviation.

[a] Cut-off score $\geq$ 54, based on findings from previous research using the TGI-SR (Boelen et al., 2019).

Primary Analyses

Change in PGD Scores Over Time Depending on Condition

A two-factorial mixed ANOVA was conducted to examine if PGD scores changed differently over time depending on the condition that participants were assigned to (ESM or control). The ANOVA revealed no significant interaction effect between condition and time on PGD scores, $F(1, 156) = 1.528$, $p = .218$, partial $\eta 2 = .010$. In other words, PGD scores did not change (i.e., increase or decrease) significantly more from T1 to T1b/T2 for those in the ESM condition compared to the control condition.

When focusing on main effects, results uncovered that there was a statistically significant main effect of time, $F(1, 156) = 27.634$, $p < .001$, partial $\eta 2 = .150$. This main effect demonstrated that, when not taking into account the condition that participants were assigned to, PGD scores at T1b/T2 were significantly lower than PGD scores at T1. Alternatively, no statistically significant main effect of condition was observed, $F(1, 156) = .045$, $p = .833$, partial $\eta 2 = .000$ (see Table 3). This finding indicates that there was no statistically significant difference in mean PGD scores between the ESM condition and the control condition.

Exploring Gender as a Moderator

Furthermore, the study aimed to explore gender as a potential moderator of the relationship between time (T1 and T1b/T2) and condition (ESM or control) on PGD scores. Even though no statistically significant relationship was found between time and condition on PGD scores, a three-factorial mixed ANOVA was run nonetheless. The analysis of variance demonstrated that there was no statistically significant three-way interaction between gender, condition, and time, $F(1, 154) = .314$, $p = .576$, partial $\eta 2 = .002$. In other words, the effect of time and condition on PGD scores was not significantly different for women and men. Besides that, no significant two-way interactions were found, however, main effects of time and gender were uncovered, see Table 3.

Table 3

Efficacy of Condition, Time, and Gender on PGD Symptom Levels

Effect		Type III Sum of Squares	*df*	Mean Square	*F*	Sig.	Partial Eta Squared
			Model 1				
Time	Sphericity Assumed	2809.520	1	2809.520	27.634	< .001	.150
Condition	Sphericity Assumed	13.806	1	13.806	.045	.833	.000
Time * Condition	Sphericity Assumed	155.393	1	155.393	1.528	.218	.010
Error(Time)	Sphericity Assumed	15860.129	156	101.667			
			Model 2				
Time	Sphericity Assumed	1020.730	1	1020.730	9.985	.002	.061
Condition	Sphericity Assumed	44.171	1	44.171	.146	.703	.001
Gender	Sphericity Assumed	1266.161	1	1266.161	4.172	.043	.026
Time * Condition	Sphericity Assumed	29.172	1	29.172	.285	.594	.002
Time * Gender	Sphericity Assumed	88.046	1	88.046	.861	.355	.006
Condition * Gender	Sphericity Assumed	136.232	1	136.232	.449	.504	.003
Time * Condition * Gender	Sphericity Assumed	32.110	1	32.110	.314	.576	.002
Error(Time)	Sphericity Assumed	15742.126	154	102.222			

Note. *df* = degrees of freedom.

Discussion

The aim of this study was to examine (1) if self-reporting grief symptoms multiple times a day significantly decreased PGD symptoms over time compared to being waitlisted and (2) if gender moderated this relationship in that, within the ESM condition PGD symptoms decreased more for men than for women. This is, to my knowledge, the first study explicitly examining the potential reactivity effect of ESM on PGD symptomatology. In doing so, 158 Dutch, recently bereaved individuals recorded the severity of eleven grief reactions in the past 3 hours, five times a day for 2 weeks. Results revealed that participants demonstrated significant improvement over time in PGD symptom levels, yet, those in the ESM condition did not show significantly larger changes over time than participants in the control condition. Furthermore, the effect of time and condition on PGD scores did not significantly differ depending on gender.

Based on Lenferink, van Eersel, and Franzen's (2022) findings that ESM may be able to serve as a way of self-monitoring PGD symptoms, it was hypothesised that participating in an ESM phase would lead to a significant decrease in PGD symptoms compared to being waitlisted. However, contrary to what was expected, the current study found that although PGD symptoms decreased from T1 to T1b/T2, PGD symptom levels did not decrease significantly more for those in the ESM condition compared to the control condition. Perhaps self-reporting PGD symptoms multiple times a day does not result in significant reductions in PGD symptom levels compared to being waitlisted because simply registering one's symptoms is not adequate to enhance individuals' awareness of cognitive and emotional processes, as well as behavioural patterns. Alternatively, intervention studies from depression literature suggest that combining ESM with personalised feedback enhances its effectiveness compared to solely engaging in ESM (Kramer et al., 2014; Suen et al., 2022). Including a personalised feedback component presented participants with a summary of their current affective state in a variety of daily contexts, as well as the relationship with their mood state, which provided participants with an explicit overview of their symptoms. Reflecting symptoms back to participants

through ESM-derived feedback may have enabled individuals to recognise the variability in symptom levels. This may have helped them acknowledge times with little to no PGD symptoms and identify real or imagined causal relationships between activities, emotions, or time of day and symptom levels potentially increasing participants sense of control over their grief reactions leading to more pronounced and long-lasting effects on PGD symptomatology. Overall, as the current findings suggest that ESM in itself is not enough to reduce PGD symptom levels in recently bereaved individuals future research should consider examining the effectiveness of an intervention combining ESM and personalised feedback.

Alternatively, there might truly be no difference in PGD symptom level change over time between participating in an ESM phase and being waitlisted. It cannot be ruled out that the PGD score changes over time in both groups were attributable to study procedures that could have had an activating effect, such as the interviews at the beginning of the study. During the baseline interviews, participants had the opportunity to share their bereavement experiences in a supportive environment with a trained interviewer. Doing so may have encouraged participants to start actively processing the loss as well as increasing awareness of one's grief reactions, regardless of the condition participants were assigned to. Therefore, findings in the current study preliminarily suggest that merely enrolling in a grief study and completing a baseline interview may be enough to reduce PGD symptoms in recently bereaved individuals. However, as this is the first RCT examining the reactivity effect of ESM on PGD symptom levels, replication is imperative to corroborate these findings.

On the other hand, it is also possible that the change in PGD symptom levels over time is significantly different for the ESM condition compared to the control condition, but that the effect could not be detected in the current sample. Upon inspection of the standard deviations, the variability of PGD scores in both conditions at all time points appeared to be relatively large ($SD \approx 14$), potentially resulting in a statistical lack of power. As PGD symptoms tend to naturally decrease

over time (Szuhany et al., 2021) symptom levels at 3 months post-loss compared to 6 months post-loss can differ substantially, therefore the time point at which individuals enrolled in the study may have resulted in considerable disparity in PGD scores between participants. Besides that, sample sizes in both conditions were somewhat unequal both potentially leading to a loss in power to detect group differences. Indeed, although not found to be significant, descriptive statistics did suggest the existence of an interaction effect as mean PGD scores decreased more for those in the ESM condition (from *M* = 56.90 (*SD* = 14.19) to *M* = 49.49 (*SD* = 14.15)) than for participants in the control condition (from *M* = 55.91 (*SD* = 13.96) to *M* = 51.32 (*SD* = 14.92)). Besides that, over half of participants (54.3%, *n* = 38) reported that they gained something positive from participating in the ESM phase (although keep in mind the risk of agreement bias (Chang et al., 2019)). These factors seem to suggest that individuals did experience some utility from participating in the ESM phase even though this was not reflected in the inferential statistics. Thus, to determine if there is truly no significant effect of ESM on PGD symptom level change compared to being waitlisted, replication is needed involving a large sample with comparable numbers of participants in each condition using an even more concise time frame of time since loss.

Furthermore, based on findings from a previous study (Schut et al., 1997), it was expected that men would benefit more from emotion-focused items, while women would benefit more from action-focused items. Thus, the hypothesis was that the ESM phase, which could be considered emotion-focused as it required participants to think about their current emotional state, would reduce PGD symptoms over time significantly more for men than for women. However, results did not indicate gender differences in the reactivity effect of ESM on PGD symptom levels over time.

The lack of gender differences in reactivity to the ESM phase may be attributable to the fact that, although completing the ESM assessments encouraged participants to contemplate their current emotional state, the ESM phase did not exclusively comprise emotion-focused items. It could be argued that despite the majority of ESM items appearing to be emotion-focused (i.e., In the past

three hours, I felt sad because of his or her death/I felt bitterness or anger because of his or her death/I felt emotionally numb because of his or her death), some of the ESM items (i.e., In the past three hours, I avoided places, objects, or thoughts that reminded me that he or she is dead/it was difficult for me to do something), as well as the contextual items (i.e., In the past three hours, what activity did you spend the most time on), could be considered action-focused instead as they targeted the behavioural aspects of grief. Thus, since a combination of emotion-focused and action-focused items were used, women and men may have benefitted equally from the ESM phase resulting in no significant gender differences in the effect of ESM on PGD symptom levels over time.

However, caution is advised when interpreting the results due to the substantial differences in sample size between women (*n* = 134) and men (*n* = 24) in the current study. While gender could be a significant moderator, in that the effect of condition on PGD symptom level change over time does vary depending on gender, the imbalance in sample size may have restricted the study's ability to detect such differences. The prevalent overrepresentation of women, consistent with broader bereavement research trends (Stroebe et al., 2003), complicates the examination of gender effects in convenience samples. It is therefore empirical that future research explores the moderating effect of gender on PGD symptoms over time depending on condition, using population-representative samples.

The present study had several strengths. Firstly, the sample size of 158 participants is relatively large compared to previous ESM studies (van Berkel et al., 2017). Furthermore, as the current study was a RCT a control group was used to examine the effectiveness of ESM on PGD symptom change compared to being waitlisted. Moreover, participants' background and loss-related characteristics were found to be comparable between conditions.

Numerous limitations should be considered when interpreting the current findings. First of all, the majority of participants were college or university-educated and experienced a natural loss. Furthermore, no individuals indicated having been bereaved because of homicide. Besides that, no

one in the current sample identified as non-binary. The results may therefore not be generalisable to individuals who (i) are not college or university-educated, (ii) lost a loved one due to traumatic circumstances, specifically homicide, or (iii) do not identify as woman or man. Secondly, participants were mainly recruited through an online grief monitoring tool. Because of this, the sample consisted of individuals who self-selected to take part and had actively searched for information about bereavement symptoms and grief care. Therefore, the current study may have attracted individuals with higher PGD symptom levels compared to the population of recently bereaved individuals, again impacting the generalisability of findings. Although challenging and time-consuming, future research should contemplate using random sampling techniques to recruit participants in order to reduce self-selection bias. Thirdly, because retrospective data was used to determine PGD symptom levels, there is a risk of recall bias influencing reporting. Participants may have overreported negative mood states (Sato & Kawahara, 2011), provided answers biased to their current mood or cognitive state (Faul & LaBar, 2022; Nahleen et al., 2019), or answered based on the most intense and proximate events experienced during the assessed period (Horwitz et al., 2023) which could have led to an over- or under-estimation of symptoms. To minimise recall bias, future studies could compare aggregated ESM scores with PGD symptom levels reported by participants in the interviews to ensure alignment or they could choose to investigate temporal effects within the ESM data (e.g., comparing score at the start versus the end of the study period) rather than relying primarily on retrospective questionnaires. Finally, although there is preliminary evidence that telephone administration of the TGI-CA provides reliable and valid data (Lenferink et al., 2023), it is unclear to what extent the results would have differed if virtual self-reporting had been used instead. Moreover, interviewer differences could not be controlled for which, in the absence of tests of inter-rater reliability, might raise concerns regarding the reliability of measurement. Despite standardised training and efforts to minimise interviewer impact by using scripts, variations in consolation, interest, and/or time allocation by research assistants may have occurred. If participants perceived an interviewer as more supportive and felt validated, they may have been more inclined to share

difficult emotions and provide truthful information about their symptoms thereby influencing responding above and beyond the condition that an individual was assigned to. Thus, future studies should consider either utilising a single interviewer for conducting all interviews or recording interviews and providing feedback to research assistants to minimise interviewer differences.

Despite these limitations, the present study has advanced our understanding of the effect of self-reporting symptoms multiple times a day on PGD symptom levels. Even though ESM was not found to significantly decrease PGD symptoms over time compared to the control condition, self-monitoring symptoms using ESM may still help recently bereaved individuals gain some insight into their grief symptoms. Importantly, as this is the first RCT of its kind, replication of the study is necessary to determine if there is truly no effect of ESM on PGD symptom levels compared to being waitlisted amongst Dutch recently bereaved individuals. If so, research should focus on examining the effectiveness of combining ESM and personalised feedback as an intervention to reduce PGD symptoms. Moreover, even if ESM does not directly reduce PGD symptoms, it might still be a valuable tool for treatment monitoring purposes in clinical practice (Bos et al., 2015). Furthermore, self-monitoring through ESM can provide clinicians and patients with additional information that can contribute to the personalisation of treatment (Verhagen et al., 2016). For instance, if the worsening of symptoms tends to be linked to a certain time of day, clinicians can tailor treatment interventions accordingly.

In conclusion, this study demonstrated that while participants' levels of PGD symptoms significantly improved over time, the ESM condition did not exhibit significantly greater changes over time than the control condition. Several explanations may account for these findings, such as the absence of a personalised feedback component or insufficient statistical power to detect group differences within the current sample. Alternatively, changes in PGD symptom levels over time in both conditions may have been attributable to study procedures having an activating effect. Additionally, the influence of time and condition on PGD symptom levels did not significantly vary by

gender. However, sample characteristics limit the generalisability of findings and self-selection bias, recall bias, and interviewer differences may have influenced the results. Pending replication, it can tentatively be concluded that solely self-monitoring symptoms through ESM does not reduce PGD symptom levels more than being waitlisted.

References

American Psychiatric Association. (2022). Prolonged Grief Disorder. In *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425596.dsm05

Boelen, P. A., Djelantik, A. A. A. M. J., de Keijser, J., Lenferink, L. I. M., & Smid, G. E. (2019). Further validation of the Traumatic Grief Inventory-Self Report (TGI-SR): A measure of persistent complex bereavement disorder and prolonged grief disorder. *Death Studies*, *43*(6), 351–364. https://doi.org/10.1080/07481187.2018.1480546

Boelen, P. A., & Lenferink, L. I. M. (2022). Prolonged grief disorder in DSM-5-TR: Early predictors and longitudinal measurement invariance. *Australian and New Zealand Journal of Psychiatry*, *56*(6), 667–674. https://doi.org/10.1177/00048674211025728

Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? *The American Psychologist*, *59*(1), 20–28. https://doi.org/10.1037/0003-066X.59.1.20

Bos, F. M., Schoevers, R. A., & aan het Rot, M. (2015). Experience sampling and ecological momentary assessment studies in psychopharmacology: A systematic review. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, *25*(11), 1853–1864. https://doi.org/10.1016/j.euroneuro.2015.08.008

Burke, L. A., Neimeyer, R. A., Bottomley, J. S., & Smigelsky, M. A. (2019). Prospective Risk Factors for Intense Grief in Family Members of Veterans Who Died of Terminal Illness. *Illness, Crisis & Loss*, *27*(3), 147–171. https://doi.org/10.1177/1054137317699580

Chang, E. M., Gillespie, E. F., & Shaverdian, N. (2019). Truthfulness in patient-reported outcomes: Factors affecting patients' responses and impact on data quality. *Patient Related Outcome Measures*, *10*, 171–186. https://doi.org/10.2147/PROM.S178344

Creighton, G., Oliffe, J. L., Butterwick, S., & Saewyc, E. (2013). After the death of a friend: Young men's grief and masculine identities. *Social Science & Medicine*, *84*, 35–43. https://doi.org/10.1016/j.socscimed.2013.02.022

Djelantik, A. A. A. M. J., Smid, G. E., Kleber, R. J., & Boelen, P. A. (2017). Early indicators of problematic grief trajectories following bereavement. *European Journal of Psychotraumatology*, *8*(sup6), 1423825. https://doi.org/10.1080/20008198.2018.1423825

Doering, B. K., Barke, A., Vogel, A., Comtesse, H., & Rosner, R. (2022). Predictors of Prolonged Grief Disorder in a German Representative Population Sample: Unexpectedness of Bereavement Contributes to Grief Severity and Prolonged Grief Disorder. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.853698

Faul, L., & LaBar, K. S. (2022). Mood-congruent memory revisited. *Psychological Review, 130*(6), 1421-1456. https://doi.org/10.1037/rev0000394

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience Sampling Method: Measuring the Quality of Everyday Life*. SAGE Publications, Inc.

Holmes, T. H., & Rahe, R. H. (1967). The Social Readjustment Rating Scale. *Journal of Psychosomatic Research*, *11*(2), 213–218. https://doi.org/10.1016/0022-3999(67)90010-4

Horwitz, A. G., Zhao, Z., & Sen, S. (2023). Peak-end bias in retrospective recall of depressive symptoms on the PHQ-9. *Psychological Assessment*, *35*(4), 378–381. https://doi.org/10.1037/pas0001219

Jones, K., Robb, M., Murphy, S., & Davies, A. (2019). New understandings of fathers' experiences of grief and loss following stillbirth and neonatal death: A scoping review. *Midwifery*, *79*, 102531. https://doi.org/10.1016/j.midw.2019.102531

Jordan, A. H., & Litz, B. T. (2014). Prolonged grief disorder: Diagnostic, assessment, and treatment considerations. *Professional Psychology: Research and Practice*, *45*(3), 180–187. https://doi.org/10.1037/a0036836

Kauer, S. D., Reid, S. C., Crooke, A. H. D., Khor, A., Hearps, S. J. C., Jorm, A. F., Sanci, L., & Patton, G. (2012). Self-monitoring using mobile phones in the early stages of adolescent depression: Randomized controlled trial. *Journal of Medical Internet Research*, *14*(3), e67. https://doi.org/10.2196/jmir.1858

Kersting, A., Brähler, E., Glaesmer, H., & Wagner, B. (2011). Prevalence of complicated grief in a representative population-based sample. *Journal of Affective Disorders*, *131*(1–3), 339–343. https://doi.org/10.1016/j.jad.2010.11.032

Komischke-Konnerup, K. B., Zachariae, R., Johannsen, M., Nielsen, L. D., & O'Connor, M. (2021). Co-occurrence of prolonged grief symptoms and symptoms of depression, anxiety, and posttraumatic stress in bereaved adults: A systematic review and meta-analysis. *Journal of Affective Disorders Reports*, *4*, 100140. https://doi.org/10.1016/j.jadr.2021.100140

Kramer, I., Simons, C. J. P., Hartmann, J. A., Menne-Lothmann, C., Viechtbauer, W., Peeters, F., Schruers, K., van Bemmel, A. L., Myin-Germeys, I., Delespaul, P., van Os, J., & Wichers, M. (2014). A therapeutic application of the experience sampling method in the treatment of depression: A randomized controlled trial. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, *13*(1), 68–77. https://doi.org/10.1002/wps.20090

Lenferink, L. I. M., Boelen, P. A., Heeke, C., & Franzen, M. (2019). *The TGI-CA Assessment after Loss in Europe (The TALE-Study): Psychometric evaluation of an interview to assess DSM-5 and ICD-11 pathological grief criteria*. https://doi.org/10.17605/OSF.IO/A6HMC

Lenferink, L. I. M., Eisma, M. C., Smid, G. E., de Keijser, J., & Boelen, P. A. (2022). Valid measurement of DSM-5 persistent complex bereavement disorder and DSM-5-TR and ICD-11 prolonged grief disorder: The Traumatic Grief Inventory-Self Report Plus (TGI-SR+). *Comprehensive Psychiatry*, *112*, 152281. https://doi.org/10.1016/j.comppsych.2021.152281

Lenferink, L. I. M., Franzen, M., Ten Klooster, P. M., Knaevelsrud, C., Boelen, P. A., & Heeke, C. (2023). The Traumatic Grief Inventory-Clinician Administered: A psychometric evaluation of a new interview for

ICD-11 and DSM-5-TR prolonged grief disorder severity and probable caseness. *Journal of Affective Disorders*, *330*, 188–197. https://doi.org/10.1016/j.jad.2023.03.006

Lenferink, L. I. M., Nickerson, A., de Keijser, J., Smid, G. E., & Boelen, P. A. (2020). Trajectories of grief, depression, and posttraumatic stress in disaster-bereaved people. *Depression and Anxiety*, *37*(1), 35–44. https://doi.org/10.1002/da.22850

Lenferink, L. I. M., van Eersel, J. H. W., & Franzen, M. (2022). Is it acceptable and feasible to measure prolonged grief disorder symptoms in daily life using experience sampling methodology? *Comprehensive Psychiatry*, *119*, 152351. https://doi.org/10.1016/j.comppsych.2022.152351

Lundorff, M., Holmgren, H., Zachariae, R., Farver-Vestergaard, I., & O'Connor, M. (2017). Prevalence of prolonged grief disorder in adult bereavement: A systematic review and meta-analysis. *Journal of Affective Disorders*, *212*, 138–149. https://doi.org/10.1016/j.jad.2017.01.030

McDevitt-Murphy, M. E., Luciano, M. T., & Zakarian, R. J. (2018). Use of Ecological Momentary Assessment and Intervention in Treatment With Adults. *Focus (American Psychiatric Publishing)*, *16*(4), 370–375. https://doi.org/10.1176/appi.focus.20180017

Myin-Germeys, I., Kasanova, Z., Vaessen, T., Vachon, H., Kirtley, O., Viechtbauer, W., & Reininghaus, U. (2018). Experience sampling methodology in mental health research: New insights and technical developments. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA)*, *17*(2), 123–132. https://doi.org/10.1002/wps.20513

Nahleen, S., Nixon, R. D. V., & Takarangi, M. K. T. (2019). Current PTSD symptomatology distorts memory for past symptoms. *Psychiatry Research*, *274*, 330–334. https://doi.org/10.1016/j.psychres.2019.02.052

Nielsen, M. K., Carlsen, A. H., Neergaard, M. A., Bidstrup, P. E., & Guldin, M.-B. (2019). Looking beyond the mean in grief trajectories: A prospective, population-based cohort study. *Social Science & Medicine*, *232*, 460–469. https://doi.org/10.1016/j.socscimed.2018.10.007

Possemato, K., Maisto, S. A., Wade, M., Barrie, K., McKenzie, S., Lantinga, L. J., & Ouimette, P. (2015). Ecological momentary assessment of PTSD symptoms and alcohol use in combat veterans. *Psychology of Addictive Behaviors: Journal of the Society of Psychologists in Addictive Behaviors*, *29*(4), 894–905. https://doi.org/10.1037/adb0000129

Prigerson, H. G., Horowitz, M. J., Jacobs, S. C., Parkes, C. M., Aslan, M., Goodkin, K., Raphael, B., Marwit, S. J., Wortman, C., Neimeyer, R. A., Bonanno, G. A., Block, S. D., Kissane, D., Boelen, P., Maercker, A., Litz, B. T., Johnson, J. G., First, M. B., & Maciejewski, P. K. (2009). Prolonged grief disorder: Psychometric validation of criteria proposed for DSM-V and ICD-11. *PLoS Medicine*, *6*(8), e1000121. https://doi.org/10.1371/journal.pmed.1000121

Ptacek, J. T., Smith, R. E., & Zanas, J. (1992). Gender, appraisal, and coping: A longitudinal analysis. *Journal of Personality*, *60*(4), 747–770. https://doi.org/10.1111/j.1467-6494.1992.tb00272.x

Sato, H., & Kawahara, J. (2011). Selective bias in retrospective self-reports of negative mood states. *Anxiety, Stress, & Coping*, *24*(4), 359–367. https://doi.org/10.1080/10615806.2010.543132

Schut, H. A., Stroebe, M. S., van den Bout, J., & de Keijser, J. (1997). Intervention for the bereaved: Gender differences in the efficacy of two counselling programmes. *The British Journal of Clinical Psychology*, *36*(1), 63–72. https://doi.org/10.1111/j.2044-8260.1997.tb01231.x

Snippe, E., Simons, C. J. P., Hartmann, J. A., Menne-Lothmann, C., Kramer, I., Booij, S. H., Viechtbauer, W., Delespaul, P., Myin-Germeys, I., & Wichers, M. (2016). Change in daily life behaviors and depression: Within-person and between-person associations. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *35*(5), 433–441. https://doi.org/10.1037/hea0000312

Stelzer, E. M., Atkinson, C., O'Connor, M. F., & Croft, A. (2019). Gender differences in grief narrative construction: A myth or reality? *European Journal of Psychotraumatology*, *10*(1), 1688130. https://doi.org/10.1080/20008198.2019.1688130

Stroebe, M., Stroebe, W., & Schut, H. (2003). Bereavement research: Methodological issues and ethical concerns. *Palliative Medicine*, *17*(3), 235–240. https://doi.org/10.1191/0269216303pm768rr

Suen, Y. N., Yau, J. Y., Wong, P. S., Li, Y. K., Hui, C. L. M., Chan, S. K. W., Lee, E. H. M., Chang, W. C., & Chen, E. Y. H. (2022). Effect of brief, personalized feedback derived from momentary data on the mental health of women with risk of common mental disorders in Hong Kong: A randomized clinical trial. *Psychiatry Research*, *317*, 114880. https://doi.org/10.1016/j.psychres.2022.114880

Szuhany, K. L., Malgaroli, M., Miron, C. D., & Simon, N. M. (2021). Prolonged Grief Disorder: Course, Diagnosis, Assessment, and Treatment. *FOCUS*, *19*(2), 161–172. https://doi.org/10.1176/appi.focus.20200052

Telford, C., McCarthy-Jones, S., Corcoran, R., & Rowse, G. (2012). Experience Sampling Methodology studies of depression: The state of the art. *Psychological Medicine*, *42*(6), 1119–1129. https://doi.org/10.1017/S0033291711002200

Treml, J., Kaiser, J., Plexnies, A., & Kersting, A. (2020). Assessing prolonged grief disorder: A systematic review of assessment instruments. *Journal of Affective Disorders*, *274*, 420–434. https://doi.org/10.1016/j.jad.2020.05.049

Vachon, H., Bourbousson, M., Deschamps, T., Doron, J., Bulteau, S., Sauvaget, A., & Thomas-Ollivier, V. (2016). Repeated self-evaluations may involve familiarization: An exploratory study related to Ecological Momentary Assessment designs in patients with major depressive disorder. *Psychiatry Research*, *245*, 99–104. https://doi.org/10.1016/j.psychres.2016.08.034

van Berkel, N., Ferreira, D., & Kostakos, V. (2017). The Experience Sampling Method on Mobile Devices. *ACM Computing Surveys*, *50*(6), 1-40. https://doi.org/10.1145/3123988

van Os, J., Verhagen, S., Marsman, A., Peeters, F., Bak, M., Marcelis, M., Drukker, M., Reininghaus, U., Jacobs, N., Lataster, T., Simons, C., ESM-MERGE Investigators PhD, Lousberg, R., Gülöksüz, S., Leue, C., Groot, P. C., Viechtbauer, W., & Delespaul, P. (2017). The experience sampling method as an

mHealth tool to support self-monitoring, self-insight, and personalized health care in clinical practice. *Depression and Anxiety*, *34*(6), 481–493. https://doi.org/10.1002/da.22647

Verhagen, S. J. W., Hasmi, L., Drukker, M., van Os, J., & Delespaul, P. A. E. G. (2016). Use of the experience sampling method in the context of clinical trials. *Evidence-Based Mental Health*, *19*(3), 86–89. https://doi.org/10.1136/ebmental-2016-102418

Wen, F. H., Prigerson, H. G., Chou, W. C., Huang, C. C., Hu, T. H., Chiang, M. C., Chuang, L. P., & Tang, S. T. (2022). How symptoms of prolonged grief disorder, posttraumatic stress disorder, and depression relate to each other for grieving ICU families during the first two years of bereavement. *Critical Care*, *26*(1), 336. https://doi.org/10.1186/s13054-022-04216-5

World Health Organization. (2022). *ICD-11: International classification of diseases* (11th ed.). https://icd.who.int/

World Health Organization, (n.d.). *Gender and Health*. https://www.who.int/health-topics/gender

Yoon, H., Park, G. R., & Kim, J. (2022). Psychosocial trajectories before and after spousal loss: Does gender matter? *Social Science & Medicine*, *294*, 114701. https://doi.org/10.1016/j.socscimed.2022.114701

Zisook, S., Iglewicz, A., Avanzino, J., Maglione, J., Glorioso, D., Zetumer, S., Seay, K., Vahia, I., Young, I., Lebowitz, B., Pies, R., Reynolds, C., Simon, N., & Shear, M. K. (2014). Bereavement: Course, consequences, and care. *Current Psychiatry Reports*, *16*(10), 482. https://doi.org/10.1007/s11920-014-0482-8