cpe.psychopen.eu | ISSN 2625-3410 | March 2024 | Volume 6 | Issue 1 |

# CLINICAL PSYCHOLOGY IN EUROPE

The Official Academic Journal of the European Association of Clinical Psychology and Psychological Treatment

PsychOpen GOLD

# CLINICAL PSYCHOLOGY IN EUROPE

The Official Academic Journal of the
European Association of Clinical Psychology and Psychological Treatment





PsychOpen GOLD

CLINICAL PSYCHOLOGY IN EUROPE

March 2024 | Volume 6 | Issue 1

# Contents



PsychOpen GOLD is a publishing service by
Leibniz Institute for Psychology Information (ZPID), Germany.
www.leibniz-psychology.org

PsychOpen GOLD

Clinical Psychology in Europe

PsychOpen GOLD

Editorial

Check for updates

# Mind the Gap – Ideas for Making Clinical Research More Relevant for Practitioners and Patients

Max Berg[1], Lea Schemer[2], Lukas Kirchner[1], Saskia Scholten[2]

**[1]** *Clinical Psychology Group, University of Marburg, Marburg, Germany.* **[2]** *Department of Clinical Psychology and Psychotherapy, RPTU Kaiserslautern-Landau, Landau, Germany.*

Clinical Psychology in Europe, 2024, Vol. 6(1), Article e12419, https://doi.org/10.32872/cpe.12419

**Received:** 2023-07-17 • **Accepted:** 2024-03-06 • **Published (VoR):** 2024-03-28

**Handling Editor:** Cornelia Weise, Philipps-University of Marburg, Marburg, Germany

**Corresponding Author:** Max Berg, University of Marburg, Gutenbergstraße 18, 35032 Marburg, Germany. E-mail: max.berg@uni-marburg.de

Randomized controlled trials (RCTs) are widely considered to be the gold standard for demonstrating *efficacy* in psychotherapy research. However, the *clinical utility* of "typical" RCTs for establishing routine care therapies has been a topic of long-standing debate in our field (Persons & Silberschatz, 1998). "Typical" refers to a study with a small to moderate sample size that targets a disorder according to a standardized diagnostic manual and is often waiting-list controlled (Carey & Stiles, 2016). Practitioners frequently express criticism about the external validity of such RCTs (Gyani et al., 2015; Safran et al., 2011). In qualitative investigations, therapists describe the "unrepresentativeness of RCTs" as a reason for why they do not regard clinical research as an important foundation for their everyday decision making (Gyani et al., 2015). A review suggested that the perceived "inflexibility" of manuals could also be related to the lack of interest of many practitioners and that therapists wonder whether the "standardized instructions" provided in them are useful for their heterogeneous clinical use cases (Speers et al., 2022).

Similarly, from a methodological point of view, the inference to intra-individual variability from group-level research was challenged (Fisher et al., 2018). Furthermore, the substantial heterogeneity in treatment effects suggests that even if patients with the same diagnoses are treated with the same treatment by the same therapist, they respond differently (Herzog & Kaiser, 2022). Given the methodological challenges and the skepticism of therapists, we argue that the criticism regarding clinical science should be taken seriously. In this editorial, we present five ideas for improving psychotherapy research and for addressing the research practice gap.

# Five Ideas for Psychotherapy Research

## Idea One: Focus on Transdiagnostic Mechanisms

Diagnoses in clinical psychology typically do not present homogeneous entities, and comorbidity rates between "different" disorders are commonly high (Rief et al., 2023). For example, different patients exhibit largely heterogeneous symptom dynamics in depression and novel clinical research is starting to acknowledge this (Fried et al., 2023). Also, we know from a large body of research that pathological mechanisms are not limited to a single disorder, but oftentimes pose transdiagnostic problems (Dalgleish et al., 2020). Transdiagnostic mechanisms include (but are not limited to), dysfunctional expectations with aberrant belief updating (Kirchner et al., 2022), social impairments (Lehmann et al., 2019), and reward insensitivity and its interplay with stress dysregulation (Martin-Soelch, 2023). Given the potential of transdiagnostic mechanisms, it seems worthwhile to allocate treatment based on them rather than solely based on diagnoses. For example, patients who exhibit a high tendency for repetitive negative thinking could be assigned to focused therapies that target this mechanism, regardless of whether they have been diagnosed with depression, generalized anxiety disorder, or both.

## Idea Two: Dismantle Treatment Protocols

A plethora of therapeutic techniques exist to target (transdiagnostic) mechanisms (Schaeuffele et al., 2021). Yet, we know little about their isolated effect because treatment manuals oftentimes with overlapping strategies – are evaluated as a *treatment package*. Such "evidence-based black boxes" are effective for treating numerous mental disorders, but their respective effect sizes and response rates remain moderate (Ormel et al., 2022). Future research should dismantle treatment protocols and evaluate the effect of specific techniques. Applying dismantled techniques instead of treatment protocols might be closer to clinical practice anyway, where the implementation of complex procedures is limited due to time, comorbidity patterns, and financial resources. The dismantling of treatment packages may also necessitate a departure from traditional therapy orientations. Competence-oriented frameworks (Rief, 2021), or process-based therapy (Moskow et al., 2023) are two approaches that could promote a more "toolbox oriented" thinking.

## Idea Three: Monitor Individual Trajectories With Sufficient Resolution

In clinical research and practice, diagnostic instruments are usually collected at only a few points in time (e.g., before and after treatment). To date, few projects exist that collect intensive longitudinal data (i.e., session-by-session data or ecological momentary assessment) in clinical trials and routine care settings (Lutz et al., 2022). These methods would allow to monitor individual trajectories, compare patients to similar cases and

provide computerized treatment suggestions, while reducing therapist biases regarding outcome estimation (Lutz et al., 2022). M-Path and Shiny apps are digital implementations of such efforts (Mestdagh et al., 2023). However, just because appropriate tools are available does not mean they are already being frequently used. Barriers, particularly in terms of usability and knowledge of digital technologies, can make it difficult for clinicians to use digital innovations. Therefore, it is vital that the curricula of psychology students are expanded from science- to practice-oriented use of data literacy and computer science.

## Idea Four: Use Causal Inference Methods for Routine Care Data

Large psychopathology data sets exist in routine care, but we need to sample and process them in a way that allows for causal inference. It was suggested that we can develop alternatives to RCTs for estimating the causal effect of a given treatment on an outcome. In addition to established approaches like propensity score matching (Lee & Little, 2017), single-case experimental designs can be utilized as an ideographic alternative for RCTs. These designs utilize an experimental manipulation that compares the individual response of a patient at different time points. (e.g., during treatment delivery versus waiting-periods). Single-case experimental designs have the potential to empower practitioners to become scientist-practitioners of their own clinical practice (Kazdin, 2019).

"Synthetic waitlists" can also bring causal inference into psychotherapy research (Kaiser et al., 2023). Here, machine learning algorithms select patients from waiting lists, based on the multidimensional similarity to a given patient under treatment. This, in turn, allows to estimate the probability that a specific patient would have reached a certain outcome without receiving therapy. If this probability is low, then a significant part of the improvement can be attributed to the treatment. Utilizing synthetic waitlists allow us to harvest routine data as an additional source of information and estimate the effect of therapeutic strategies under realistic, everyday conditions.

## Idea Five: Utilize the Expertise of Practitioners and the Lived Experience of Patients

Participatory science actively engages various stakeholders throughout the entire research process (Slattery et al., 2020). For example, patients should be involved as experts in the development of clinically relevant research questions (Birnie et al., 2019), the optimization of treatment manuals (Schemer et al., 2023), and for the planning of upcoming research projects (Slattery et al., 2020). Similarly, practitioners can be involved to facilitate the clinical usefulness of technological advances or to find ways to overcome practical barriers to implement an effective therapeutic strategy. Here, we face the challenge of involving practitioners who are distant or even skeptical about clinical research. Yet, a participatory approach would help to address the research-practice gap

by involving groups for whom practically relevant and effective clinical science is in their vital self-interest.

## Conclusion

In conclusion, addressing the research-practice gap requires a shift towards dismantling the effect of specific therapeutic techniques on better-operationalized transdiagnostic mechanisms. Monitoring individual trajectories and using innovative methods for inference can provide valuable insights into therapy effectiveness, if needed at an individual level. Finally, an active involvement of non-scientists can create research that is interesting and engaging for different stakeholders.

**Funding:** The post-doc position of Max Berg was funded by the PsyChange initiative (funding number: 56040018), by the Hessian Ministry of Science and Arts when this editorial was written. This work was also supported by the DYNAMIC center, funded by the LOEWE program of the Hessian Ministry of Science and Arts (grant number: LOEWE1/16/519/03/09.001(0009)/98).

**Acknowledgments:** The PsyChange initiative by the Hessian Ministry of Science and Arts organized think-tanks, symposia, and expert interviews with the goal of conducting meta-science to improve psychological treatments. This work was endorsed and made possible by the initiative and Max Berg wants to thank all PsyChange members for the ongoing support.

**Competing Interests:** All authors declare no competing interests, financial or otherwise. The ideas expressed in this editorial are solely based on the consensual personal opinions of the authors.

**Preprint Disclosure:** A preprint can be found at the PsyArXiv server: https://doi.org/10.31234/osf.io/2qvhy

## References

Birnie, K. A., Dib, K., Ouellette, C., Dib, M. A., Nelson, K., Pahtayken, D., Baerg, K., Chorney, J., Forgeron, P., Lamontagne, C., Noel, M., Poulin, P., & Stinson, J. (2019). Partnering for pain: A priority setting partnership to identify patient-oriented research priorities for pediatric chronic pain in Canada. *Canadian Medical Association Journal Open, 7*(4), E654–E664. https://doi.org/10.9778/cmajo.20190060

Carey, T. A., & Stiles, W. B. (2016). Some problems with randomized controlled trials and some viable alternatives. *Clinical Psychology & Psychotherapy, 23*(1), 87–95. https://doi.org/10.1002/cpp.1942

Dalgleish, T., Black, M., Johnston, D., & Bevan, A. (2020). Transdiagnostic approaches to mental health problems: Current status and future directions. *Journal of Consulting and Clinical Psychology, 88*(3), 179–195. https://doi.org/10.1037/ccp0000482

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences of the United States of America, 115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Fried, E. I., Proppert, R. K. K., & Rieble, C. L. (2023). Building an early warning system for depression: Rationale, objectives, and methods of the WARN-D study. *Clinical Psychology in Europe, 5*(3), Article e10075. https://doi.org/10.32872/cpe.10075

Gyani, A., Shafran, R., Rose, S., & Lee, M. J. (2015). A qualitative investigation of therapists' attitudes towards research: Horses for courses? *Behavioural and Cognitive Psychotherapy, 43*(4), 436–448. https://doi.org/10.1017/S1352465813001069

Herzog, P., & Kaiser, T. (2022). Is it worth it to personalize the treatment of PTSD? – A variance-ratio meta-analysis and estimation of treatment effect heterogeneity in RCTs of PTSD. *Journal of Anxiety Disorders, 91*, Article 102611. https://doi.org/10.1016/j.janxdis.2022.102611

Kaiser, T., Brakemeier, E. L., & Herzog, P. (2023). What if we wait? Using synthetic waiting lists to estimate treatment effects in routine outcome data. *Psychotherapy Research, 33*(8), 1043–1057. https://doi.org/10.1080/10503307.2023.2182241

Kazdin, A. E. (2019). Single-case experimental designs: Evaluating interventions in research and clinical practice. *Behaviour Research and Therapy, 117*, 3–17. https://doi.org/10.1016/j.brat.2018.11.015

Kirchner, L., Eckert, A., Berg, M., Endres, D., Straube, B., & Rief, W. (2022). *Better safe than sorry? An active inference approach to biased social inference in depression.* PsyArXiv Preprints. https://doi.org/10.31234/osf.io/bp9re

Lee, J., & Little, T. D. (2017). A practical guide to propensity score analysis for applied clinical research. *Behaviour Research and Therapy, 98*, 76–90. https://doi.org/10.1016/j.brat.2017.01.005

Lehmann, K., Maliske, L., Böckler, A., & Kanske, P. (2019). Social impairments in mental disorders: Recent developments in studying the mechanisms of interactive behavior. *Clinical Psychology in Europe, 1*(2), Article e33143. https://doi.org/10.32872/cpe.v1i2.33143

Lutz, W., Schwartz, B., & Delgadillo, J. (2022). Measurement-based and data-informed psychological therapy. *Annual Review of Clinical Psychology, 18*(1), 71–98. https://doi.org/10.1146/annurev-clinpsy-071720-014821

Martin-Soelch, C. (2023). The (neuro)-science behind resilience: A focus on stress and reward. *Clinical Psychology in Europe, 5*(1), Article e11567. https://doi.org/10.32872/cpe.11567

Mestdagh, M., Verdonck, S., Piot, M., Niemeijer, K., Kilani, G., Tuerlinckx, F., Kuppens, P., & Dejonckheere, E. (2023). m-Path: An easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice [Technology and code]. *Frontiers in Digital Health, 5*, Article 1182175. https://doi.org/10.3389/fdgth.2023.1182175

Moskow, D. M., Ong, C. W., Hayes, S. C., & Hofmann, S. G. (2023). Process-based therapy: A personalized approach to treatment. *Journal of Experimental Psychopathology, 14*(1), Article 20438087231152848. https://doi.org/10.1177/20438087231152848

Ormel, J., Hollon, S. D., Kessler, R. C., Cuijpers, P., & Monroe, S. M. (2022). More treatment but no less depression: The treatment-prevalence paradox. *Clinical Psychology Review, 91*, Article 102111. https://doi.org/10.1016/j.cpr.2021.102111

Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology, 66*(1), 126–135. https://doi.org/10.1037/0022-006X.66.1.126

Rief, W. (2021). Moving from tradition-based to competence-based psychotherapy. *BMJ Mental Health, 24*(3), 115–120. https://doi.org/10.1136/ebmental-2020-300219

Rief, W., Hofmann, S. G., Berg, M., Forbes, M. K., Pizzagalli, D. A., Zimmermann, J., Fried, E., & Reed, G. M. (2023). Do we need a novel framework for classifying psychopathology? A discussion paper. *Clinical Psychology in Europe, 5*(4), Article e11699. https://doi.org/10.32872/cpe.11699

Safran, J. D., Abreu, I., Ogilvie, J., & DeMaria, A. (2011). Does psychotherapy research influence the clinical practice of researcher–clinicians? *Clinical Psychology: Science and Practice, 18*(4), 357–371. https://doi.org/10.1111/j.1468-2850.2011.01267.x

Schaeuffele, C., Schulz, A., Knaevelsrud, C., Renneberg, B., & Boettcher, J. (2021). CBT at the crossroads: The rise of transdiagnostic treatments. *International Journal of Cognitive Therapy, 14*(1), 86–113. https://doi.org/10.1007/s41811-020-00095-2

Schemer, L., Hess, C. W., Van Orden, A. R., Birnie, K. A., Harrison, L. E., Glombiewski, J. A., & Simons, L. E. (2023). Enhancing exposure treatment for youths with chronic pain: Co-design and qualitative approach. *Journal of Participatory Medicine, 15*, Article e41292. https://doi.org/10.2196/41292

Slattery, P., Saeri, A. K., & Bragge, P. (2020). Research co-design in health: A rapid overview of reviews. *Health Research Policy and Systems, 18*(1), Article 17. https://doi.org/10.1186/s12961-020-0528-9

Speers, A. J. H., Bhullar, N., Cosh, S., & Wootton, B. M. (2022). Correlates of therapist drift in psychological practice: A systematic review of therapist characteristics. *Clinical Psychology Review, 93*, Article 102132. https://doi.org/10.1016/j.cpr.2022.102132

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

# Uncertainty Breeds Anxiety and Depression: The Impact of the Russian Invasion in Ukraine on a Swedish Clinical Population Receiving Internet-Based Psychotherapy

Jón Ingi Hlynsson [1], Oskar Gustafsson [2], Per Carlbring [1]

**[1]** *Department of Psychology, Stockholm University, Stockholm, Sweden.* **[2]** *Department of Statistics, Stockholm University, Stockholm, Sweden.*

Clinical Psychology in Europe, 2024, Vol. 6(1), Article e12083, https://doi.org/10.32872/cpe.12083

**Received:** 2023-05-31 • **Accepted:** 2024-01-15 • **Published (VoR):** 2024-03-28

**Handling Editor:** Cornelia Weise, Philipps-University of Marburg, Marburg, Germany

**Corresponding Author:** Jón Ingi Hlynsson, Department of Psychology, University of Iceland, Sæmundargata 12, 102 Reykjavík, Iceland. Phone: +354 774 50 50. E-mail: jih10@hi.is

**Supplementary Materials:** Materials [see Index of Supplementary Materials]

## Abstract

**Background:** Recent global crises, such as the COVID-19 pandemic and the 2022 Russian invasion of Ukraine, have contributed to a rise in the global prevalence of anxiety and depressive disorders. This study examines the indirect impact of the Ukraine war on emotional disorders within a Swedish clinical population.

**Method:** The sample comprised participants (n = 1,222) actively engaged in an internet-based psychotherapeutic intervention (cognitive-behavioral, psychodynamic, and waitlist) when the war broke out. The Patient Health Questionnaire-9 scale and the Generalized Anxiety Disorder-7 scale were used to measure depression and anxiety.

**Results:** Anxiety and depressive symptom severity increased following the war's onset, with an average weekly increase of 0.77-points for anxiety ($p = .001$, Cohen's $d = 0.08$) and 0.09-points for depression ($p = .70$, Cohen's $d = 0.01$); however, the increase was negligible for depression. Furthermore, higher socioeconomic status (SES) predicted declines in depression and anxiety during the study period, with a 0.69-point average weekly decrease in anxiety ($p < .001$, Cohen's $d = 0.32$) and a 1.09-point decrease in depression ($p < .001$, Cohen's $d = 0.48$) per one unit increase in SES, suggesting that SES may serve as a protective factor that buffers against psychopathological development during crises.

**Conclusions:** These findings have implications for mitigating the development of psychopathology during crises and interpreting treatment efficacy estimates during such events.

Our findings also emphasize the potential of internet-based psychotherapy in addressing emotional disorders during crises. This study presents up-to-date information about the reaction of treatment-seeking individuals to abrupt uncertainty.

## Keywords

anxiety, depression, Russian–Ukrainian war, uncertainty-inducing event, clinical trial, internet-based psychotherapy, emotional disorders

### Highlights

- The 2022 Russian invasion of Ukraine rapidly exacerbated anxiety symptom severity.
- Socioeconomic status may buffer against psychopathology during heightened uncertainty.
- Spatially distant uncertainty-inducing events can elevate the risk for psychopathology.
- Increased anxiety during crises may confound treatment efficacy estimations.

In recent years, the world has faced numerous global crises with devastating consequences for mental health. For instance, depression prevalence rose significantly after the 2008 global financial crisis (Guerra & Eboreime, 2021), and anxiety and depression rates worldwide increased by roughly 25% during the COVID-19 pandemic (Ettman et al., 2020; World Health Organization, 2022). Similarly, the Russian invasion of Ukraine on February 24th, 2022, resulted in increased prevalence rates of anxiety and depression among Ukrainians (Osokina et al., 2023; Xu et al., 2023) and Europeans (Riad et al., 2022; Skwirczyńska et al., 2022). Although these crises differ, they share a common characteristic: an increase in symptomatology of emotional disorders in response to an increase in external uncertainty.

Emotional disorders are characterized by frequent experiences of negative emotions, along with maladaptive reactions to and regulation of these experiences. These maladaptive reactions contribute to the persistence of negative emotions and the maintenance of the presenting disorder symptoms (cf. negative feedback loop; Bullis et al., 2019). Effectively managing uncertainty is already a critical adaptive challenge for humans. However, when environmental uncertainty abruptly increases, as during a global pandemic or war outbreak, adaptive information processing becomes even more hindered by internal disorder and uncertainty. These features, known as psychological entropy (Hirsh et al., 2012), tend to increase during crises, in turn, raising the likelihood of psychopathological development. For instance, anxiety and depressive symptoms were significantly higher during the COVID-19 pandemic compared to pre-pandemic rates (Gao et al., 2020; Xiong et al., 2020), with worldwide prevalence rates rising by 25% (Ettman et al., 2020; World Health Organization, 2022) and pandemic-related media exposure increased the odds of presenting with anxiety and combined anxiety and depression (Gao et al., 2020). Similarly, economic recessions (e.g., the 2008 global financial crisis) are associated with

an overall increase in depression and anxiety prevalence rates, with low socioeconomic status as a significant risk factor (Frasquilho et al., 2016; Gili et al., 2013; Guerra & Eboreime, 2021).

Focusing on the recent[1] war outbreak in Ukraine, a study by Riad et al. (2022) found that Czech university students reported high levels of concern about the ongoing conflict, with increased age correlating with higher levels of concern and media exposure engagement predicting anxiety and depression severity. Similarly, Skwirczyńska et al. (2022) discovered a positive association between war-related fear and anxiety severity in a Polish student sample. Intriguingly, access to monetary savings emerged as a protective factor that reduced the odds of presenting anxiety symptoms. One interpretation of Skwirczyńska et al.'s (2022) findings is that socioeconomic status, as indicated by access to monetary savings, buffers against anxiety symptom development (cf. Guerra & Eboreime, 2021). In summary, the war outbreak in Ukraine has noticeably affected the European population.

## Internet-Based Therapy

In recent years, a disparity has emerged between the demand for psychotherapy and its availability. As a result, the utilization of internet-based psychotherapeutic interventions has risen substantially to address this gap (Andersson et al., 2019). Internet-based psychotherapeutic treatments leverage technological advancements to create a contemporary alternative to traditional therapy. Typically, internet-based therapy consists of structured, manualized psychotherapy delivered online through modules containing self-help texts and the option to communicate with a therapist via encrypted messages (Andersson & Carlbring, 2022). Designed to parallel conventional face-to-face therapy in length and content (Andersson et al., 2016), internet-based therapy demonstrates equivalent overall therapeutic efficacy (Hedman-Lagerlöf et al., 2023). Meta-analytic findings support the treatment efficacy of internet-based therapy for emotional disorders, revealing moderate to large effect sizes for anxiety and depressive disorders (Andersson et al., 2019; Hedman-Lagerlöf et al., 2023).

## Aim of the Present Study

This study aims to assess the effects of indirect experiences of the war outbreak in Ukraine on the severity of anxiety and depression among individuals seeking treatment through an internet-based intervention, hereafter collectively referred to as "treatment-seeking individuals". Although this study was conducted in Sweden, which is a neighboring country but not directly bordering Ukraine (i.e., approximately 1500 kilometers

1) It should be noted that tensions between Russia and Ukraine began in 2014, but escalated into a full-blown war in February 2022, following Russia's invasion of Ukraine (cf. Michailova, 2022).

separate Sweden and Ukraine), previous studies suggest that the war outbreak in Ukraine has increased the prevalence rates of anxiety and depression in the general European population (Riad et al., 2022; Skwirczyńska et al., 2022). Indeed, surges in exposures to psychological threats (e.g., media exposure to crisis-related content) can jeopardize individuals' sense of personal security and exacerbate psychopathological development (Gao et al., 2020; Jayuphan et al., 2020; Riad et al., 2022). Consequently, we predicted a divergence in weekly therapeutic efficacy trends among treatment-seeking individuals following the war outbreak, as indicated by a spike in anxiety and depression. To our knowledge, this is the first study investigating the effects of the war in Ukraine on emotional disorders in a clinical population and thereby aims to provide up-to-date information about the reaction of treatment-seeking individuals to abrupt uncertainty.

## Hypotheses

This study has two core hypotheses: Scores on the 1) PHQ-9 and 2) GAD-7 will be significantly elevated following the outbreak of war in Ukraine when compared to a baseline established by the trend in scores observed over the preceding four weeks, adjusting for treatment group assignment, socioeconomic status, education level, age, and gender. Additionally, high socioeconomic status is hypothesized to be a protective factor that buffers against further development of psychopathology following the war outbreak.

# Method

## Participants and Recruitment

The present study utilizes data from an ongoing clinical trial (ClinicalTrials.gov identifier: NCT05016843) that is being conducted in Sweden. Participants were recruited online through a website outlining the study's aims and components (Vlaescu et al., 2016). The study was advertised on Facebook and also spread through word of mouth. Participants did not receive any monetary compensation for their involvement in the study. The only form of compensation provided was the inherent benefits derived from participation in the treatment interventions. See Figure S1, Supplementary Materials, for a flow chart illustration of the study design.

### Sample Size

All participants ($n$ = 1,222) actively engaged in the study between January 24th, 2022, and March 24th, 2022, were included. This two-month period was chosen to adequately represent treatment efficacy before and after the war outbreak in Ukraine on February 24th, 2022.

### Eligibility Criteria

Eligibility criteria were assessed during the study's screening phase. Participants were required to: a) be at least 18 years of age; b) read and write in Swedish; c) have an internet connection via their mobile phone or computer; and d) experience at least mild anxiety symptoms (i.e., GAD-7 ≥ 5 points) or mild to moderate depression symptoms (i.e., PHQ-9 ≥ 10 points), or both. Participants were excluded if they: a) were currently seeking other psychological treatment; b) had begun or adjusted psychopharmacological treatment for anxiety, worry, or depression within the nearest month from screening; or c) had severe depression (i.e., PHQ-9 ≥ 20 points) or suicidality (i.e., PHQ-9, item nine score > 2 points) indicated during screening.

## Measures

Demographic variables and anxiety and depression measurements were collected during screening, followed by weekly measurements of anxiety and depression.

### Demographics

Demographic variables collected during screening included age, gender, socioeconomic status[2], marital status, household composition, level of education, employment status, mental health characteristics, and prior psychopharmaceutical medication usage.

### Patient Health Questionnaire 9-Item Scale (PHQ-9)

The Patient Health Questionnaire 9-item scale (PHQ-9) is a self-report questionnaire that quantifies depression severity (Kroenke et al., 2001). Each item is rated on a scale from 0 to 3, with total scores ranging from 0 to 27. A score of 10 or higher is a diagnostic indicator of depression (Kroenke et al., 2001, 2010). The PHQ-9 consistently demonstrates good accuracy and discrimination ability in clinical settings and the general population (Kocalevent et al., 2013; Kroenke et al., 2001, 2010) as well as when administered via the internet (Martin-Key et al., 2022). In this study, the PHQ-9 exhibited adequate internal reliability during screening, Cronbach's alpha = 0.66, 95% CI [0.63, 0.68], indicating acceptable internal consistency. It should be noted that this internal consistency reliability estimate suffers from a restriction of range and an analysis of the whole sample at screening (both included and excluded participants) yielded Cronbach's alpha = 0.81, 95% CI [0.80, 0.62] (Hlynsson & Carlbring, 2023).

2) Socioeconomic status was indirectly measured with a self-rated scale; participants rated their socioeconomic status in relation to others on a scale from 1 to 5 (see Table 1 for response options).

### Generalized Anxiety Disorder 7-Item Scale (GAD-7)

The Generalized Anxiety Disorder 7-item scale (GAD-7) is a self-report questionnaire that assesses anxiety and screens for generalized anxiety disorder (Spitzer et al., 2006). Each item is rated on a scale from 0 to 3, with total scores ranging from 0 to 21. A score of 8 or higher is a diagnostic indicator of anxiety disorders (Luo et al., 2019; Spitzer et al., 2006). The items align with DSM-5 criteria (American Psychiatric Association, 2022) and are sensitive to various anxiety disorders (Kroenke et al., 2010) in both clinical settings and the general population, as well as when administered online (Byrd-Bredbenner et al., 2021; Johnson et al., 2019; Löwe et al., 2008; Martin-Key et al., 2022). In this study, the GAD-7 demonstrated good internal reliability during screening, Cronbach's alpha = 0.77, 95% CI [0.75, 0.79], indicating excellent internal consistency. It should be noted that this internal consistency reliability estimate suffers from a restriction of range and an analysis of the whole sample at screening (both included and excluded participants) yielded Cronbach's alpha = 0.85, 95% CI [0.83, 0.85] (Hlynsson & Carlbring, 2023).

## Treatment Interventions

Data was collected as part of an ongoing clinical trial (Mechler et al., 2022) comparing cognitive-behavioral therapy (unified protocol [UP]; Barlow et al., 2017) with psychodynamic affective phobia (AP) therapy (Julien & O'Connor, 2017). The trial comprised three factors: a) type of internet-based treatment intervention; b) treatment length; and c) effects of access to a clinician-moderated discussion forum. Participants were randomly assigned via a factorial assignment mechanism to one of twelve conditions: UP, AP, or a waitlist, each for either 8 or 16 weeks, and each with or without access to a clinician-moderated forum.

## Data Analysis

The data was analyzed using R Studio (R Core Team, 2021). A panel-data regression analysis was conducted, in which PHQ-9 and GAD-7 scores were separately predicted by the treatment time course in weeks (e.g., data provided between January 24th and January 30th, 2022, was assigned the number 1 corresponding to week one) and a dummy variable containing information about whether data corresponded to the time period before or after the war outbreak (i.e., all data corresponding to dates before February 24th, 2022, was coded as 0 and other data as 1), while adjusting for relevant covariates. In addition, Cohen's *d* effect sizes were computed to interpret the magnitude of all associations. Hemphill's (2003) interpretive framework for effect sizes, derived from an empirical assessment of the magnitude of the average effect sizes produced in psychological studies, was used to interpret effect size magnitudes. The correlational effect size guidelines provided by Hemphill (2003) were converted into Cohen's *d* effect sizes (Ruscio, 2008). Cohen's *d* effect sizes below 0.4 were considered small in magnitude,

effect sizes between 0.4 and 0.6 were considered moderate, and effect sizes above 0.6 were considered large.

To preserve power and minimize missing data, participants were only compared during the first 8 weeks of treatment/waitlist. This is because data was only collected for half of the participants for 8 weeks (i.e., participants were either assigned to 8 or 16 weeks, and thus observations corresponding to weeks 9-16 would be missing for half of the sample due to the study design). A separate analysis wherein only participants assigned to a 16-week treatment intervention was conducted to corroborate the findings of the present analysis (see Table S1, Supplementary Materials). Moreover, since data was stratified by treatment group assignment and the experiment was conducted over several weeks, a heteroscedasticity and autocorrelation consistent (HAC) covariance matrix estimation was used to obtain a robust estimation of the linear models' standard errors (Cribari-Neto & da Silva, 2011).

Additionally, due to a large amount of missing observations in the dataset (i.e., 53% of observations for depression and anxiety), the data was also modeled using a Full Information Maximum Likelihood (FIML) estimation (cf. Hesser, 2015; Hoffart et al., 2022). FIML estimation allows for parameter estimates despite missing data by estimating patterns of missingness (Baraldi & Enders, 2010). This additional analysis was conducted to assess the convergence between FIML estimation and HAC covariance matrix estimation (i.e., compare the results obtained from the two methods). Isomorphic parameter estimates from both methods (i.e., in terms of signs and significance) will be taken as indicators of a stable and generalizable parameter estimation. In an effort to approach a model that might suggest potential causal effects of the war outbreak on anxiety and depressive symptom severity, all variables considered relevant were included in the analysis (Rohrer, 2018). A directed acyclic graph of the hypothesized causal associations and interdependencies in the assumed data-generating process was constructed using DAGitty to guide the choice of variables to adjust and not to adjust for in the present analysis (see Figure S2, Supplementary Materials; Textor et al., 2016).

# Results

## Sample Characteristics

Descriptive statistics for the sample demographics are summarized in Table 1.

During screening, PHQ-9 scores ranged from 1 to 19 ($M$ = 11.76, $SD$ = 4.16), and GAD-7 scores ranged from 0 to 21 ($M$ = 9.74, $SD$ = 4.16). In the four weeks leading up to the war outbreak, PHQ-9 scores ranged from 0 to 27 ($M$ = 8.84, $SD$ = 5.14), and GAD-7 scores ranged from 0 to 21 ($M$ = 7.79, $SD$ = 4.80). In the four weeks following the war outbreak, PHQ-9 scores ranged from 0 to 27 ($M$ = 8.26, $SD$ = 5.36), and GAD-7 scores ranged from 0 to 21 ($M$ = 7.83, $SD$ = 5.14).

**Table 1**

*Demographical Descriptive Statistics*

| Participant Characteristics | Waitlist[a], *n* = 560 | Psychodynamic Affect Phobia Therapy[b], *n* = 348 | Cognitive Behavior Therapy[c], *n* = 314 | Total, *n* = 1,222 |
|---|---|---|---|---|
| **Age** | 43 (12) | 43 (12) | 44 (13) | 42 (12) |
| **Education** | | | | |
| Elementary School | 14 (2.5%) | 13 (3.7%) | 7 (2.2%) | 34 (2.8%) |
| High School | 128 (23%) | 89 (26%) | 84 (27%) | 301 (25%) |
| College-level education (< 3 years) | 155 (28%) | 96 (28%) | 80 (25%) | 331 (27%) |
| College-level education (> 3 years) | 263 (47%) | 150 (43%) | 143 (46%) | 556 (45%) |
| **Sex** | | | | |
| Female | 483 (86%) | 302 (87%) | 270 (86%) | 1,055 (86%) |
| Male | 73 (13%) | 45 (13%) | 43 (14%) | 161 (13%) |
| Other | 4 (0.7%) | 1 (0.3%) | 1 (0.3%) | 6 (0.5%) |
| **Self-rated socioeconomic status** | | | | |
| Much worse than others | 23 (4.1%) | 20 (5.7%) | 12 (3.8%) | 55 (4.5%) |
| Worse than others | 132 (24%) | 88 (25%) | 60 (19%) | 280 (23%) |
| About the same as others | 234 (42%) | 149 (43%) | 145 (46%) | 528 (43%) |
| Better than others | 152 (27%) | 86 (25%) | 80 (25%) | 318 (26%) |
| Much better than others | 19 (3.4%) | 5 (1.4%) | 17 (5.4%) | 41 (3.4%) |
| **Children under 18 in the house** | | | | |
| No | 347 (62%) | 195 (56%) | 191 (61%) | 733 (60%) |
| Yes | 206 (37%) | 144 (41%) | 116 (37%) | 466 (38%) |
| Complicated/Sometimes | 7 (1.2%) | 9 (2.6%) | 7 (2.2%) | 23 (1.9%) |
| **Prior medication for anxiety/depression** | 150 (27%) | 85 (24%) | 77 (25%) | 312 (26%) |
| **Current occupation** | | | | |
| Working | 394 (70%) | 226 (65%) | 220 (70%) | 840 (69%) |
| Studying | 71 (13%) | 48 (14%) | 45 (14%) | 164 (13%) |
| Seeking work | 32 (5.7%) | 19 (5.5%) | 11 (3.5%) | 62 (5.1%) |
| Retired | 26 (4.6%) | 20 (5.7%) | 15 (4.8%) | 61 (5.0%) |
| Parental leave | 5 (0.9%) | 8 (2.3%) | 2 (0.6%) | 15 (1.2%) |
| Sick leave | 32 (5.7%) | 27 (7.8%) | 21 (6.7%) | 80 (6.5%) |

[a]Aggregated from four groups: Waitlist for 8 weeks, with discussion forum access (*n* = 126), Waitlist for 8 weeks, no discussion forum access (*n* = 122), Waitlist for 16 weeks, with discussion forum access (*n* = 154), Waitlist for 16 weeks, no discussion forum access (*n* = 158).

[b]Aggregated from four groups: Affect Phobia for 8 weeks, with discussion forum access (*n* = 59), Affect Phobia for 8 weeks, no discussion forum access (*n* = 61), Affect Phobia for 16 weeks, with discussion forum access (*n* = 111), Affect Phobia for 16 weeks, no discussion forum access (*n* = 117).

[c]Aggregated from four groups: Unified Protocol for 8 weeks, with discussion forum access (*n* = 46), Unified Protocol for 8 weeks, no discussion forum access (*n* = 58), Unified Protocol for 16 weeks, with discussion forum access (*n* = 100), Unified Protocol for 8 weeks, no discussion forum access (*n* = 110).

### Missing Data

For the eight instances when data was provided, a Fisher's exact test comparing the propensity for data being differentially missing between the first four and latter four instances revealed non-significant differences for both the PHQ-9 ($p$ = .168) and GAD-7 ($p$ = .204). Furthermore, no obvious trends of missingness were discernible as a function of age, gender, or SES.

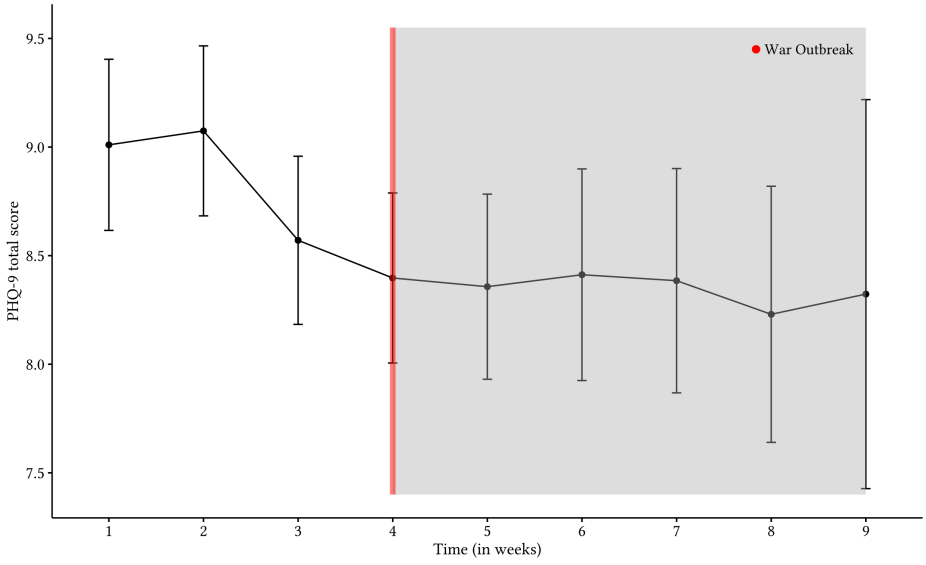
## The Effects of the War Outbreak

### Symptoms of Depression in Response to the Outbreak

The outbreak of war did not significantly increase average levels of depression. Scores on the PHQ-9 slightly increased following the war outbreak, $t(4566) = 0.39$, $p$ = .699, wherein comparing two individuals of the same socioeconomic status, treatment group, education level, age, and gender, while adjusting for the date on which data was provided, revealed a 0.09-point increase in average levels of depression, 95% CI [-0.38, 0.56]; effect size: $d$ = 0.01, following the outbreak of war. The data was most compatible with values ranging from a 0.38-point decrease to a 0.56-point increase in scores on the PHQ-9. As such, the results do not indicate that the war outbreak significantly affected the severity of depression (see Figure 1).

**Figure 1**

*Graphical Depiction of Unadjusted Raw-Mean Scores and 95% Confidence Intervals for Depression Each Week, Over the Course of Treatment for All Treatment Groups*

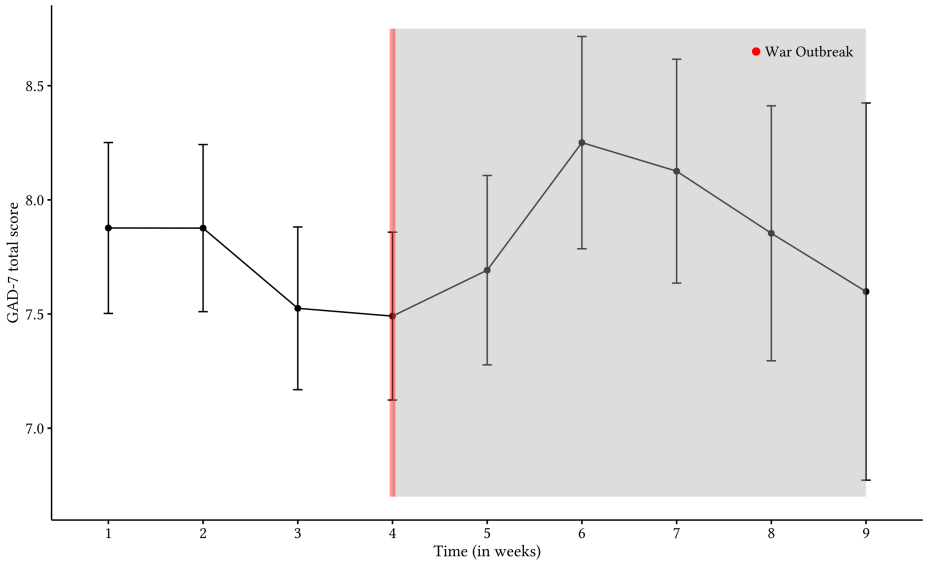


### Symptoms of Anxiety in Response to the Outbreak

The war outbreak significantly increased average anxiety levels. Anxiety scores on the GAD-7 rose following the outbreak, $t(4566) = 3.23$, $p = .001$. Comparing two individuals with the same socioeconomic status, treatment group, education level, age, and gender, and adjusting for the data collection date, a 0.77-point increase in anxiety severity, 95% CI [0.30, 1.23]; effect size: $d = 0.08$, was observed after the war outbreak. The data was most compatible with values ranging from a 0.30-point to a 1.23-point increase in GAD-7 scores. A general decline in anxiety symptom severity was detected prior to the war outbreak which then increased abruptly in the wake of the war outbreak before rapidly declining to pre-war outbreak levels (see Figure 2). Consequently, the results suggest that the war outbreak exacerbated anxiety severity.

**Figure 2**

*Graphical Depiction of Unadjusted Raw-Mean Scores and 95% Confidence Intervals for Anxiety Each Week, Over the Course of Treatment for All Treatment Groups*




### Socioeconomic Status as a Protective Factor

Socioeconomic status was inversely associated with anxiety severity over time, $t(4566) = -3.61$, $p < .001$, during this study. When comparing two individuals on the same date, within the same treatment group, of the same age, gender, and education level, while adjusting for the outbreak of war, a 1-point increase in self-rated socioeconomic status was associated with a 0.69-point average decrease in scores on the GAD-7, 95% CI [-1.06, -0.31]; effect size: $d = 0.32$. The data was most compatible with values ranging from

a 1.06-point decrease to a 0.31-point decrease in scores on the GAD-7. Thus, anxiety symptom severity is, on average, lower for people with relatively higher socioeconomic status when controlling for the time course of treatment, war outbreak, gender, and treatment group, in turn, suggesting that socioeconomic status may be a potential protective factor for anxiety symptoms during a war outbreak (cf. entropy increase). Adding an interaction term between the war outbreak dummy variable and socioeconomic status did not increase the model fit nor alter the coefficient estimates.

Socioeconomic status was also inversely associated with depression severity over time, $t(4566) = -5.28$, $p < .001$, during this study. When comparing two individuals on the same date, within the same treatment group, of the same age, gender, and education level, while adjusting for the outbreak of war, a 1-point increase in self-rated socioeconomic status was associated with a 1.09-point average decrease in scores on the PHQ-9, 95% CI [-1.49, -0.68]; effect size: $d = 0.48$. The data was most compatible with values ranging from a 1.49-point decrease to a 0.68-point decrease in scores on the PHQ-9. Thus, depressive symptom severity is, on average, lower for people with relatively higher socioeconomic status when controlling for the time course of treatment, war outbreak, gender, and treatment group, in turn, suggesting that socioeconomic status may be a potential protective factor for depressive symptoms during a war outbreak (cf. entropy increase). Adding an interaction term between the war outbreak dummy variable and socioeconomic status did not increase the model fit nor alter the coefficient estimates.

## Additional Analyses

### Full Information Maximum Likelihood (FIML) Estimation

To further support the previously reported results, linear models for the PHQ-9 and GAD-7 were analyzed using FIML estimations (see Table S2, Supplementary Materials). This analysis produced parameter estimates that were consistent with HAC covariance matrix estimation results reported earlier (i.e., equivalent parameter estimates and *p*-values). Moreover, an additional analysis that adjusted for all background variables at our disposal also produced parameter estimates that were consistent with both the HAC covariance matrix estimation and FIML results. Taken together, the parameter estimates seem stable in the current analysis, and patterns of missing data do not appear to significantly impact the results.

### Treatment Group and Treatment Efficacy Analyses

Analyses of the differential effects of the war outbreak and overall treatment efficacy were conducted (see Supplementary Materials).

## Discussion

The present study aimed to elucidate the effects of the outbreak of war in Ukraine following the Russian invasion on February 24th on measures of anxiety and depressive symptom severity. To our knowledge, this is the first study on the indirect effects of the war in Ukraine on emotional disorders in a clinical population, thereby providing up-to-date information about the reaction of treatment-seeking individuals to uncertainty-inducing events. The results indicate that anxiety symptoms significantly increased in response to the war outbreak, as predicted, although this effect was small in magnitude (cf. Hemphill, 2003). Anxiety symptom severity generally declined before the outbreak of war, spiked following the war outbreak, before rapidly declining to pre-war outbreak levels. However, contrary to our hypothesis, the war outbreak had a negligible effect on depressive symptoms. Depressive symptoms gradually declined throughout the duration of the study and did not spike in response to the war outbreak. Finally, socioeconomic status had a moderate effect on decreased anxiety symptoms and decreased depressive symptoms over the course of treatment, irrespective of the war outbreak. These findings thus provide support for the notion that socioeconomic status serves as a protective factor against psychopathology in times of heightened uncertainty.

The finding that anxiety symptom severity increased in response to the war outbreak, but depressive symptom severity did not, may relate to how anxiety and depression are differentially associated with intolerance of uncertainty. As noted in the introduction, intolerance of uncertainty, which may underpin many psychopathological impairments to daily functioning, has been suggested to be more pronounced in anxiety disorders than depression (Jensen et al., 2016). However, meta-analytic findings suggest that intolerance of uncertainty lacks etiological specificity to differentiate anxiety and depression (Gentes & Ruscio, 2011). Nonetheless, the semantic link between anxiety and intolerance of uncertainty is reflected in the American Psychiatric Association's (2022, p. 215) definition of anxiety as the "anticipation of [a] future threat," which coincides with the definition of intolerance of uncertainty (i.e., responding to uncertainty-inducing events with discomfort and anxiety which, in turn, further increases negative affectivity, cf. psychological entropy; Hirsh et al., 2012; Jensen et al., 2016). Furthermore, even though the effect of the war outbreak on anxiety symptoms is small in magnitude by most statistical standards, it is important to consider the clinical implications of uncertainty-inducing events on anxiety symptoms within a treatment-seeking population and place the effect in a broader context. For instance, the magnitude of the effect between increased anxiety symptoms in response to the war outbreak is slightly larger than the association between aspirin consumption and heart attack prevention (Rosenthal, 1991, p. 136; see also Hemphill, 2003). Moreover, this effect size mirrors typical effect sizes that research on the effect of disasters on mental health disorders produces, where pooled effect estimates range from 0.05 and 0.20 (Keya et al., 2023).

The present study has limitations. In line with previous studies (e.g., Guerra & Eboreime, 2021; Skwirczyńska et al., 2022), we found socioeconomic status to buffer against psychopathological development following the abrupt increase in external uncertainty due to the war outbreak. However, the interpretation of this effect may be limited by using self-reported socioeconomic status, where participants self-rated their socioeconomic status in relation to others. Another limitation is our lack of control for media exposure. Previous studies indicate frequency of media exposure to covary with anxiety and depression symptom severity (Gao et al., 2020; Riad et al., 2022). As such, without control for participant exposure to media coverage of the war, effects of the war outbreak on anxiety and depression symptom severity may have been attenuated (or even augmented).

Furthermore, this study is limited by a lack of qualitative interviews to provide insight into participant's experiences and perceptions of the war outbreak and its effects on their mental health. Future studies could ameliorate this limitation by incorporating an ecological momentary assessment protocol (e.g., Verhagen et al., 2022), wherein data on exposure to war-related media and self-reported affectedness of the war outbreak is collected with high frequency concomitantly with indices of anxiety and depression. No clinical interviews were conducted to accurately detect whether participants qualified for a diagnosis of an anxiety or depressive disorder. However, only treatment-seeking participants with scores indicative of an emotional disorder were included in the study, and the PHQ-9 and the GAD-7 routinely emerge as good indicators of depressive and anxiety disorders (Byrd-Bredbenner et al., 2021; Johnson et al., 2019; Martin-Key et al., 2022). Additionally, this study is limited by design; temporal precedence was established but true causality cannot be inferred from the present analysis.

Finally, the large number of missing observations in the measures of anxiety and depression severity somewhat limits the statistical analyses. The possibility that participants selectively neglected to provide data when they suffered most severely from depression and/or anxiety cannot be eliminated. However, there were no discernible trends in the missingness of data. Moreover, modelling the data with state-of-the-art statistical procedures for handling missing data (i.e., FIML and robust HAC versions of the general linear model) did not influence the statistical conclusion of the results as it yielded isomorphic parameter estimations.

The present study has numerous strengths. Firstly, this study is the first analysis of the impact of the war in Ukraine in a clinical sample, and thus provides up-to-date information about the reaction of treatment-seeking individuals to abrupt uncertainty. Additionally, although greater average variability in indicators of depression and anxiety is to be expected in clinical samples (Hirsh et al., 2012; Sauer-Zavala & Barlow, 2021), a clear upward spike in average levels of depression and anxiety severity in response to the war outbreak was discernible. Secondly, measures of anxiety and depression severity were obtained weekly throughout the treatment intervention, allowing for a representa-

tive estimation of the psychopathological response to the war outbreak. Thirdly, given that psychopathological development surges in response to abrupt uncertainty-inducing events (cf. entropy increase; Guerra & Eboreime, 2021; Hirsh et al., 2012; Lim et al., 2022; Osokina et al., 2023; Riad et al., 2022), our study may have buffered psychopathological development among Swedish treatment-seeking individuals. Other strengths include the exclusive inclusion of treatment-seeking individuals and an adequately large sample size.

The present study may have implications for how abrupt uncertainty-inducing events can be mitigated at a population level. Briefly, our results suggest that anxiety symptom severity rises in conjunction with increased environmental uncertainty (cf. entropy increase); a particularly interesting finding considering the geographical distance between Sweden and Ukraine, which exceeds 1500 km. The study underscores the need for heightened vigilance and support for individuals predisposed to psychopathology when confronted with sudden, uncertainty-inducing events, irrespective of their physical proximity. However, it is important to approach these findings with caution. The study did not directly measure participants' perceptions of the war outbreak or ascertain which specific aspects of the conflict were most impactful to them. Given this limitation, the direct influence of the war outbreak on the observed increase in anxiety symptoms remains speculative. Nevertheless, providing readily accessible health care services, such as government-funded internet-based psychotherapy, in the aftermath of such events could be beneficial. This approach may help alleviate societal impacts and reduce the overall burden of such events, particularly for individuals with below-average socioeconomic status who might encounter additional challenges in the wake of uncertainty-inducing events.

Finally, this study holds implications for clinicians in practice. It suggests that when psychotherapy is provided during crises, a sudden increase in anxiety symptoms can, in general, be expected in response to heightened environmental uncertainty (cf. entropy increase). However, statistically controlling for this crisis-related increase reveals that overall severity of anxiety symptoms continues to decrease throughout the course of treatment. As such, an increase in anxiety symptoms during crises situations should not automatically be interpreted as an indicator of unsuccessful treatment. Instead, it should be recognized as a potential confounding factor in estimating treatment efficacy. Furthermore, this effect differed between treatment group assignments (see Figures S3 and S4, Supplementary Materials).

## Conclusion

The present study highlights the impact of the Ukrainian war outbreak on emotional disorders, particularly anxiety symptoms, in a clinical population. Anxiety symptom severity seems to be sensitive to the conflict's influence, experiencing an increase of up to 1.22 points following the war outbreak. Moreover, socioeconomic status may serve as a protective factor against the development of psychopathological disorders in

the wake of uncertainty-inducing events. Lastly, this study reinforces previous findings demonstrating the effectiveness of internet-based psychotherapeutic interventions in alleviating emotional disorder symptoms.

**Funding:** We declare that no financial support was received for the research, authorship, and/or publication of this article.

**Acknowledgments:** The authors have no additional (i.e., non-financial) support to report.

**Competing Interests:** The authors have declared that no competing interests exist.

**Ethics Statement:** The study was approved in 2021 by the Swedish Ethical Review Authority (Dnr 2021-00034) and again in 2022 following an addendum to the initial proposal outlining our intent to analyze data related to the war (Dnr 2022-01362-02).

**Preregistration:** This study was not preregistered.

**Twitter Accounts:** @JonHlynsson

**Data Availability:** The data that support the findings of this study are available from the corresponding author, JIH, upon reasonable request.

## Supplementary Materials

The Supplementary Materials include the following items (see Hlynsson et al., 2024):

- A flow chart illustrating the study design.
- An analysis featuring participants exclusively assigned to a 16-week treatment intervention to corroborate the findings reported in this paper.
- A Directed Acyclic Graph illustrating the hypothesized causal model and associations between variables.
- Full Information Maximum Likelihood (FIML) estimates contrasted with the heteroscedasticity and autocorrelation consistent (HAC) covariance matrix estimation.
- An analysis of the effect of treatment group assignment during the study period.
- An assessment of overall treatment efficacy during the study period.

### Index of Supplementary Materials

Hlynsson, J. I., Gustafsson, O., & Carlbring, P. (2024). *Supplementary materials to "Uncertainty breeds anxiety and depression: The impact of the Russian invasion in Ukraine on a Swedish clinical population receiving internet-based psychotherapy"* [Additional information]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.14140

## References

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (DSM-5-TR). American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787

Andersson, G., & Carlbring, P. (2022). 6.12—Internet interventions in clinical psychology. In G. J. G. Asmundson (Ed.), *Comprehensive clinical psychology* (2nd ed., pp. 194–205). Elsevier. https://doi.org/10.1016/B978-0-12-818697-8.00211-9

Andersson, G., Carlbring, P., & Lindefors, N. (2016). History and current status of ICBT. In N. Lindefors & G. Andersson (Eds.), *Guided Internet-based treatments in psychiatry* (pp. 1–16). Springer International. https://doi.org/10.1007/978-3-319-06083-5_1

Andersson, G., Carlbring, P., Titov, N., & Lindefors, N. (2019). Internet interventions for adults with anxiety and mood disorders: A narrative umbrella review of recent meta-analyses. *Canadian Journal of Psychiatry, 64*(7), 465–470. https://doi.org/10.1177/0706743719839381

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 48*(1), 5–37. https://doi.org/10.1016/j.jsp.2009.10.001

Barlow, D. H., Farchione, T. J., Sauer-Zavala, S., Murray Latin, H., Ellard, K. K., Bullis, J. R., Bentley, K. H., Boettcher, H. T., & Cassiello-Robbins, C. (2017). *Unified Protocol for transdiagnostic treatment of emotional disorders: Therapist guide* (2nd ed.). Oxford University Press. https://doi.org/10.1093/med-psych/9780190685973.001.0001

Bullis, J. R., Boettcher, H., Sauer-Zavala, S., Farchione, T. J., & Barlow, D. H. (2019). What is an emotional disorder? A transdiagnostic mechanistic definition with implications for assessment, treatment, and prevention. *Clinical Psychology: Science and Practice, 26*(2), Article e12278. https://doi.org/10.1037/h0101755

Byrd-Bredbenner, C., Eck, K., & Quick, V. (2021). GAD-7, GAD-2, and GAD-mini: Psychometric properties and norms of university students in the United States. *General Hospital Psychiatry, 69*, 61–66. https://doi.org/10.1016/j.genhosppsych.2021.01.002

Cribari-Neto, F., & da Silva, W. B. (2011). A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model. *AStA Advances in Statistical Analysis, 95*(2), 129–146. https://doi.org/10.1007/s10182-010-0141-2

Ettman, C. K., Abdalla, S. M., Cohen, G. H., Sampson, L., Vivier, P. M., & Galea, S. (2020). Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. *JAMA Network Open, 3*(9), Article e2019686. https://doi.org/10.1001/jamanetworkopen.2020.19686

Frasquilho, D., Matos, M. G., Salonna, F., Guerreiro, D., Storti, C. C., Gaspar, T., & Caldas-de-Almeida, J. M. (2016). Mental health outcomes in times of economic recession: A systematic literature review. *BMC Public Health, 16*(1), Article 115. https://doi.org/10.1186/s12889-016-2720-y

Gao, J., Zheng, P., Jia, Y., Chen, H., Mao, Y., Chen, S., Wang, Y., Fu, H., & Dai, J. (2020). Mental health problems and social media exposure during COVID-19 outbreak. *PLoS One, 15*(4), Article e0231924. https://doi.org/10.1371/journal.pone.0231924

Gentes, E. L., & Ruscio, A. M. (2011). A meta-analysis of the relation of intolerance of uncertainty to symptoms of generalized anxiety disorder, major depressive disorder, and obsessive–compulsive disorder. *Clinical Psychology Review, 31*(6), 923–933. https://doi.org/10.1016/j.cpr.2011.05.001

Gili, M., Roca, M., Basu, S., McKee, M., & Stuckler, D. (2013). The mental health risks of economic crisis in Spain: Evidence from primary care centres, 2006 and 2010. *European Journal of Public Health, 23*(1), 103–108. https://doi.org/10.1093/eurpub/cks035

Guerra, O., & Eboreime, E. (2021). The impact of economic recessions on depression, anxiety, and trauma-related disorders and illness outcomes—A scoping review. *Behavioral Sciences, 11*(9), Article 119. https://doi.org/10.3390/bs11090119

Hedman-Lagerlöf, E., Carlbring, P., Svärdman, F., Riper, H., Cuijpers, P., & Andersson, G. (2023). Therapist-supported Internet-based cognitive behaviour therapy yields similar effects as face-to-face therapy for psychiatric and somatic disorders: An updated systematic review and meta-analysis. *World Psychiatry, 22*(2), 305–314. https://doi.org/10.1002/wps.21088

Hemphill, J. F. (2003). Interpreting the magnitudes of correlation coefficients. *The American Psychologist, 58*(1), 78–79. https://doi.org/10.1037/0003-066X.58.1.78

Hesser, H. (2015). Modeling individual differences in randomized experiments using growth models: Recommendations for design, statistical analysis and reporting of results of internet interventions. *Internet Interventions, 2*(2), 110–120. https://doi.org/10.1016/j.invent.2015.02.003

Hirsh, J. B., Mar, R. A., & Peterson, J. B. (2012). Psychological entropy: A framework for understanding uncertainty-related anxiety. *Psychological Review, 119*(2), 304–320. https://doi.org/10.1037/a0026767

Hlynsson, J. I., & Carlbring, P. (2023). *Diagnostic accuracy and clinical utility of the PHQ-2 and GAD-2: A comparison with long-format measures for depression and anxiety* [Unpublished manuscript]. Stockholm University.

Hoffart, A., Bauer, D. J., Johnson, S. U., & Ebrahimi, O. V. (2022). Anxiety in the adult population from the onset to termination of social distancing protocols during the COVID-19: A 20-month longitudinal study. *Scientific Reports, 12*(1), Article 1. https://doi.org/10.1038/s41598-022-22686-z

Jayuphan, J., Sangthong, R., Hayeevani, N., Assanangkornchai, S., & McNeil, E. (2020). Mental health problems from direct vs indirect exposure to violent events among children born and growing up in a conflict zone of southern Thailand. *Social Psychiatry and Psychiatric Epidemiology, 55*(1), 57–62. https://doi.org/10.1007/s00127-019-01732-8

Jensen, D., Cohen, J. N., Mennin, D. S., Fresco, D. M., & Heimberg, R. G. (2016). Clarifying the unique associations among intolerance of uncertainty, anxiety, and depression. *Cognitive Behaviour Therapy, 45*(6), 431–444. https://doi.org/10.1080/16506073.2016.1197308

Johnson, S. U., Ulvenes, P. G., Øktedalen, T., & Hoffart, A. (2019). Psychometric properties of the general anxiety disorder 7-item (GAD-7) scale in a heterogeneous psychiatric sample. *Frontiers in Psychology, 10*, Article 1713. https://doi.org/10.3389/fpsyg.2019.01713

Julien, D., & O'Connor, K. P. (2017). Recasting psychodynamics into a behavioral framework: A review of the theory of psychopathology, treatment efficacy, and process of change of the

affect phobia model. *Journal of Contemporary Psychotherapy, 47*(1), 1–10. https://doi.org/10.1007/s10879-016-9324-9

Keya, T. A., Leela, A., Habib, N., Rashid, M., & Bakthavatchalam, P. (2023). Mental health disorders due to disaster exposure: A systematic review and meta-analysis. *Cureus, 15*(4), Article e37031. https://doi.org/10.7759/cureus.37031

Kocalevent, R.-D., Hinz, A., & Brähler, E. (2013). Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *General Hospital Psychiatry, 35*(5), 551–555. https://doi.org/10.1016/j.genhosppsych.2013.04.006

Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Kroenke, K., Spitzer, R. L., Williams, J. B. W., & Löwe, B. (2010). The patient health questionnaire somatic, anxiety, and depressive symptom scales: A systematic review. *General Hospital Psychiatry, 32*(4), 345–359. https://doi.org/10.1016/j.genhosppsych.2010.03.006

Lim, I. C. Z. Y., Tam, W. W. S., Chudzicka-Czupała, A., McIntyre, R. S., Teopiz, K. M., Ho, R. C., & Ho, C. S. H. (2022). Prevalence of depression, anxiety and post-traumatic stress in war- and conflict-afflicted areas: A meta-analysis. *Frontiers in Psychiatry, 13*, Article 978703. https://doi.org/10.3389/fpsyt.2022.978703

Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. *Medical Care, 46*(3), 266–274. https://doi.org/10.1097/MLR.0b013e318160d093

Luo, Z., Li, Y., Hou, Y., Zhang, H., Liu, X., Qian, X., Jiang, J., Wang, Y., Liu, X., Dong, X., Qiao, D., Wang, F., & Wang, C. (2019). Adaptation of the two-item generalized anxiety disorder scale (GAD-2) to Chinese rural population: A validation study and meta-analysis. *General Hospital Psychiatry, 60*, 50–56. https://doi.org/10.1016/j.genhosppsych.2019.07.008

Martin-Key, N. A., Spadaro, B., Funnell, E., Barker, E. J., Schei, T. S., Tomasik, J., & Bahn, S. (2022). The current state and validity of digital assessment tools for psychiatry: Systematic review. *JMIR Mental Health, 9*(3), Article e32824. https://doi.org/10.2196/32824

Mechler, J., Lindqvist, K., Carlbring, P., Topooco, N., Falkenström, F., Lilliengren, P., Andersson, G., Johansson, R., Midgley, N., Edbrooke-Childs, J., Dahl, H.-S. J., Sandell, R., Thorén, A., Ulberg, R., Bergsten, K. L., & Philips, B. (2022). Therapist-guided internet-based psychodynamic therapy versus cognitive behavioural therapy for adolescent depression in Sweden: A randomised, clinical, non-inferiority trial. *The Lancet Digital Health, 4*(8), e594–e603. https://doi.org/10.1016/S2589-7500(22)00095-4

Michailova, S. (2022). An attempt to understand the war in Ukraine – An escalation of commitment perspective. *British Journal of Management, 33*(4), 1673–1677. https://doi.org/10.1111/1467-8551.12633

Osokina, O., Silwal, S., Bohdanova, T., Hodes, M., Sourander, A., & Skokauskas, N. (2023). Impact of the Russian invasion on mental health of adolescents in Ukraine. *Journal of the American*

*Academy of Child and Adolescent Psychiatry, 62*(3), 335–343. https://doi.org/10.1016/j.jaac.2022.07.845

R Core Team. (2021). *R: A language and environment for statistical computing* [Manual]. https://www.R-project.org/

Riad, A., Drobov, A., Krobot, M., Antalová, N., Alkasaby, M. A., Peřina, A., & Koščík, M. (2022). Mental health burden of the Russian–Ukrainian War 2022 (RUW-22): Anxiety and depression levels among young adults in Central Europe. *International Journal of Environmental Research and Public Health, 19*(14), Article 8418. https://doi.org/10.3390/ijerph19148418

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science, 1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Rosenthal, R. (1991). The evaluation of meta-analytic procedures and meta-analytic results. In *Meta-analytic procedures for social research* (pp. 127–136). SAGE. https://doi.org/10.4135/9781412984997

Ruscio, J. (2008). A probability-based measure of effect size: Robustness to base rates and other factors. *Psychological Methods, 13*(1), 19–30. https://doi.org/10.1037/1082-989X.13.1.19

Sauer-Zavala, S., & Barlow, D. H. (2021). *Neuroticism: A new framework for emotional disorders and their treatment*. The Guilford Press.

Skwirczyńska, E., Kozłowski, M., Nowak, K., Wróblewski, O., Sompolska-Rzechuła, A., Kwiatkowski, S., & Cymbaluk-Płoska, A. (2022). Anxiety assessment in Polish students during the Russian–Ukrainian war. *International Journal of Environmental Research and Public Health, 19*(20), Article 20. https://doi.org/10.3390/ijerph192013284

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Textor, J., van der Zander, B., Gilthorpe, M. S., Liśkiewicz, M., & Ellison, G. T. H. (2016). Robust causal inference using directed acyclic graphs: The R package 'dagitty'. *International Journal of Epidemiology, 45*(6), 1887–1894. https://doi.org/10.1093/ije/dyw341

Verhagen, S., van Os, J., & Delespaul, P. (2022). 5—Ecological momentary assessment and other digital technologies for capturing daily life in mental health. In D. J. Stein, N. A. Fineberg, & S. R. Chamberlain (Eds.), *Mental health in a digital world* (pp. 81–108). Academic Press. https://doi.org/10.1016/B978-0-12-822201-0.00017-4

Vlaescu, G., Alasjö, A., Miloff, A., Carlbring, P., & Andersson, G. (2016). Features and functionality of the Iterapi platform for internet-based psychological treatment. *Internet Interventions, 6*, 107–114. https://doi.org/10.1016/j.invent.2016.09.006

World Health Organization. (2022). *Mental health and COVID-19: Early evidence of the pandemic's impact: Scientific brief, 2 March 2022*. https://www.who.int/publications/i/item/WHO-2019-nCoV-Sci_Brief-Mental_health-2022.1

Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M. W., Gill, H., Phan, L., Chen-Li, D., Iacobucci, M., Ho, R., Majeed, A., & McIntyre, R. S. (2020). Impact of COVID-19 pandemic on mental health in the

general population: A systematic review. *Journal of Affective Disorders, 277*, 55–64. https://doi.org/10.1016/j.jad.2020.08.001

Xu, W., Pavlova, I., Chen, X., Petrytsa, P., Graf-Vlachy, L., & Zhang, S. X. (2023). Mental health symptoms and coping strategies among Ukrainians during the Russia-Ukraine war in March 2022. *The International Journal of Social Psychiatry, 69*(4), 957–966. https://doi.org/10.1177/00207640221143919

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

# Developing a Brief Cognitive Task Intervention to Reduce Long-Standing Intrusive Memories of Trauma: A Feasibility Study With Remote Delivery for Women in Iceland

Johann Palmar Hardarson[1], Beau Gamble[2], Kristjana Thorarinsdottir[1], Elín Sjöfn Stephensen[1], Marie Kanstrup[2], Thorsteinn Gudmundsson[1], Unnur Valdimarsdóttir[3,4,5], Arna Hauksdottir[3], Andri S. Bjornsson[1], Michelle L. Moulds[6§], Emily A. Holmes[7§]

**[1]** *Department of Psychology, University of Iceland, Reykjavik, Iceland.* **[2]** *Department of Psychology, Uppsala University, Uppsala, Sweden.* **[3]** *The Center of Public Health Sciences, University of Iceland, Reykjavik, Iceland.* **[4]** *Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA, USA.* **[5]** *The National University Hospital of Iceland, University of Iceland, Reykjavik, Iceland.* **[6]** *School of Psychology, UNSW Sydney, Sydney, Australia.* **[7]** *Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden.*

[§]*These authors contributed equally to this work.*

Clinical Psychology in Europe, 2024, Vol. 6(1), Article e11237, https://doi.org/10.32872/cpe.11237

**Received:** 2023-01-27 • **Accepted:** 2023-11-17 • **Published (VoR):** 2024-03-28

**Handling Editor:** Cornelia Weise, Philipps-University of Marburg, Marburg, Germany

**Corresponding Author:** Johann Palmar Hardarson, Department of Psychology, University of Iceland, Sæmundargata 12, 102 Reykjavik, Iceland. Phone: +354 525 4240. E-mail: jph6@hi.is

**Supplementary Materials:** Code, Data, Materials, Preregistration [see Index of Supplementary Materials]

## Abstract

**Background:** There is emerging evidence that a brief cognitive task intervention may reduce the frequency of intrusive memories, even long-standing memories of older trauma. However, evaluations to date have involved in-person researcher contact. We investigated the feasibility and acceptability of remote delivery to women (n = 12) in Iceland who had experienced trauma on average two decades earlier.

**Method:** Participants monitored intrusive memories in a daily diary for one week (i.e., baseline phase), completed (at least) two guided, remote intervention sessions (e.g., via secure video platform), and were encouraged to continue to use the intervention self-guided.
**Results:** Eight participants completed the primary outcome and reported fewer intrusive memories in Week 5 (M = 6.98, SD = 5.73) compared to baseline (M = 25.98, SD = 29.39) – a 68% reduction. Intrusions decreased at each subsequent time point; at 3-months (n = 7) there was a 91% reduction compared to baseline. Other psychological symptoms reduced and functioning improved. Importantly, participant ratings and qualitative feedback support feasibility and acceptability.
**Conclusion:** Findings suggest the feasibility of remote delivery of the brief imagery-competing task intervention by non-specialists (who were not mental health professionals) and hold promise for developing psychotherapeutic innovations supporting women with intrusive memories even decades after trauma.

## Keywords

trauma, intrusive memories, intervention, feasibility study, mental imagery

### Highlights

- There is a high global prevalence of trauma exposure and mental health resources are limited.
- The intervention delivered remotely by non-specialists to women in Iceland was feasible and acceptable.
- Participants reported fewer intrusive memories at 5 weeks post-intervention relative to the baseline phase.
- This method could complement existing therapies, in cases of long waitlists or lack of access.

Effective brief, low intensity interventions are needed to address mental health problems on a global scale. Such an intervention has been developed to target intrusive trauma memories (Holmes et al., 2009; Iyadurai et al., 2018; Kanstrup, Singh, et al., 2021). The intervention draws on cognitive neuroscience (Monfils & Holmes, 2018), specifically targeting the potential effect of taxing working memory on altering re-consolidation of trauma memories (Visser et al., 2018). It comprises three components: (1) briefly bringing a trauma memory to mind, (2) engaging in a visuospatial task such as the computer game 'Tetris' for approximately 20 minutes, whilst (3) employing mental rotation during gameplay. Studies in the laboratory (using trauma analogues; e.g., James et al., 2015) and with trauma exposed samples (e.g., women who experienced traumatic childbirth, Horsch et al., 2017; emergency department patients, Iyadurai et al., 2018; Kanstrup, Singh, et al., 2021) demonstrate that receiving the intervention in the initial hours and days posttrauma results in fewer intrusive memories relative to receiving a placebo control.

There is also emerging evidence that this intervention reduces long-standing intrusive memories up to decades old; e.g., in people with chronic PTSD (Kanstrup, Kontio, et al., 2021; Kessler et al., 2018). Further, a pilot case study (Thorarinsdottir et al., 2021) and brief case series ($N = 3$; Thorarinsdottir et al., 2022) with Icelandic women with a chronic trauma history provided preliminary evidence of its capacity to reduce intrusive memories in this group. Not only were treatment gains (i.e., reduced intrusions) maintained at 3 month follow-up, other clinical symptoms (e.g., depression, anxiety) reduced and functioning (e.g., concentration, sleep) improved.

Essential to an intervention's scope for scalability is its capacity for effective remote delivery, eliminating the need for in-person contact. Ideally, scalable interventions should be deliverable by non-specialists who have received remotely-delivered training. Whilst the abovementioned case study (Thorarinsdottir et al., 2021) and case series (Thorarinsdottir et al., 2022) provide encouraging preliminary evidence of the cognitive task intervention's effectiveness, both studies included some aspects of in-person recruitment and/or intervention delivery, and the intervention was delivered by a qualified clinical psychologist.

In line with the goal of establishing scalability, the current study (i) investigated the feasibility of a fully remote delivered, researcher-guided form of the intervention, and (ii) explored pre- to post-intervention changes in the number of intrusive memories. In addition, we delivered some aspects of the intervention in digitalized format; i.e., via brief animated film-clips (e.g., to explain the target symptom).

We investigated feasibility in a sample of trauma-exposed women in Iceland who reported intrusive memories of long-standing trauma. We assessed the feasibility of delivering the intervention in a fully remote format based on the number of sessions completed, dropout rates and reasons, and adverse events. We also investigated the feasibility of conducting remote training and supervision of non-specialists (psychology students) to train them to deliver the intervention. Finally, we assessed intervention acceptability via participants' ratings and qualitative feedback.

Consistent with previous studies (e.g., Kanstrup, Singh, et al., 2021), we predicted that, compared to the baseline phase (Week -1), participants would report fewer intrusive memories in the fifth week after the second intervention session, as assessed via a daily diary (primary outcome). Second, we predicted that the intervention would lead to reductions in related psychological symptoms (e.g., anxiety, depression), and improved functioning (e.g., in concentration, sleep, social relationships) (secondary outcomes). We also aimed to explore whether the frequency of targeted intrusive memories decreased relative to the frequency of non-targeted intrusive memories.

# Method

## Participants

Women in a sub-study of the Stress and Gene Analysis (SAGA) cohort (a population-based longitudinal study of Icelandic women investigating trauma history, www.áfallasaga.is) were screened for eligibility. The sub-study (the Social Trauma Project) involves a comparative analysis of two sub-samples extracted from the SAGA cohort (i.e., women with likely PTSD or no PTSD). Participants were assessed (in person) with the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5; Weathers et al., 2018) and the Mini International Neuropsychiatric Interview for DSM-4 (MINI; Sheehan et al., 1998). These diagnostic interviews were adminstered by fully qualified clinical psychologists and students who were completing their Masters in Clinical Psychology.

Inclusion criteria were: (a) having experienced at least one Criterion A trauma according to the Diagnostic and Statistical Manual of Mental Disorders (5th Ed.; DSM–5; American Psychiatric Association, 2013); (b) reporting at least two intrusive memories in the previous week (consistent with the criterion of a minimum of 1-2 intrusions per week required to endorse this symptom on the CAPS-5); (c) reporting being bothered by intrusive memories over the past month (i.e., scoring at least a moderate score on PCL-5 item 1); (d) able and willing to complete 3-9 sessions with the researcher; (e) willing to monitor intrusive memories; (f) having access to a smartphone; (g) able to speak Icelandic and read study materials in Icelandic. Exclusion criteria (assessed with the MINI) were: (a) current psychotic disorder; (b) current manic episode; (c) being acutely suicidal.

Twelve women were enrolled in the study (mean age = 42.42 years, *SD* = 12.03; mean duration since time of trauma (target memory) = 20.73 years, *SD* = 14.65). Primary traumas were sexual violence ($n$ = 5), witness to death or serious injury ($n$ = 3), physical violence ($n$ = 3), and motor vehicle accident ($n$ = 1). Eight participants completed the intervention and the primary outcome; 7 participants completed the 3-month follow-up.

## Design

Participants monitored intrusive memories of trauma in a daily diary for one week (i.e., baseline phase, Week -1), followed by at least two guided intervention sessions with a researcher remotely (via telephone or secure video platform) over the following week (Week 0). Participants could opt to complete up to four additional guided intervention sessions (i.e., maximum of 6 sessions), until the completion of the primary outcome (i.e., Week 5). After the first guided session, participants were encouraged to use the intervention on their own throughout the study.

Participants continued to monitor their intrusive memories in the daily diary throughout Weeks 0-5. Follow-up questionnaires were completed at Week 1, 1-month, and 3-months after the second intervention session. The primary outcome was the

change in total number of intrusive memories from the baseline week (Week -1) to the fifth week after the second intervention session (Week 5). Participants also monitored (in a daily diary) the number of intrusive memories they experienced for one week, beginning the day of completing the 3-month follow-up questionnaires.

The study had a *repeated AB design*, such that the length of baseline ('A,' preintervention, monitoring only) and intervention ('B') phases differed across each intrusive memory; i.e., depending on when it was targeted. The baseline phase could thus be used as a control period for each individual memory – i.e., to compare the number of intrusive memories before and after the intervention.

## Training and Supervision of Psychology Students to Deliver the Intervention

The intervention was delivered to the first participant by KT, a licensed clinical psychologist who had received training in delivering the intervention via two workshops led by EAH and MK, and had experience in intervention delivery (Thorarinsdottir et al., 2021). The intervention was delivered to the next 11 participants by four MSc students in clinical psychology and two BSc students at the University of Iceland who received remotely-delivered training and ongoing supervision.

To allow remote training during the COVID-19 pandemic, we developed a beta version of an online training course (via the platform www.talentlms.com) in the style of a 'MOOC' (massive open online course), which included material in the form of text, images, animated videos, video roleplay assignments, quizzes, and written reflections (Oakley & Sejnowski, 2019). Alongside the MOOC, training was delivered by KT, JPH, and MK (all with intervention delivery experience [Kanstrup, Kontio, et al., 2021; Kanstrup, Singh, et al., 2021; Thorarinsdottir et al., 2021, 2022]) and supervised by ASB and EAH.

A training group (trainees, trainers, facilitator (BG) and supervisors) met via Zoom for seven one-hour weekly sessions (Sept-Nov 2020), as trainees worked through the online course. Trainees could discuss the MOOC, observe experienced trainers roleplaying and ask questions. Trainees uploaded video roleplays online, assessed by KT and JPH with rating scales ranging from 0 ('absence') to 6 ('excellence') covering nine components (e.g., *'Explanation of the target symptom (intrusive memories)*). Trainees were required to score at least 4 ('competent') on all scales before delivering the intervention.

The group continued to attend weekly Zoom supervision meetings (Jan-July 2021) with the option of individual supervision (from KT).

## Measures

### Primary Outcome Measure

*Intrusive memory diary.* Participants monitored their intrusive memories in a daily paper diary used in previous research (e.g., Iyadurai et al., 2018; Kanstrup, Singh, et al., 2021) and validated (Singh, Ahmed Pihlgren, et al., 2023). Primary outcome was the change in total number of intrusive memories of the traumatic event recorded in the diary (morning, afternoon, evening and night) from baseline week (Week -1) to the fifth week after the second intervention session (Week 5).

### Secondary Outcome Measures

In line with the goal of investigating feasibility and in the interest of brevity, we report data for the first five pre-registered 'Secondary Outcome measures' which examine symptoms of PTSD, depression, and anxiety along with intrusive memories. Findings for the remaining measures including functional measures are presented in Appendix A of the Supplementary Materials.

*Intrusive memory diary.* Change in the total number of intrusive memories recorded in the diary daily during the week of receiving the first two intervention sessions (Week 0), the subsequent four weeks (Weeks 1-4) and at 3-month follow-up, compared to the Baseline week (Week -1).

*Unwanted Memories of Trauma* (*UMT*; Hackmann et al., 2004). Six items measuring the frequency of unwanted memories of the trauma in the previous week, the level of distress, nowness, reliving, disconnectedness associated with intrusions, and the degree to which different triggers are associated with memories of the trauma.

*Posttraumatic Stress Disorder Checklist 5 (PCL-5*; Weathers et al., 2013). A 20-item measure assessing the severity of PTSD symptoms.

*Patient Health Questionnaire-9 (PHQ-9;* Kroenke et al., 2001). A 9-item measure of the severity of depression symptoms.

*Generalized Anxiety Disorder-7* (*GAD-7*; Spitzer et al., 2006). A 7-item screening tool assessing the presence and severity of GAD symptoms.

*Sheehan Disability Scale* (*SDS*; Leon et al., 1997). A measure of functional impairment in work/school, social and family life domains. Items were adapted to assess functional impairment associated with intrusive memories.

*World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0*; World Health Organization, 2010). A 12-item questionnaire measuring difficulties due to health conditions, including mental problems. Lower scores indicate better functioning.

*Impact of intrusive memories on concentration, sleep and stress – Ratings.* Self-rated items assessing the impact of intrusive memories on concentration, sleep and stress in the past week. Two items assess general concentration difficulties and impairments in

concentration, two items assess sleep disturbances, and one item assesses the impact of intrusive memories on stress.

*Rating of how long intrusive memories disrupt concentration.* A single item assessing the estimated average duration of disruption to concentration, rated on a 6-point scale (from 0 = *<1 minute* to 6 = *> 60 minutes*).

*Impact of intrusive memories on functioning.* A 2-item measure assessing the impact of intrusive memories on daily functioning. The first question is: "*Have the intrusive memories affected your ability to function in your daily life in the past week?* (from 0 = *not at all* to 10 = *affected very much*), followed by the open-ended question: *"If yes, how?"*.

*General impact of intrusive memories – Ratings.* Two items assessing the impact of intrusive memories.

### Other Outcome Measures

Only data for the pre-registered 'Other Outcome Measures' that examine feasibility, adherence, and acceptability are reported and described, in the interest of conciseness. For a comprehensive review of the remaining Other Outcome Measures, please refer to the CTR (NCT04709822); the corresponding findings can be found in Appendix B of the Supplementary Materials.

*Self-guided intervention adherence – Usage of the gameplay intervention in daily life.* Two items assessing participants' use of the gameplay intervention in everyday life: *"How many times did you manage to play Tetris after you experienced an intrusive memory?"* If relevant participants were asked a follow-up open-ended question, i.e.: "*Which of your intrusive memories did you target when you played on your own?"*.

*Intrusion diary adherence.* A single item assessing participants' adherence to completing the intrusion diary accurately.

*Acceptability ratings.* Acceptability of the intervention was assessed with two rating items. Acceptability was also assessed with two open-ended questions ("*How did you feel about playing Tetris after you had an intrusive memory?"* and *"Did you find the intervention helpful? If yes, how?").*

*Credibility/expectancy scale.* Prior to completing the intervention for the first time, participants provided ratings of treatment expectancy[1] as well as the degree to which they found the rationale for intervention credible. Wording of the items was adapted for the current study.

1) The CTR states that this scale contained 5 items, but only 4 items were included owing to an administrative error.

## Procedure

Participants were recruited between January and May 2021. Women who participated in the Social Trauma Project sub-study of the SAGA cohort who met the inclusion criteria were contacted (described in Thorarinsdottir et al., 2021).

### Baseline Session

Eligible participants were invited to a remote meeting (i.e., baseline session) with a researcher. Participants were given a brief verbal description of the study, presented with an information sheet containing study details, and provided informed consent by signing an e-consent form in the electronic registration system REDCap. All but one participant indicated that they had a printer and were emailed the diary. A paper diary was delivered to the remaining participant.

Participants then watched a brief video titled *What are intrusive memories?* The researcher asked a series of questions to check their understanding of the content, then sent a link to a second video, *Identifying your intrusive memories.* Participants were then asked to generate a list of intrusive memories they were experiencing. The researcher emphasised that they should not provide a detailed description of each intrusion, but rather summarise each briefly in only a few words (e.g., "*dark room*"). The researcher recorded each intrusive memory in REDCap and shared the screen containing the list of intrusive memories with each participant.

Next, participants watched the third video, *Keeping count of your intrusive memories,* which explained how to monitor intrusive memories. The researcher then explained how to use the intrusive memory diary to monitor their daily intrusions in the week ahead (i.e., baseline, Week -1). Participants also completed baseline questionnaires, and an appointment was scheduled for the first intervention session.

### First Intervention Session

At the start of the session, the researcher explained that the session would involve using the intervention to target one of the participant's intrusive memories, then sent them a link to the fourth video (*What is the intervention?*) which provided a rationale.

Together the participant and researcher then chose an intrusion to target (typically the most frequent or distressing). Next, the researcher asked participants to briefly bring the memory to mind so they could 'see it in their mind's eye', but without discussing its content. The researcher sent a link to a fifth video (*How to play Tetris using mental rotation*), which included instructions about how to play Tetris, and emphasised the importance of mental rotation (i.e., mentally rotating upcoming blocks in the game to visualise how to best place them). After viewing, the researcher directed participants to open www.tetris.com in their web browser and share their screen with the researcher. Participants then had the opportunity to practice playing Tetris (if they wanted to) and were then instructed to engage in gameplay using mental rotation for at least 20

minutes. Next, an appointment was scheduled for the second intervention session, and participants were also given instructions as to how to use the intervention at home. The last three participants also watched a final video, *Tetris and the brain,* which re-iterated the rationale for the intervention and its hypothesised mechanisms.

### Second Intervention Session

Participants received a second intervention session approximately one week later, targeting the same intrusion (i.e., if intrusions persisted) or a different intrusive memory that they wished to reduce. Participants were informed that they had the choice of continuing to use the intervention alone (i.e., self-guided) or scheduling further intervention session/s (up to 6 sessions) with researcher support.

Participants continued to monitor the frequency of both targeted and non-targeted intrusive memories in the daily diary throughout Weeks 0-5.

### Follow-Up Assessments

Participants completed follow-up questionnaires at Week 1, 1-month, and 3-months after the second intervention session. Participants also monitored their intrusive memories for one week at the 3-month follow-up.

At each intervention session and assessment, participants were asked about the occurrence of any adverse events since the previous contact.

This study was registered on ClinicalTrials.gov (NCT04709822) on 14/1/2021. It was approved by the National Bioethics Committee in Iceland (ID: No. VSNb2017110046/03.01, dated 1/10/2019; amendments: (i) 17-238-V23, dated 23/6/2020; (ii) 17-238-V27, dated 24/11/2020; (iii) 17-238-V29-S1, dated 2/3/2021; (iv) 17-238-V30, dated 30/3/2021; (v) 17-238-V31, dated 13/4/2021). Participants provided their informed consent digitally. All sessions followed a written protocol. No serious adverse events or adverse events related to the intervention were reported.

# Results

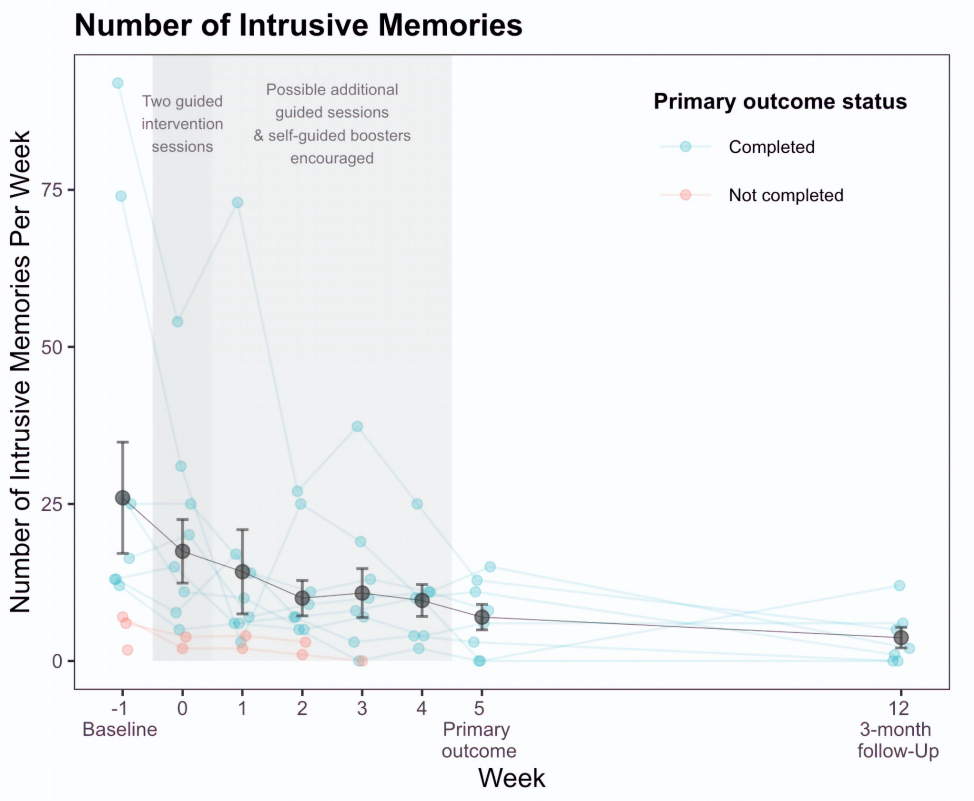
## Analytic Approach

As a feasibility trial, we adopted a descriptive approach to reporting the results. Whilst we collected both qualitative and quantitative data, only quantitative findings are reported here. Analyses were conducted (by BG) using R, Version 4.0.242 ('psych' package, version 2.0.8, for descriptive analyses). De-identified summary data, codebook and R scripts are available on the Open Science Framework (Gamble et al., 2022). Whilst descriptive statistics for all participants are reported below, we also provide data for completers only (i.e., per protocol analyses) on the OSF (Gamble et al., 2022).

Figure 1 presents the number of intrusive memories reported in the daily diary for each participant. Table 1 reports the means, *SD*s, and effect sizes (as Cohen's *d* along with 95% CIs) for (i) number of intrusive memories reported in the daily diary at each assessment point, and (ii) secondary outcome measures at each assessment point. Table 2 reports measures of adherence and credibility.

**Figure 1**

*Number of Intrusive Memories for All Participants (n = 12): Treatment Completers (n = 8) and Non-Completers (n = 4)*




## Primary Outcome

The primary outcome was the change in the total number of intrusive memories recorded in the daily diary from baseline (Week -1) to Week 5. Participants reported fewer intrusive memories of the traumatic event in the fifth week after the second intervention session ($M$ = 6.98, $SD$ = 5.73, range: 0-15) compared to the baseline week (Week -1; $M$ =

25.98, *SD* = 29.39, range: 2-92) – a difference that reflected a 68% reduction in the number of intrusions (i.e., for participants who completed the primary outcome, *n* = 8).

## Secondary Outcomes

We explored whether participants reported fewer intrusive memories in the daily diary at Week 0, Weeks 1-4 and at 3-month follow-up (relative to Week -1, baseline phase), as well as reductions in other psychological symptoms (e.g., anxiety, depression) over the course of the study (see Table 1 for means).

**Table 1**

*Number of Intrusive Memories Reported in the Daily Diary and Self-Report Measures of Posttraumatic Stress Symptoms, Depression and Anxiety for All Participants (n = 12)*

| Outcome | *n* | *M* | *SD* | Cohen's *d* Comparison to baseline | Cohen's *d* 95% CI | |
|---|---|---|---|---|---|---|
| | | | | | *LL* | *UL* |
| **Number of intrusive memories (daily diary)** | | | | | | |
| Baseline (Week -1) | 11 | 25.98 | 29.39 | | | |
| Week 0 | 10 | 17.46 | 16.01 | -0.27 | -0.55 | 0.00 |
| Week 1 | 10 | 14.20 | 21.21 | -0.52 | -1.05 | 0.02 |
| Week 2 | 10 | 10.00 | 8.91 | -0.22 | -0.39 | -0.05 |
| Week 3 | 9 | 10.81 | 11.69 | -0.38 | -0.66 | 0.10 |
| Week 4 | 8 | 9.62 | 7.19 | -0.64 | -1.22 | -0.07 |
| Week 5 | 8 | 6.98 | 5.73 | -1.16 | -2.47 | 0.15 |
| 3-month | 7 | 3.71 | 4.35 | -0.81 | -1.52 | -0.11 |
| **UMT[a] (frequency)** | | | | | | |
| Baseline | 12 | 3.25 | 1.14 | | | |
| Week 1 | 9 | 3.56 | 1.33 | 0.09 | -0.71 | 0.89 |
| 1-month | 8 | 2.38 | 0.74 | -1.52 | -2.98 | -0.06 |
| 3-month | 7 | 1.29 | 0.76 | -2.36 | -3.59 | -1.14 |
| **UMT[a] (distress)** | | | | | | |
| Baseline | 12 | 45.42 | 15.08 | | | |
| Week 1 | 9 | 41.33 | 19.40 | -0.28 | -0.79 | 0.24 |
| 1-month | 8 | 28.50 | 22.58 | -0.94 | -1.69 | -0.18 |
| 3-month | 7 | 24.00 | 29.45 | -0.76 | -1.52 | 0.01 |
| **UMT[a] (nowness)** | | | | | | |
| Baseline | 12 | 38.83 | 25.46 | | | |
| Week 1 | 9 | 25.56 | 22.56 | -0.63 | -1.50 | 0.25 |
| 1-month | 8 | 23.00 | 23.86 | -0.77 | -1.75 | 0.22 |
| 3-month | 7 | 10.43 | 26.72 | -1.10 | -2.32 | 0.12 |

| Outcome | *n* | *M* | *SD* | Cohen's *d* Comparison to baseline | Cohen's *d* 95% CI | |
|---|---|---|---|---|---|---|
| | | | | | *LL* | *UL* |
| **UMT[a] (reliving)** | | | | | | |
| Baseline | 12 | 40.58 | 24.83 | | | |
| Week 1 | 9 | 39.22 | 29.53 | -0.20 | -1.12 | 0.73 |
| 1-month | 8 | 30.38 | 29.40 | -0.36 | -1.40 | 0.69 |
| 3-month | 7 | 17.71 | 26.02 | -0.83 | -1.91 | 0.26 |
| **UMT[a] (disconnectedness)** | | | | | | |
| Baseline | 12 | 61.42 | 24.99 | | | |
| Week 1 | 9 | 72.78 | 17.37 | 0.38 | -0.60 | 1.35 |
| 1-month | 8 | 59.88 | 17.11 | -0.03 | -0.97 | 0.92 |
| 3-month | 7 | 32.14 | 37.81 | -0.81 | -2.30 | 0.68 |
| **UMT[a] (triggers)** | | | | | | |
| Baseline | 12 | 55.67 | 24.63 | | | |
| Week 1 | 9 | 52.89 | 28.87 | -0.36 | -1.04 | 0.32 |
| 1-month | 8 | 35.88 | 27.76 | -0.65 | -1.29 | -0.01 |
| 3-month | 7 | 28.71 | 28.62 | -0.79 | -1.77 | 0.18 |
| **PCL-5[b]** | | | | | | |
| Baseline | 12 | 36.42 | 16.81 | | | |
| Week 1 | 9 | 28.89 | 16.83 | -0.42 | -0.77 | -0.07 |
| 1-month | 8 | 28.50 | 21.27 | -0.31 | -0.73 | 0.10 |
| 3-month | 7 | 18.71 | 17.53 | -0.71 | -1.37 | -0.05 |
| **PHQ-9[c]** | | | | | | |
| Baseline | 12 | 11.83 | 5.70 | | | |
| Week 1 | 9 | 9.89 | 4.76 | -0.20 | -0.53 | 0.13 |
| 1-month | 8 | 8.62 | 3.46 | -0.37 | -0.71 | -0.02 |
| 3-month | 7 | 8.29 | 5.71 | -0.46 | -0.98 | 0.07 |
| **GAD-7[d]** | | | | | | |
| Baseline | 12 | 8.75 | 6.17 | | | |
| Week 1 | 9 | 6.00 | 4.39 | -0.16 | -0.34 | 0.02 |
| 1-month | 8 | 5.62 | 4.00 | -0.30 | -0.57 | -0.03 |
| 3-month | 7 | 5.00 | 4.62 | -0.44 | -1.18 | 0.31 |

[a]Unwanted Memories of Trauma. [b]Posttraumatic Stress Disorder Checklist 5. [c]Patient Health Questionnaire-9. [d]Generalized Anxiety Disorder-7.

## Change in Total Number of Intrusive Memories

Participants recorded fewer intrusive memories in the diary during the week of receiving the first two intervention sessions (Week 0), the subsequent four weeks (Weeks 1-4) and at 3-month follow-up, relative to the baseline week (i.e., Week -1). At 3-month follow-up,

there was a 91% reduction in the number of intrusive memories reported relative to baseline (Week -1) (i.e., for participants who completed the 3-month follow-up, $n$ = 7).

### Symptoms of PTSD

*Unwanted Memories of Trauma (UMT).* Overall, participants reported increased frequency of intrusive memories from baseline to Week 1; however, ratings of frequency declined across subsequent time points. Similarly, ratings of disconnectedness increased from baseline to Week 1, but progressively diminished at each subsequent time point. For the remaining items (distress, nowness, reliving, triggers), participants' ratings steadily decreased from baseline to 3-month follow-up.

*Posttraumatic Stress Disorder Checklist 5 (PCL-5).* PCL-5 scores decreased at each time point, from baseline to 3-month follow-up.

### Depression and Anxiety Symptoms

*Patient Health Questionnaire-9 (PHQ-9).* PHQ-9 scores decreased at each successive time point, from baseline to 3-month follow-up.

*Generalized Anxiety Disorder-7 scale (GAD-7).* Anxiety symptoms decreased at each assessment point, from baseline to 3-month follow-up.

### Functioning

*Sheehan Disability Scale (SDS).* Ratings of functional impairment decreased at each assessment point for all domains, indicating improved self-reported functioning from baseline to 3-month follow-up.

*World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0).* Consistent with the SDS, participants reported improvements in functioning at each timepoint, across the course of the study.

*Impact of intrusive memories on concentration, sleep, and stress – Ratings.* Participants reported improved concentration at each timepoint. Ratings of the impact of intrusive memories on sleep decreased at each timepoint from baseline to 1-month follow-up. However, participants rated an increased impact of their intrusions on sleep from the 1-month to 3-month follow-up. Notably, the mean rating at 3-month follow-up was lower than that reported at baseline. Regarding nightmares, ratings indicated an increased impact of intrusions on nightmares from baseline to Week 1, with decreased ratings at each subsequent timepoint (with a mean of 0 at 3-month follow-up). Finally, although participants reported an overall reduction in the impact of intrusive memories on stress across the study, the means fluctuated across assessment points. Specifically, ratings reduced (indicating that intrusions had less impact on stress) from baseline to Week 1, then increased at 1-month follow-up, and subsequently decreased at 3-month follow-up.

*Rating of how long intrusive memories disrupted concentration on average.* Ratings indicated that the duration of time that intrusive memories disrupted concentration

decreased from baseline to 3-month follow-up. Whilst duration of disruption increased from baseline to 1 week, it decreased at each subsequent timepoint.

*Impact of intrusive memories on functioning.* Ratings of the impact of intrusive memories decreased at each assessment point, from baseline to 3-month follow-up.

*General impact of intrusive memories – Ratings.* Ratings of the vividness of intrusive memories and intrusion-related distress reduced from baseline to 3-month follow-up. Despite these overall reductions there was some fluctuation across assessment points.

We also planned to explore the relative differences in the number of intrusive memories (reported during the baseline phase and Week -1) targeted by the intervention and non-targeted intrusive memories. However, we were unable to conduct these planned exploratory analyses because participants' untargeted intrusive memories were not sufficiently frequent to conduct the comparisons. Whilst such analyses have been carried out in a previous investigation (Kessler et al., 2018), we note that participants in that study were inpatients with complex PTSD who reported frequent intrusive memories of multiple traumas. By comparison, in the current feasibility trial participants reported a smaller number of key intrusive memories, which were the focus of the intervention – and non-targeted intrusions were less frequent.

## Feasibility

*Feasibility of delivering the intervention in fully-remote format.* Twelve participants commenced the trial, of whom 8 completed the primary outcome. Seven treatment completers completed the required two intervention sessions, and one completed four sessions. Of the four non-completers, two completed two intervention sessions, one completed one session, and one completed zero sessions. Two completed the Week 1 follow-up but could not be contacted to obtain the primary outcome. The other two dropped out before the Week 1 follow-up due to unrelated stressors. No adverse events were reported.

*Feasibility of remote training and supervision to deliver the intervention.* All four MSc students and the two BSc students attended all remote training sessions and completed the online training course. Online supervision sessions proved feasible, enabling interactions between the students and trainers in real-time, and the opportunity for practical teaching components such as role-plays.

*Feasibility of training non-specialists (i.e., BSc and MSc students) to deliver the intervention to a competent standard.* All trainees were judged to reach competence (defined as scoring a '4' or greater on all competency rating scales), demonstrating the feasibility of training non-specialists to deliver the intervention. The median competency score across trainees for all rating scales was 5.00, based on the final round of video roleplays completed prior to delivering the intervention to participants.

## Adherence

Participants' ratings of self-guided intervention adherence (i.e., usage in everyday life) indicated that (across all time points) the average number of times participants played Tetris after experiencing an intrusive memory was 3.54 (*SD* = 2.95). These ratings were relatively consistent from Week 1 to 3-month follow-up (range = 2.86 – 4.00; see Table 2). There were high levels of adherence to completing the daily diary: of the 8 participants who completed the primary outcome, the mean percentage of missing days (across all weeks) was 2.27% (*SD* = 4.02%). In addition, participants' self-rated accuracy in completing the diary indicated consistently high levels of accuracy (*M* = 7.96, *SD* = 1.19) across all time points.

**Table 2**

*Self-Report Measures of Ratings of Adherence and Credibility/Expectancy for All Participants (n = 12)*

| Outcome | *n* | *M* | *SD* |
|---|---|---|---|
| **Self-Guided Intervention Adherence** | | | |
| Baseline | | | |
| Week 1 | 9 | 3.67 | 3.20 |
| 1-month | 8 | 4.00 | 3.38 |
| 3-month | 7 | 2.86 | 2.34 |
| **Intrusive Memory Diary Adherence** | | | |
| Baseline | 11 | 8.27 | 1.01 |
| Week 0 | 10 | 7.70 | 1.42 |
| Week 1 | 10 | 7.90 | 1.10 |
| Week 2 | 10 | 8.00 | 1.15 |
| Week 3 | 9 | 7.44 | 1.33 |
| Week 4 | 8 | 8.06 | 1.02 |
| Week 5 | 8 | 7.94 | 1.21 |
| Week 12 | 6 | 8.67 | 1.37 |
| **Credibility/Expectancy – How logical** | | | |
| Baseline | 12 | 74.67 | 20.03 |
| **Credibility/Expectancy – How useful** | | | |
| Baseline | 12 | 75.42 | 18.42 |
| **Credibility/Expectancy – How strongly recommend to a friend** | | | |
| Baseline | 12 | 65.50 | 17.96 |
| **Credibility/Expectancy – How much improvement expected** | | | |
| Baseline | 12 | 69.67 | 16.52 |

### Acceptability

Participants' ratings indicated acceptability. Specifically, participants indicated that they would recommend the intervention to a friend ($M$ = 6.50, $SD$ = 3.51), and considered gameplay an acceptable way to reduce intrusive memories ($M$ = 6.25, $SD$ = 3.20).

### Credibility/Expectancy

Overall, participants rated high levels of intervention expectancy and credibility ($M$ = 71.31, $SD$ = 14.43).

## Discussion

We investigated the feasibility and acceptability of a remotely delivered, researcher-guided imagery-competing task intervention targeting intrusive memories of long-standing trauma in a sample of women in Iceland. Twelve participants commenced the trial, of whom 8 completed the primary outcome. No intervention related adverse events were reported. These data confirm the feasibility of remote delivery of the researcher-guided form of the intervention and good client engagement. The trial also confirmed the feasibility of conducting remote (i.e., fully online) training for non-specialists without clinical psychology qualifications. Finally, participants' ratings indicated acceptability of the remote version of this brief guided intervention.

Another goal was to explore pre- to post-intervention changes in the number of intrusive memories reported in a daily diary. As predicted, participants reported fewer intrusive memories in Week 5 relative to baseline; specifically, a 68% reduction. By 3-month follow-up, there was a 91% reduction relative to baseline. This pattern of improvement was also observed across other psychological outcomes: depression and anxiety symptoms reduced, and self-reported functioning improved. These encouraging results extend the findings of our previous case series' using the same imagery competing task intervention, in which participants ($n$ = 1, $n$ = 3) reported a reduced frequency of intrusive memories by 38% to 56% in the intervention phase, with continued reductions observed at 1- and 3-month follow-ups (Thorarinsdottir et al., 2021, 2022). Notably, in the current study a similar pattern of outcomes was achieved despite having fewer sessions and remote delivery of the intervention by non-specialists (i.e., individuals without professional training in mental health). Current data complement existing evidence of the efficacy of this intervention when delivered in acute settings in the initial hours and days following traumatic events (e.g., in women who experienced traumatic childbirth, Horsch et al., 2017; emergency department patients, (Iyadurai et al., 2018; Kanstrup, Singh, et al., 2021), and also in targeting established memories of trauma in frontline healthcare workers (Iyadurai et al., 2023; Ramineni et al., 2023).

We highlight that these findings are preliminary and come with several limitations, including a small sample size and (in keeping with the aims of a feasibility trial) the

absence of a control condition. We note that intrusions increased from baseline to Week 1, and cannot rule out the possibility that the frequency of intrusive memories may have increased during the baseline phase due to participants being asked to monitor their occurrence, particularly in the context of long-standing trauma. This is an aspect that future studies may wish to carefully monitor and explore further – potentially through employing a larger sample size and extending the study period. Incorporating a longer baseline could also be beneficial to ascertain whether any observed increase is maintained or is transient. Further, this feasibility study cannot clarify the underlying mechanisms of the intervention; specifically, whether intrusions reduce owing to memory reconsolidation (Astill Wright et al., 2021), mental imagery interference (Baddeley & Andrade, 2000), a combination of the two, or other factors.

Should these beneficial outcomes be replicated and extended in randomised controlled trials, they will have important applied implications. Specifically, removing the need for in-person contact will increase the capacity to deliver the intervention at scale, and disseminate it to vulnerable traumatised populations (e.g., refugees), potentially overcoming challenges related to geographical location, language, and other barriers to access (Holmes et al., 2017; Kazlauskas, 2017). Similarly, eliminating the need for highly qualified trauma specialists to deliver the intervention competently will further increase scope for scalability.

**Funding:** EAH received funding from The Oak Foundation (OCAY-18-442) in main support of this study; ASB received funding from the Icelandic Research Fund (11709-0270) and AH as well (200095-5601; 2106-0365). EAH also receives funding from the Swedish Research Council (2020-00873), AFA Insurance (200342), and the Wellcome Trust (223016/Z/21/Z). Funders were not involved in the study design, collection, analysis, and interpretation of data; writing of the report; and/or decision to submit the report for publication.

**Acknowledgments:** The authors thank Bergrun Mist Johannesdottir, Freyja Agustsdottir, Marianna Hlif Jonasdottir, and Roshildur Arna Olafsdottir for their involvement in data collection. We also thank Berglind Gudmundsdottir, Edda Bjork Thordardottir, and Thorhildur Olafsdottir for assistance with study design. Finally, we thank Katy Metcalf for producing the animated videos.

**Competing Interests:** EAH has written books on mental imagery with Guilford Press and Oxford University Press and receives occasional honoraria for conference keynotes and clinical workshops. EAH is on the Board of Trustees of the MQ Foundation. EAH has developed the intervention approach to reduce intrusive memories and training in using it (ANEMONE ™). EAH salary is part funded by the Wellcome Trust (223016/Z/21/Z) via consultancy to P1vital Products Ltd (a study not part of the current work).

**Author Contributions:** Conceptualization (EAH, ASB, BG, JPH), Methodology (all), Formal analysis, data curation, visualization (BG, JPH, MM), Writing – Original Draft (BG, MM, TG), Supervision (ASB, EAH, JPH, MK, MM, KT), Funding acquisition (EAH, ASB, AH). All authors critically revised the manuscript approved the final version for publication.

**Ethics Statement:** This study was approved by the National Bioethics Committee in Iceland (ID: No. VSNb2017110046/03.01, dated 1/10/2019; amendments: (i) 17-238-V23, dated 23/6/2020; (ii) 17-238-V27, dated 24/11/2020; (iii) 17-238-V29-S1, dated 2/3/2021; (iv) 17-238-V30, dated 30/3/2021; (v) 17-238-V31, dated 13/4/2021).

**Data Availability:** Participants provided their consent for de-identified summary data to be made openly available for secondary research; this data can be accessed at the OSF (Gamble et al., 2022). We have aimed to follow FAIR data principles; i.e., such that data are findable, accessible, interoperable, and reusable. Study materials may be made available upon reasonable request with an appropriate materials transfer agreement (MTA) with EAH (Uppsala University). We note delivery of this intervention requires training and supervision.

## Supplementary Materials

The Supplementary Materials include the following items:

- the pre-registration protocol for the study, registered on ClinicalTrials.gov (NCT04709822) on 2021-01-14 (see Bjornsson, 2021)
- de-identified summary data, codebook and R scripts (see Gamble et al., 2022)
- online appendices (see Hardarson et al., 2024)

### Index of Supplementary Materials

Bjornsson, A. S. (2021). *Remote delivery of a brief visuospatial interference intervention to reduce intrusive memories of trauma* (ClinicalTrials.gov ID NCT04709822) [Pre-registration protocol]. ClinicalTrials.gov. https://clinicaltrials.gov/study/NCT04709822

Gamble, B., Kanstrup, M., Stephensen, E. S., Bjornsson, A. S., Holmes, E. A., Singh, L., Hardarson, J. P., Þórarinsdóttir, K., & Moulds, M. L. (2022). *Remote delivery of a brief visuospatial interference intervention to reduce intrusive memories among trauma exposed women: A feasibility study / De-identified data and analysis scripts* [De-identified summary data, codebook and R scripts]. OSF. https://osf.io/2b6nh

Hardarson, J. P., Gamble, B., Thorarinsdottir, K., Stephensen, E. S., Kanstrup, M., Gudmundsson, T., Valdimarsdóttir, U., Hauksdottir, A., Bjornsson, A. S., Moulds, M. L., & Holmes, E. A. (2024). *Supplementary materials to "Developing a brief cognitive task intervention to reduce long-standing intrusive memories of trauma: A feasibility study with remote delivery for women in Iceland"* [Online appendices]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.14093

## References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). American Psychiatric Publishing.

Astill Wright, L., Horstmann, L., Holmes, E. A., & Bisson, J. I. (2021). Consolidation/reconsolidation therapies for the prevention and treatment of PTSD and re-experiencing: A systematic review and meta-analysis. *Translational Psychiatry, 11*(1), Article 453. https://doi.org/10.1038/s41398-021-01570-w

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General, 129*(1), 126–145. https://doi.org/10.1037/0096-3445.129.1.126

Hackmann, A., Ehlers, A., Speckens, A., & Clark, D. M. (2004). Characteristics and content of intrusive memories in PTSD and their changes with treatment. *Journal of Traumatic Stress, 17*(3), 231–240. https://doi.org/10.1023/B:JOTS.0000029266.88369.fd

Holmes, E. A., Ghaderi, A., Eriksson, E., Lauri, K. O., Kukacka, O. M., Mamish, M., James, E. L., & Visser, R. M. (2017). 'I can't concentrate': A feasibility study with young refugees in Sweden on developing science-driven interventions for intrusive memories related to trauma. *Behavioural and Cognitive Psychotherapy, 45*(2), 97–109. https://doi.org/10.1017/S135246581600062X

Holmes, E. A., James, E. L., Coode-Bate, T., & Deeprose, C. (2009). Can playing the computer game "Tetris" reduce the build-up of flashbacks for trauma? A proposal from cognitive science. *PLoS One, 4*(1), Article e4153. https://doi.org/10.1371/journal.pone.0004153

Horsch, A., Vial, Y., Favrod, C., Harari, M. M., Blackwell, S. E., Watson, P., Iyadurai, L., Bonsall, M. B., & Holmes, E. A. (2017). Reducing intrusive traumatic memories after emergency caesarean section: A proof-of-principle randomized controlled study. *Behaviour Research and Therapy, 94*, 36–47. https://doi.org/10.1016/j.brat.2017.03.018

Iyadurai, L., Blackwell, S. E., Meiser-Stedman, R., Watson, P. C., Bonsall, M. B., Geddes, J. R., Nobre, A. C., & Holmes, E. A. (2018). Preventing intrusive memories after trauma via a brief intervention involving Tetris computer game play in the emergency department: A proof-of-concept randomized controlled trial. *Molecular Psychiatry, 23*(3), 674–682. https://doi.org/10.1038/mp.2017.23

Iyadurai, L., Highfield, J., Kanstrup, M., Markham, A., Ramineni, V., Guo, B., Jaki, T., Kingslake, J., Goodwin, G. M., Summers, C., Bonsall, M. B., & Holmes, E. A. (2023). Reducing intrusive memories after trauma via an imagery-competing task intervention in COVID-19 intensive care staff: A randomised controlled trial. *Translational Psychiatry, 13*(1), Article 290. https://doi.org/10.1038/s41398-023-02578-0

James, E. L., Bonsall, M. B., Hoppitt, L., Tunbridge, E. M., Geddes, J. R., Milton, A. L., & Holmes, E. A. (2015). Computer game play reduces intrusive memories of experimental trauma via reconsolidation-update mechanisms. *Psychological Science, 26*(8), 1201–1215. https://doi.org/10.1177/0956797615583071

Kanstrup, M., Kontio, E., Geranmayeh, A., Olofsdotter Lauri, K., Moulds, M. L., & Holmes, E. A. (2021). A single case series using visuospatial task interference to reduce the number of visual intrusive memories of trauma with refugees. *Clinical Psychology & Psychotherapy, 28*(1), 109–123. https://doi.org/10.1002/cpp.2489

Kanstrup, M., Singh, L., Göransson, K. E., Widoff, J., Taylor, R. S., Gamble, B., Iyadurai, L., Moulds, M. L., & Holmes, E. A. (2021). Reducing intrusive memories after trauma via a brief cognitive task intervention in the hospital emergency department: An exploratory pilot randomised controlled trial. *Translational Psychiatry, 11*(1), Article 30. https://doi.org/10.1038/s41398-020-01124-6

Kazlauskas, E. (2017). Challenges for providing health care in traumatized populations: Barriers for PTSD treatments and the need for new developments. *Global Health Action, 10*(1), Article 1322399. https://doi.org/10.1080/16549716.2017.1322399

Kessler, H., Holmes, E. A., Blackwell, S. E., Schmidt, A.-C., Schweer, J. M., Bücker, A., Herpertz, S., Axmacher, N., & Kehyayan, A. (2018). Reducing intrusive memories of trauma using a visuospatial interference intervention with inpatients with posttraumatic stress disorder (PTSD). *Journal of Consulting and Clinical Psychology, 86*(12), 1076–1090. https://doi.org/10.1037/ccp0000340

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Leon, A. C., Olfson, M., Portera, L., Farber, L., & Sheehan, D. V. (1997). Assessing psychiatric impairment in primary care with the Sheehan Disability Scale. *International Journal of Psychiatry in Medicine, 27*(2), 93–105. https://doi.org/10.2190/T8EM-C8YH-373N-1UWD

Monfils, M. H., & Holmes, E. A. (2018). Memory boundaries: Opening a window inspired by reconsolidation to treat anxiety, trauma-related, and addiction disorders. *The Lancet Psychiatry, 5*(12), 1032–1042. https://doi.org/10.1016/S2215-0366(18)30270-0

Oakley, B. A., & Sejnowski, T. J. (2019). What we learned from creating one of the world's most popular MOOCs. *npj Science of Learning, 4*(1), Article 7. https://doi.org/10.1038/s41539-019-0046-0

Ramineni, V., Millroth, P., Iyadurai, L., Jaki, T., Kingslake, J., Highfield, J., Summers, C., Bonsall, M. B., & Holmes, E. A. (2023). Treating intrusive memories after trauma in healthcare workers: A Bayesian adaptive randomised trial developing an imagery-competing task intervention. *Molecular Psychiatry, 28*, 2985–2994. https://doi.org/10.1038/s41380-023-02062-7

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): The development and validation of a structured diagnostic psychiatric interview. *The Journal of Clinical Psychiatry, 59*(Suppl 20), 22–33.

Singh, L., Ahmed Pihlgren, S., Holmes, E. A., & Moulds, M. L. (2023). Using a daily diary for monitoring intrusive memories of trauma: A translational data synthesis study exploring convergent validity. *International Journal of Methods in Psychiatric Research, 32*(1), Article e1936. https://doi.org/10.1002/mpr.1936

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Thorarinsdottir, K., Holmes, E. A., Hardarson, J., Hedinsdottir, U., Kanstrup, M., Singh, L., Hauksdottir, A., Halldorsdottir, T., Gudmundsdottir, B., Valdimarsdottir, U., Thordardottir, E. B., Gamble, B., & Bjornsson, A. (2021). Reducing intrusive memories of childhood trauma using a visuospatial intervention: Case study in Iceland. *JMIR Formative Research, 5*(11), Article e29873. https://doi.org/10.2196/29873

Thorarinsdottir, K., Holmes, E. A., Hardarson, J., Stephenssen, E. S., Jonasdottir, M. H., Kanstrup, M., Singh, L., Hauksdottir, A., Halldorsdottir, T., Gudmundsdottir, B., Thordardottir, E., Valdimarsdottir, U., & Bjornsson, A. (2022). Using a brief mental imagery competing task to reduce the number of intrusive memories: Exploratory case series with trauma-exposed women. *JMIR Formative Research, 6*(7), Article e37382. https://doi.org/10.2196/37382

Visser, R. M., Lau-Zhu, A., Henson, R. N., & Holmes, E. A. (2018). Multiple memory systems, multiple time points: How science can inform treatment to control the expression of unwanted emotional memories. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373*(1742), Article 20170209. https://doi.org/10.1098/rstb.2017.0209

Weathers, F. W., Bovin, M. J., Lee, D. J., Sloan, D. M., Schnurr, P. P., Kaloupek, D. G., Keane, T. M., & Marx, B. P. (2018). The Clinician-Administered PTSD Scale for DSM–5 (CAPS-5): Development and initial psychometric evaluation in military veterans. *Psychological Assessment, 30*(3), 383–395. https://doi.org/10.1037/pas0000486

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). *The PTSD Checklist for DSM-5 (PCL-5).* Scale available from the National Center for PTSD. https://www.ptsd.va.gov

World Health Organization. (2010). *Measuring health and disability: Manual for WHO Disability Assessment Schedule WHODAS 2.0*. World Health Organization.

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

# Feasibility and Acceptability of a Mobile App for Prolonged Grief Disorder Symptoms

Anaïs Aeschlimann[1], Nicolas Gordillo[2], Taro Ueno[3], Andreas Maercker[1], Clare Killikelly[1,4]

**[1]** *Department of Psychology, University of Zurich, Zurich, Switzerland.* **[2]** *Department of Informatics, University of Zurich, Zurich, Switzerland.* **[3]** *SUSMED Inc., Tokyo, Japan.* **[4]** *Department of Psychiatry, University of British Columbia, Vancouver, Canada.*

Clinical Psychology in Europe, 2024, Vol. 6(1), Article e10881, https://doi.org/10.32872/cpe.10881

**Received:** 2022-12-19 • **Accepted:** 2024-01-15 • **Published (VoR):** 2024-03-28

**Handling Editor:** Cornelia Weise, Philipps-University of Marburg, Marburg, Germany

**Corresponding Author:** Clare Killikelly, Department of Psychology, University of Zurich, Binzmuehlestrasse 14/17, CH-8050 Zurich, Switzerland. E-mail: c.killikelly@psychologie.uzh.ch

**Supplementary Materials:** Materials [see Index of Supplementary Materials]

## Abstract

**Background:** Mobile apps provide a unique platform for mental health assessment and monitoring. They can provide real time, accessible data on symptoms of mental disorders that may yield rich data for detailed clinical assessment and help individuals gain insight into their current mental state. We developed one of the first apps for tracking symptoms of prolonged grief disorder.
**Method:** In this pilot feasibility study, we assess the feasibility and acceptability of a new mobile app mGAGE for use once a day for 3 weeks. 27 participants completed mental health assessments at t1 and t2.
**Results:** Adherence to the app protocol was very high with 100% for the first two weeks of use. A surprising finding was the improvement of grief symptoms at t2. Debriefing interviews revealed general qualitative categories including positive feedback, negative feedback and specific recommendations. Overall, the app was found to be feasible for use for the first two weeks and acceptable for bereaved individuals.
**Conclusions:** This app could provide valuable data for in depth clinical assessment, may support individuals to gain greater insight into their symptoms and may have a therapeutic effect in terms of improved grief symptoms. Implications for future studies including use in larger intervention studies are discussed.

## Keywords

prolonged grief disorder, e-mental health, mobile app, self-monitoring intervention

### Highlights

- Prolonged grief disorder is a new mental health disorder in the ICD-11: This is one of the first mobile apps developed to assess the new ICD-11 PGD.
- The app is found to be easy to use and acceptable by bereaved individuals. This may improve clinical assessment as it provides real time data on the intensity and stability of grief over a 2-week period.
- The act of self-monitoring may have an intervention effect; after two weeks of reporting grief symptoms there was a significant improvement in grief symptoms.

The death of a loved one represents one of the most severe life stressors (Breslau et al., 1998). In a majority of those affected (80-90%), acute grief symptoms dissipate after a certain time, however, approximately 10% experience a prolonged and severe grief reaction (Lundorff et al., 2017). With the introduction of the diagnostic category prolonged grief disorder (PGD) in 2018, the World Health Organization's International Classification of diseases (ICD-11) allows the diagnosis of disordered grief reactions for the first time. PGD is characterized by two core symptoms relating to longing or yearning for the deceased and persistent preoccupation with the deceased, accessory symptoms of emotional distress and a time and functional impairment criterion. The criteria also take into account varying cultural norms and practices, stating that symptoms must exceed the typical duration and intensity in an individual's culture and context (Maercker et al., 2013). However, the applicability of the diagnostic criteria, including the time criterion, have not yet been established beyond the Western-European contexts (Stelzer et al., 2020).

If left untreated, individuals affected by PGD are at risk of experiencing a range of further serious health and psychosocial problems including increased rates of cardiovascular problems, high blood pressure, harmful health behaviors, substance abuse, or suicidality (Fujisawa et al., 2010; Kersting et al., 2011; Maercker et al., 2008; Prigerson et al., 2009). Being a relatively recent diagnostic category however, PGD may be difficult for clinicians to differentiate from normal grief (Keeley et al., 2016). Furthermore, PGD does not respond well to interventions that are intended and effective for other bereavement-related mental health problems, e.g., depression, but rather calls for interventions specifically tailored to the precursor to PGD, complicated grief (M. K. Shear et al., 2016). Hence, a correct diagnosis is highly relevant for identifying individuals in need of treatment and providing suitable interventions.

Although recent evidence has shown the effectiveness of interventions specifically tailored to PGD, research still remains limited and further investigation is necessary (Boelen et al., 2006; Bryant et al., 2014; Killikelly & Maercker, 2017; Shear et al., 2005).

Another domain of research, which is still lacking, concerns providing the individual with the right support at the right time (Wagner, 2013; Wakefield, 2012). Currently, the long-term trajectories of PGD are poorly understood and further research on the heterogeneity in the fluctuations of symptoms over time is needed (Bonanno & Malgaroli, 2020; Sveen et al., 2018). Some individuals may need immediate psychotherapy support while others may benefit from self-help and monitoring (Johannsen et al., 2019). In addition, recall bias can be problematic for individuals who provide a one-off self-report questionnaire on symptoms. From depression to psychosis there can be variability in the presentation and reporting of symptoms day by day. Digital technology provides the opportunity to more reliably monitor symptoms daily over a longer period of time to ensure a robust assessment and valid measurement (Lenferink et al., 2022). Alongside more accurate diagnosis, symptom tracking and self-monitoring has two uses for interventions. Firstly, daily monitoring alongside psychological and behavioral interventions would allow participants to see the effect of interventions as they unfold day to day. Secondly, the mere act of tracking and monitoring symptoms could provide more insight into symptoms and their severity.

In recent years, digitalization has gained importance in the domain of mental health, bearing the potential to overcome access barriers, as well as expanding the availability and quality of mental health treatment (Chandrashekar, 2018; Torous et al., 2019). Innovative solutions to self-management of mental health problems are especially relevant, given the large treatment gap, meaning that only a small proportion of individuals in need of treatment receive professional help (Kohn et al., 2004). E-mental health interventions delivered via smartphones bring many advantages including immediate support, constant availability, anonymity, low cost, greater access and equity of mental health resources (Olff, 2015).

There has been an increase in the number of available smartphone apps for monitoring and management of mental health symptoms (Wang et al., 2018). This method of ambulant monitoring possesses a number of advantages, which have been demonstrated for various mental health problems (Myin-Germeys et al., 2011; Wichers et al., 2011). This type of data may deliver insights on symptom triggers, relapse signatures, real-time effects of treatment and can provide early detection of change in symptoms (Birchwood et al., 2000; Gleeson et al., 2005; Palmier-Claus, 2011). This has the potential of facilitating earlier interventions, preventing relapse and avoiding hospital admissions, thus also reducing costs (Whelan et al., 2015). Recently 'My Grief App' was developed by a Swedish collaborator for bereaved parents after the death of a child (Eklund et al., 2021, 2022). It includes modules focused on psychoeducation and intervention with some brief symptom tracking (one item, grief severity). In a pilot study 13 parents used the app for 4 weeks and it was found to be an acceptable and useful intervention. A follow-up randomized control trial is currently underway to assess the effectiveness in reducing PGD symptoms. In terms of grief monitoring, a recent study confirmed the feasibility

and acceptability of grief symptom monitoring using a mobile app and experience sampling methodology (Lenferink et al., 2022). Bereaved individuals responded to PGD symptom questions five times a day for two weeks. Adherence was variable with a high drop out rate of 35-40%. Our current study adds to this new wave of symptom tracking research by developing a mobile app to directly assess ICD-11 PGD symptoms using a validated questionnaire, the International Prolonged Grief Disorder Scale (IPGDS) and using a less intensive monitoring frequency.

## A Mobile Self-Report Tool to Assess PGD: mGAGE

Taking into consideration the current knowledge on the benefits of symptom monitoring and harnessing the potential of digitalization, we designed and developed the *mobile Grief Assessment Guide and E-resource* (mGAGE). mGAGE was designed employing the user-centered design process and two focus groups of bereaved individuals.

The first step in our development process was to design an app for use by bereaved individuals in general, not only those with clinically severe PGD. Focus group participants were recruited from a convenience sample and using the snowball technique and the inclusion criteria included adults who had experienced the death of a loved one at least 6 months ago. The first focus group ($n$ = 4) openly explored the need for and qualities of an app for grief, including advantages and disadvantages of such an app. The second focus group ($n$ = 4) elaborated on a preliminary design for the app and gathered feedback on the design and specific planned features of the app. Examples of the focus group questions include: what type of information about grief would you like to know? How many questions would you answer? What kind of feedback would be helpful for you?

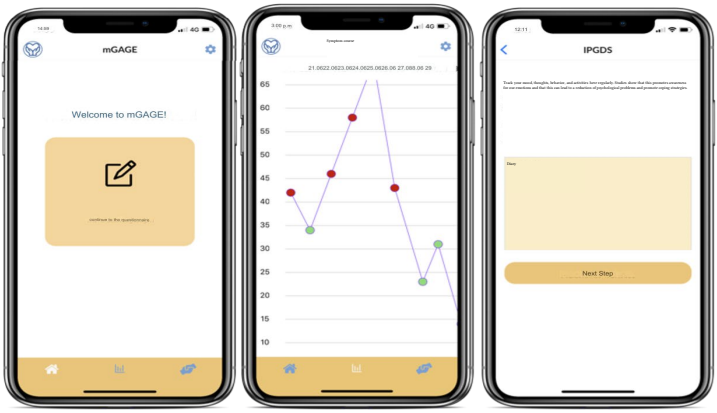
One important consideration in the further development of this app is the purpose of the app. We have sought to ensure that the user centered design (Schnall et al., 2016) is applied to all stages of the app development. User centered design is defined as an iterative developmental process that incorporates user feedback across different stages of app development. This can be achieved through interviews and focus groups. Related to this, the findings from the current study confirm that it is important to provide participants with the choice of how and when to use the app. Several participants identified the need for personalization of the app. In the next iterative round of development, we will add options to ensure that participants can choose when they complete the app, for how long they would like to use it and how frequently.

Moreover, self-monitoring tools, such as mGAGE, have the potential to promote a more empowered and active role of patients in treatment (Huber et al., 2011). Future research should explore these possibilities in the form of a randomized controlled trial with an active control group and importantly, with a clinical sample of those diagnosed with PGD.

In the current version of the mGAGE app it can be delivered via an iOS and Android app, which can be used online on mobile devices. Users create a personalized login ID and are then led to two introduction pages containing information on PGD and the mGAGE app respectively. The app sends users a daily reminder to fill out the integrated questionnaire on PGD symptoms. The completion of the questionnaire takes around 5 minutes. At the end of the questionnaire, users have the option to utilize the diary function, which additionally allows them to record their mood, thoughts, behaviors or actions (see Figure 1). After completion, users receive a feedback concerning the severity of their symptoms, as well as help-seeking recommendations. If the user scores above the cut-off score, indicating that they may suffer from PGD, a direct link to different local services and resources including psychological support appears at the bottom of the page. Furthermore, mGAGE includes a graph function (see Figure 1), which allows users to visualize their assessment history, helping them track the course of their symptoms. Additionally, mGAGE provides users with a support resource page, which lists different professional support resources for bereaved individuals.

**Figure 1**

*Screenshot of the mGAGE App Welcome Page and Symptom Tracking Screens*




The aim of the present study is to (1) evaluate the feasibility of the app in terms of adherence and use of the app as well as (2) the acceptability of the mGAGE app in terms of feedback and evaluation of the app by bereaved individuals. We also explore the variability of IPGDS scores across the 2 weeks and provide insights that may guide the app use to evaluate the effectiveness of PGD interventions.

# Method

## Study Design

We conducted an observational feasibility study with the aim of developing and assessing the feasibility and acceptability of a new mobile grief symptom tracking tool to assist bereaved individuals in self-tracking and monitoring their grief symptoms. The study protocol was approved by the ethics review board of the University of Zurich.

## Participants and Recruitment

Participants were bereaved individuals, who had experienced the loss of a close person. Inclusion criteria specified: 1) fluency in German; 2) age of 18 or older; 3) ability to provide written informed consent; 4) loss of a loved one (family or friend) at least 6 months prior; 5) use of a smartphone. Exclusion criteria were a severe mental health disorder (e.g., Major depression, suicidality, current schizophrenia) or currently being an in-patient.

Participants were remunerated for their time to complete the pre-assessment with 30 Swiss francs (CHF) and the post-assessment with 50 CHF. Furthermore, participants who were Psychology students had the option of being compensated with course credit (4 hours) as an alternative. Additionally, participants had the option of being reimbursed up to 20 CHF for travel expenses related to study participation.

Recruitment took place between August 24$^{th}$ and November 4$^{th}$, 2020. Methods of recruitment included the posting of advertisements and recruitment links on different Facebook groups and mailing lists of the University of Zurich, as well as through WhatsApp groups. Four people who had agreed to participate dropped out before t1. Reasons for drop out included personal reasons and lack of readiness to discuss the death (10 people were interested in the study but did not meet criteria).

## Measures

Participants were assessed twice in-person at the Department of Psychology of the University of Zurich: for the pre-assessment (t1) and the post-assessment (t2). Assessments were conducted by the project manager and a bachelor's student in Psychology under the supervision of a clinical psychologist. The bachelor's student was previously trained by the project manager on the study procedures, including the administration of the questionnaires. Additionally, participants were encouraged to provide daily online data regarding their grief symptoms for 3 weeks between t1 and t2. One participant continued to use the app for longer than the required 21 days (up to 27 days). However, this was only 1 participant and beyond the scope of our protocol. The duration of assessment of up to 3 weeks was determined as a limit based on previous experience sampling methods

with bereaved individuals who found that adherence to self-monitoring significantly declined after 2 weeks (Lenferink et al., 2022; Mintz et al., 2023).

### Socio-Demographic Data

Data on socio-demographic information (age, gender, nationality, marital status, education, psychotherapy experience) was collected at t1.

### Mental Health Outcomes

At t1 and t2, we assessed several mental health outcomes using the measures and descriptions listed here. Grief was assessed using the International ICD-11 Prolonged Grief Disorder Scale, including both the Standard Subscale and the Cultural Supplement, IPGDS Killikelly et al., 2020). The standard subscale consists of 13 items based on the ICD-11 definition of PGD, while the cultural supplement consists of an additional 19 possible items that can be used for treatment planning and cultural acceptability.

Symptoms are rated on a 5-point scale: 1 = almost never (less than once a month), 2 = rarely (monthly), 3 = sometimes (weekly), 4 = often (daily), and 5 = always (several times a day). An impairment item as well as screening items for the length of time since bereavement and the violation of socio-cultural norms was also included. PGD strict criteria requires the fulfillment of the following criteria: one of Items 1 or 2, 1 or more of Items 3-12 and the impairment criteria (Item 13) all rated 4 or above (Lenferink et al., 2021). PGD moderate criteria is the same as the strict criteria except all items are rated 3 or above. Finally, the Maciejewski et al. (2016) criteria includes one of Items 1 or 2, 3-5 of Items 3-12 and no impairment criteria, all rated 4 or higher (Breen & O'Connor, 2007; Maciejewski et al., 2016). Depression: The Patient Health Questionnaire (PHQ) is a self-administered version of the PRIME – MD diagnostic instrument for common mental disorders (Kroenke et al., 2001). The PHQ-9 is the depression module, which scores each of the 9 DSM – IV criteria as "0" (not at all) to "3" (nearly every day). A cutoff score of ≥ 10 has been recommended to indicate moderate to severe depression (Kroenke, Spitzer, & Williams, 2001).

Anxiety: The 7-item Generalized Anxiety Disorder Scale (GAD-7) is a practical self-report anxiety questionnaire that proved valid in primary care according to DSM-IV (Spitzer et al., 2006). Scores for all 7 items range from 0 (Not at all) and 3 (Nearly every day). A cutoff score of 8 or above is recommended to identify possible anxiety disorder.

Post traumatic stress disorder: International Trauma Questionnaire (ITQ) is the ICD-11 based post-traumatic stress disorder (PTSD) measure (Cloitre et al., 2018). The first 9 items of the scale relate to core symptoms of PTSD including re-experiencing, avoidance and sense of current threat and functional impairment. For the purposes of this current study we used only the first 9 items to assess PTSD. For the diagnostic algorithm see Cloitre et al. (2018) Furthermore, we assessed subjective wellbeing (WHO-five wellbeing index, WHO-5, Bech et al., 2003) The total raw score, ranging from 0 to 25, is

multiplied by 4: 0 indicates worst well-being and 100 indicates the best well-being. Daily mobile app data regarding grief symptoms was collected using the Standard Subscale from the IPGDS (13 items) which was recently psychometrically validated in terms of reliability and validity (Killikelly et al., 2020). Previous research using similar daily sampling or experience sampling methodology has purported the importance of allowing participants the option to personalize the assessment method and tailor questions based on their current needs or experiences. Here we pilot the use of personalized assessment by including three idiosyncratic, personally relevant items chosen from the cultural supplement alongside the standard scale of 13 items (van Os et al., 2017).

### Feasibility, Clinical Utility and Acceptability Outcomes

To examine the feasibility of the intervention, we analyzed usage data, specifically (i) percent of participant completing all entries (ii) average entries completed per week. *Acceptability* was assessed during an unstructured exit interview, where participants were asked about their experience using mGAGE and their suggestions for improvement in a potential updated version. Additionally, we employed a questionnaire on acceptability to assess the quality of health-related mobile apps at t2 (mobile app rating scale, MARS, Stoyanov et al., 2015). The MARS evaluates the quality of mobile apps (engagement, functionality, visual aesthetics, information quality and subjective quality subscales) on a scale from 1 (Inadequate) to 5 (Excellent).

## Procedure

Preceding t1, participants were sent an e-mail confirming their appointment for t1, including the study information sheet, the informed consent form, directions to the Department of Psychology and an information sheet about the mGAGE app. Additionally, participants received a reminder e-mail one day before t1.

Assessments were conducted in our offices by the project manager or a Psychology student, who received training on administering the assessment. Participants provided written informed consent. Participants then completed a questionnaire battery about their current mental health status and grief symptoms. Subsequently the participants had the option to select up to three items from the cultural supplement of the IPGDS to be included in their online questionnaire on the mGAGE app. A brief introduction to the app was given where a team member helped participants download the app and create a personal user ID, as well as elaborating all app functionalities and answering any questions participants had (see Supplementary Materials on mobile app information). Participants were instructed to complete the IPGDS within the mGAGE app once a day for three weeks. The IPGDS in this case meant the Standard Subscale of the IPGDS and the potential maximum of 3 additional individual items selected from the cultural supplement of the IPGDS by the participants. Additionally, an information sheet, summa-

rizing all app functionalities was provided. Furthermore, all participants received a list of resources and links for bereaved individuals.

Three weeks after t1, participants were invited to return for the post-assessment t2 and received a reminder e-mail one day before t2. At t2, participants were asked to complete the same questionnaires as for t1 and an additional questionnaire concerning their experience with the mGAGE app. Upon completion, a brief unstructured exit interview was conducted assessing the acceptability of the app. After this, participants received a short debriefing to discuss their experience of participating in the study.

## Analyses

All statistical analyses were conducted in SPSS version 24. Inspection of histograms and the Kolmogorov–Smirnov test statistic (i.e., significance indicates that the distribution of the data significantly differs from a normal distribution) was used to determine whether parametric or non-parametric testing was appropriate. The main outcome measures for grief (IPGDS t1 and t2) were normally distributed while all other measures were normally distributed at one time point. To ensure consistency, the results of the parametric tests (paired samples *t*-tests) were confirmed with non-parametric tests (Mann-Whitney-Wilcoxon test). Feasibility and acceptability variables included percentage of participants completing the daily entry and averaged over the week.

The debriefing interview was analyzed using qualitative thematic analysis. Firstly, the interview data was in vivo transcribed into short relevant sentences and translated into English. Secondly the text was coded and grouped into large categories following iterative categorization (Neale, 2016). Finally, the codes and themes were reviewed by CK and AA and final consensus codes were determined.

# Results

Table 1 presents the demographic characteristics of the participants. Participants were mostly University educated (40.7%) young (average age 25.7) women (88.9%). Time since loss ranged from 6 months to more than 10 years. Almost 30% of participants had previous experience with psychotherapy. Mean scores on the mental health outcome measures are presented in Table 2. Paired sample *t*-tests revealed no significant differences between t1 and t2 on all mental health measures, except for the IPGDS standard scale (13 items) (t1 mean 21.3 vs. t2 mean 18.5, $p = .002$). Table 3 presents the diagnostic algorithm findings. None of the participants met criteria for a strict diagnosis of PGD, while one met criteria for moderate PGD at t1 and two participants at t2. Three participants met criteria for PTSD at t2.

**Table 1**

*Sociodemographic, Loss-Related and Symptom Characteristics*

| | Total sample ($n$ = 27) | |
|---|---|---|
| **Variable** | ***n*** | **%** |
| **Gender** | | |
| Male | 3 | 11.1 |
| Female | 24 | 88.9 |
| **Education** | | |
| Matura[a] | 16 | 59.3 |
| College/university | 11 | 40.7 |
| **Relationship status** | | |
| Single[b] | 26 | 96.3 |
| Married | 1 | 3.7 |
| **Time since loss** | | |
| 6 to 12 months ago | 8 | 29.6 |
| 1 to 5 years ago | 13 | 48.1 |
| 10 to 20 years ago | 2 | 7.4 |
| **Other** | | |
| Previous psychotherapy for grief[c] | 8 | 29.7 |
| Current psychotherapy[c] | 2 | 7.4 |

*Note.* $N$ = 27. Participants were on average 25.70 years old ($SD$ = standard deviation).
[a]Equivalent to high school education or above in Switzerland. [b]Is defined here as non-married individuals. [c]Reflects the number and percentage of participants answering "yes" to this question.

In order to assess the feasibility of mGAGE use, variables related to adherence to the app were examined. Only one participant completed all entries over the total possible 21 days. Up to Day 15 all participants completed the required daily entry. From Day 16 to 21, there was a drop in adherence. As revealed in Figure 3, the average number of entries drops in Week 3 (67.7%) compared to Week 1 and 2 (100%). The average IPGDS score was calculated for Days 1 to 15 (with all data points complete for all participants) to examine the variability in the average score. (See Figures 2 and 3). Paired samples $t$-test compared the highest scored day (Day 2, 20.0) vs. lowest average scored day (Day 11, 18.52) and confirmed a statistically significant difference ($p$ = .049).

**Table 2**

*Questionnaire Data at Pre (T1) Assessment and Post (T2) Assessment*

| | T1 Pre | | T2 Post | | Difference test |
|---|---|---|---|---|---|
| **Variable** | ***M*** | ***SD*** | ***M*** | ***SD*** | **Paired samples *t*-test** |
| PGD sum score (13 standard items) | 21.3 | 6.4 | 18.5 | 6.3 | .002* |
| PGD cultural supplement | 26.5 | 5.9 | 26.2 | 8.7 | .779 |
| PHQ9 | 4.2 | 2.7 | 4.4 | 3.9 | .790 |
| GAD7 | 4.2 | 3.2 | 3.8 | 3.1 | .374 |
| ITQ Re | 1.3 | 1.7 | 1.6 | 1.9 | .387 |
| ITQ Av | 1.6 | 2.2 | 1.3 | 2.0 | .235 |
| ITQ Th | .6 | 1.0 | .7 | 1.3 | .574 |
| WHO wellbeing scale | 15.1 | 5.2 | 15.4 | 4.8 | .704 |
| PG13 | 17.6 | 5.7 | 16.8 | 6.4 | .204 |

*Note.* *M* = mean; *SD* = standard deviation; PGD = prolonged grief disorder; PHQ9 = patient health questionnaire 9; GAD7 = generalized anxiety disorder 7; ITQ Re = International Trauma Questionnaire – Reexperiencing subscale; ITQ Av = International Trauma Questionnaire – Avoidance subscale; ITQ Th = International Trauma Questionnaire – Threat subscale; WHO = world health organization wellbeing scale; PG13 = prolonged grief 13. *$p < .05$. **$p < .01$. ***$p < .001$.

**Table 3**

*Diagnostic Algorithm Comparison*

| | T1 Pre | | T2 Post | |
|---|---|---|---|---|
| **Diagnostic Test** | ***%*** | ***n*** | ***%*** | ***n*** |
| PGD Strict | 0.0 | | 0.0 | |
| PGD Moderate | 3.7 | 1 | 7.4 | 2 |
| Maciejewski 2016 criteria | 3.7 | 1 | 0.0 | |
| PTSD | 0.0 | | 11.1 | 3 |

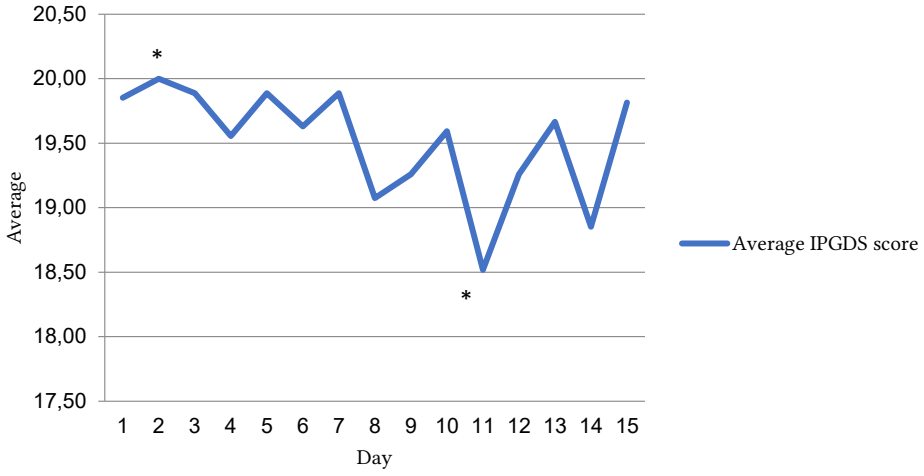
*Note.* PGD = prolonged grief disorder; PTSD = posttraumatic stress disorder.

## Acceptability of mGAGE

Overall mean scores on the MARS questionnaire ranged from 3 to 5 indicating average to very good acceptability for the app (across all questions mean score of 4). The lowest scored question was ‘would you pay for this app’ (average 2). The highest scored questions were ease of use, navigation and gestural designs (all scored average of 5). For the question ‘overall star rating’ (Question 11) participants’ mean score was 4/5.

**Figure 2**

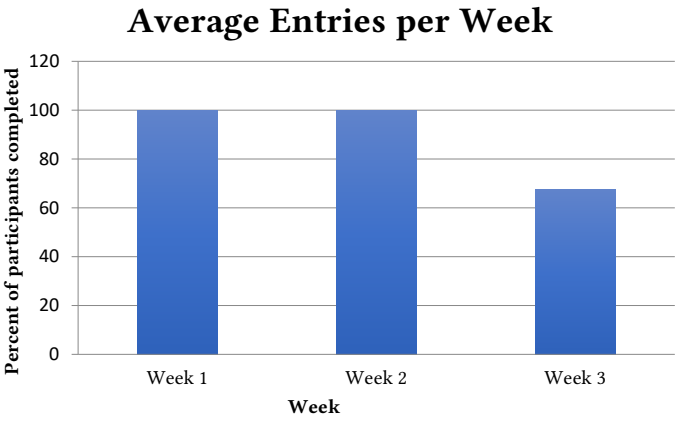
*Daily Average IPGDS Score for All Participants*




*Note.* IPGDS = International Prolonged Grief Disorder Scale.
*Statistically significant difference between average daily IPGDS on Day 2 and Day 11; $p < .05$.

**Figure 3**

*Adherence per Week*




*Note.* Percent of participants completing the daily questionnaire.

Thematic analysis was conducted on the debriefing interviews. Overall, three broad categories of responses were identified, each with related themes (see Table 4 for exemplar quotes).

**Table 4**

*Summary of Qualitative Feedback and Example Quotes*

| Themes | Example quotes |
|---|---|
| **Positive Feedback** | |
| Useful links | especially the links to get help were helpful, |
| Self-monitoring and reflection | But she liked the app, it helped her to reflect and process her grieving thoughts. |
| Ease of Use | easy to use, also good for older people, fulfilled its purpose, |
| Good design | the design was very simple but good; |
| **Negative Feedback** | |
| Effect on symptoms | He said even though he thought, that he had already processed his grief symptoms, the app made him a bit sad in the beginning |
| Notifications and technical issues | the reminders didn't work on a daily basis. Some days the app reminded her, some days not |
| Repetitive | After two weeks he thought, that the questions were getting boring and annoying, cause they were always the same |
| Design issues | but the design wasn't that appealing |
| **Specific recommendations** | |
| Personalization and timing | the participant would have liked the app to react to the questions and ask more specific questions |
| Variation | she would have preferred the questions to be in a different order each time, cause it was a bit repetitive |
| Other functions | tips like breathing exercises for when score is red (instead of just contact numbers) |
| IPGDS changes | the IPGDS should have more gradations or possibly a slider bar for more sensitivity |
| Chart improvement | it would be nice to have a summary of the scores if the app should be used for a longer period of time (e.g. a smaller graph) |
| **Other** | |
| | Even though she liked the app, she wouldn't pay for it. |
| | Her grief symptoms aren't as strong anymore as they used to be, that's why she didn't find the questions disturbing. |

*Note.* IPGDS = International Prolonged Grief Disorder Scale.

The first category was 'positive feedback' with related themes including *useful links, self-monitoring and reflection, ease of use, and good design.* The second category was 'negative feedback' with related themes including *effect on symptoms, notifications and technical issues, repetitive, and design issues.* The third category 'specific recommendations' included related themes such as *personalization and timing, variation, other functions, IPGDS changes, chart improvement.*

# Discussion

## Main Findings

This pilot study provides preliminary support for the feasibility and acceptability of the mGAGE app for use by bereaved individuals. Feasibility, assessed through adherence to the app, was extremely high for the first two weeks of use (100%). Acceptability, assessed by high ratings on the mobile app rating scale (MARS) questionnaire (mean score of 4/5), was also confirmed. Qualitative feedback from debriefing interviews revealed several themes supporting the acceptability of the app (useful links, self-monitoring and reflection, ease of use, and good design). However, participants also identified possible negative effects on mood as well as specific recommendations for improving the app. The final aim of this study was to explore variability in grief symptoms over two weeks. Here we confirmed high variability (a statistically significant difference between highest and lowest mean scores) over the two weeks. This has important implications for better understanding variation and intensity of grief scores in real time as outlined below. Additionally, the process of self-monitoring may have a therapeutic effect. At t2 participants had significantly lower scores on the International Prolonged Grief Disorder Scale (IPGDS). This significant reduction in symptoms may indicate that online self-monitoring is a beneficial intervention and could be used to supplement face to face therapy, which is in accordance with results from previous studies (e.g. Bakker & Rickard, 2018; Kauer et al., 2012). However, this result should be interpreted with caution as no control group was included in this study. The finding that one participant met moderate diagnostic criteria for PGD and two participants met criteria for PTSD at T2 although not at T1 attests to the need for a control group to unpick the effect of PGD symptom monitoring compared to other possible confounding or moderating effects on T2 outcomes.

Previous research has confirmed that adherence to e-mental health apps wavers. Response rates of around 60-80% have been found in previous e-mental health studies of depression (70%; Putnam & McSweeney, 2008), substance misuse (88.8%; Phillips et al., 2014), and trauma (67.5%; Dewey et al., 2015). Evidently the current study found exceptionally high adherence; 100% of responses completed by all participants. However, after two weeks adherence declines substantially (Week 3 average response 67.7%). This is also

found in other areas of e-mental health. After an initial burst of interest participants typically decrease their use of health apps after two weeks (Dorsey et al., 2017). Reasons for decline in use include poor user centered design, lack of incentive and decreased internal motivation (Torous et al., 2019). Importantly, it should be noted that many participants mentioned technical problems as a reason for decline in use. The current study developed the mGAGE app following a user centered iterative design and we plan to build on the current findings to redesign elements of the app to improve acceptability. Mohr et al. (2011) developed a model of 'supportive accountability' which seeks to improve adherence to e-mental health interventions by adding person-to-person contact throughout the intervention. They argue that participants are more likely to continue the intervention if they experience contact with a 'coach' who is supportive, trustworthy and kind. To improve adherence to the current app we would consider adding contact with a weekly coach to check-in and provide support and encouragement. However, based on the current findings we may consider limiting the required use of the app to two weeks. After two weeks participants have provided daily real-time data on their grief symptoms. This may be enough data to capture a comprehensive clinical picture for further assessment and treatment by a clinician. Additionally, this may provide the individual with more insight into their current symptoms.

One interesting finding of the current study was the possible therapeutic effect of self-monitoring. Previous literature has revealed that the repeated act of completing an outcome measure may be an intervention in itself (Amble et al., 2015). This study found that participants had decreased grief scores at t2. Previous studies have also found that daily self-monitoring may improve mental health symptoms. For example, after 30 days of mood monitoring with a new mobile app, MoodPrism, participants experienced a significant decrease in symptoms of depression and anxiety, and this was directly related to app use (Bakker & Rickard, 2018). The act of daily self-reflection and improved insight into symptoms is the goal of several cognitive-behavioral therapy interventions for grief and other disorders (Boelen et al., 2011; Kavanagh, 1990; Spuij et al., 2015). These are found to be predictors of positive affect, cognitive reappraisal as well as emotional self-awareness (Kauer et al., 2012; O'Toole et al., 2014). Emotional self-awareness is thought to be a key factor that may improve self-regulation and wellbeing (Barrett & Gross, 2001).

Overall participants rated the app as highly acceptable. However, they identified some important areas for improvement. Several participants identified that daily self-monitoring may not have a positive effect on mood, but instead remind bereaved individuals of their sadness and grief. One participant identified that she thought she experienced more negative symptoms after using the app daily. Another worried that it could be difficult for people to be reminded of their loss everyday (see Table 4). This is an important consideration and will be an important topic for the next phase of research, particularly when conducting research with a clinical sample (Wykes & Brown, 2016).

As mentioned above, we would consider adding the model of supportive accountability, not only to improve intrinsic motivation to complete the app but to offer participants support, as well as for risk assessment.

## Clinical Implications and Future Research

Overall, the findings of this pilot study confirm that the mGAGE app may have an important use in future randomized controlled trials and grief interventions. Firstly, it can deliver real time daily data on symptom variability and intensity, which provides researchers with a tool to accurately investigate the heterogeneity in the fluctuations of PGD symptoms over time in future research (Bonanno & Malgaroli, 2020; Sveen et al., 2018). This may also be used to inform clinical assessments and treatment options by providing more depth and richer information than a singular self-report assessment. Secondly, daily monitoring can be used to track the effectiveness of grief interventions. As change in symptom severity may vary and effects may waver, mGAGE could be used as an addition to standard, face-to-face therapy and help to track progress and effectiveness, as suggested by Torous and Roux (2017). It could also be used in conjunction with existing online grief interventions (Wagner et al., 2006). Additionally, mGAGE has the potential to aid clinicians in identifying individuals with PGD. By facilitating early identification of cases and accelerating access to appropriate treatment, it could also prevent hospitalization or relapses (Whelan et al., 2015). Thirdly, the act of self-monitoring may be a useful intervention by increasing emotional self-awareness. Developing insight into the severity and variability of symptoms may be an effective therapeutic tool (Bakker & Rickard, 2018).

## Limitations

This study was limited in the following ways. The sample size was small and homogeneous. It was a largely female sample of the same age group and education level. We did not include individuals from a clinical sample. This is the next required step to ensure acceptability for PGD diagnosis. There were also several technical issues with the app (such as server unavailable, only worked with internet connection) that may have prevented optimal data collection. In terms of the qualitative debriefing interviews, there may have been a social desirability effect as the participants were not blinded to the interviewer. The finding that two participants met criteria for PTSD or moderate PGD at T2 needs further investigation. Currently no participants at t1 or t2 met 'strict' criteria for PGD is reassuring as our intention was not to investigate a clinical sample in this pilot study. The diagnostic algorithm for PGD is currently under debate with no clear consensus on whether the 'strict' 'moderate' or another algorithm for ICD-11 PGD may yield the most reliable and valid diagnosis (Boelen & Lenferink, 2020). In the present study if a participant met criteria for moderate PGD, currently this does not necessarily

indicate disorder. However, it may mean elevated symptoms. This should be monitored and follow up with a case control analysis or an RCT including a control group. Another significant limitation in the sample that requires follow up is the heterogeneity in the time criteria. Time since death is a significant predictor of grief symptom severity and although all included participants experienced a death more than 6 months ago, there was still large variability in the duration since loss. The impact of time since loss on daily sampling and grief symptom variability should be assessed in future studies.

**Funding:** CK and TU have established a research partnership to develop a mobile app. This is funded by the Innovation Partnership Grant, IPG 05-122017 from the University of Zurich. This study is part of a larger study in collaboration with SUSMED inc.

**Acknowledgments:** We would like to acknowledge the participation of bereaved participants during the COVID-19 pandemic and the hard work of interns at the Universtiy of Zurich.

**Competing Interests:** AA, CK, NG and AM report no conflict of interest. Taro Ueno is member of Japanese based startup company, SUSMED inc.

**Author Contributions:** CK and AA wrote the manuscript and conducted the study. NG developed the app. AM and TU advised on the study and the project management.

**Ethics Statement:** Ethical approval was obtained from the University of Zurich.

## Supplementary Materials

The Supplementary Materials contain the following item (for access, see Aeschlimann et al., 2024):

- Information pamphlet in German and English describing the functions and features of the mGAGE app for prolonged grief disorder.

### Index of Supplementary Materials

Aeschlimann, A., Gordillo, N., Ueno, T., Maercker, A., & Killikelly, C. (2024). *Supplementary materials to "Feasibility and acceptability of a mobile app for prolonged grief disorder symptoms"* [Information pamphlet]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.14238

## References

Amble, I., Gude, T., Stubdal, S., Andersen, B. J., & Wampold, B. E. (2015). The effect of implementing the Outcome Questionnaire-45.2 feedback system in Norway: A multisite randomized clinical trial in a naturalistic setting. *Psychotherapy Research, 25*(6), 669–677. https://doi.org/10.1080/10503307.2014.928756

Bakker, D., & Rickard, N. (2018). Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *Journal of Affective Disorders, 227*, 432–442. https://doi.org/10.1016/j.jad.2017.11.016

Barrett, L. F., & Gross, J. J. (2001). Emotional intelligence: A process model of emotion representation and regulation. In T. J. Mayne & G. A. Bonanno (Eds.), *Emotions: Current issues and future directions* (pp. 286–310). Guilford Press.

Bech, P., Olsen, L., Kjoller, M., & Rasmussen, N. (2003). Measuring well-being rather than the absence of distress symptoms: A comparison of the SF-36 Mental Health subscale and the WHO-Five well-being scale. *International Journal of Methods in Psychiatric Research, 12*(2), 85–91. https://doi.org/10.1002/mpr.145

Birchwood, M., Spencer, E., & McGovern, D. (2000). Schizophrenia: Early warning signs. *Advances in Psychiatric Treatment, 6*(2), 93–101. https://doi.org/10.1192/apt.6.2.93

Boelen, P. A., van den Hout, M. A., & Bout, J. V. D. (2006). A cognitive-behavioral conceptualization of complicated grief. *Clinical Psychology: Science and Practice, 13*(2), 109–128. https://doi.org/10.1111/j.1468-2850.2006.00013.x

Boelen, P. A., de Keijser, J., van den Hout, M. A., & van den Bout, J. (2011). Factors associated with outcome of cognitive-behavioural therapy for complicated grief: A preliminary study. *Clinical Psychology & Psychotherapy, 18*(4), 284–291. https://doi.org/10.1002/cpp.720

Boelen, P. A., & Lenferink, L. I. M. (2020). Comparison of six proposed diagnostic criteria sets for disturbed grief. *Psychiatry Research, 285*, Article 112786. https://doi.org/10.1016/j.psychres.2020.112786

Bonanno, G. A., & Malgaroli, M. (2020). Trajectories of grief: Comparing symptoms from the DSM-5 and ICD-11 diagnoses. *Depression and Anxiety, 37*(1), 17–25. https://doi.org/10.1002/da.22902

Breen, L. J., & O'Connor, M. (2007). The fundamental paradox in the grief literature: A critical reflection. *Omega, 55*(3), 199–218. https://doi.org/10.2190/OM.55.3.c

Breslau, N., Kessler, R. C., Chilcoat, H. D., Schultz, L. R., Davis, G. C., & Andreski, P. (1998). Trauma and posttraumatic stress disorder in the community: The 1996 Detroit Area Survey of Trauma. *Archives of General Psychiatry, 55*(7), 626–632. https://doi.org/10.1001/archpsyc.55.7.626

Bryant, R. A., Kenny, L., Joscelyne, A., Rawson, N., Maccallum, F., Cahill, C., Hopwood, S., Aderka, I., & Nickerson, A. (2014). Treating prolonged grief disorder: A randomized clinical trial. *JAMA Psychiatry, 71*(12), 1332–1339. https://doi.org/10.1001/jamapsychiatry.2014.1600

Chandrashekar, P. (2018). Do mental health mobile apps work: Evidence and recommendations for designing high-efficacy mental health mobile apps. *mHealth, 4*(3), Article 6. https://doi.org/10.21037/mhealth.2018.03.02

Cloitre, M., Shevlin, M., Brewin, C. R., Bisson, J. I., Roberts, N. P., Maercker, A., Karatzias, T., & Hyland, P. (2018). The International Trauma Questionnaire: Development of a self-report measure of ICD-11 PTSD and complex PTSD. *Acta Psychiatrica Scandinavica, 138*(6), 536–546. https://doi.org/10.1111/acps.12956

Dewey, D., McDonald, M. K., Brown, W. J., Boyd, S. J., Bunnell, B. E., & Schuldberg, D. (2015). The impact of ecological momentary assessment on posttraumatic stress symptom trajectory. *Psychiatry Research, 230*(2), 300–303. https://doi.org/10.1016/j.psychres.2015.09.009

Dorsey, E. R., "Yvonne" Chan, Y.-F., McConnell, M. V., Shaw, S. Y., Trister, A. D., & Friend, S. H. (2017). The use of smartphones for health research. *Academic Medicine, 92*(2), 157–160. https://doi.org/10.1097/ACM.0000000000001205

Eklund, R., Eisma, M. C., Boelen, P. A., Arnberg, F. K., & Sveen, J. (2021). Mobile app for prolonged grief among bereaved parents: Study protocol for a randomised controlled trial. *BMJ Open, 11*(12), Article e052763. https://doi.org/10.1136/bmjopen-2021-052763

Eklund, R., Eisma, M. C., Boelen, P. A., Arnberg, F. K., & Sveen, J. (2022). My Grief app for prolonged grief in bereaved parents: A pilot study. *Frontiers in Psychiatry, 13*, Article 872314. https://doi.org/10.3389/fpsyt.2022.872314

Fujisawa, D., Miyashita, M., Nakajima, S., Ito, M., Kato, M., & Kim, Y. (2010). Prevalence and determinants of complicated grief in general population. *Journal of Affective Disorders, 127*(1-3), 352–358. https://doi.org/10.1016/j.jad.2010.06.008

Gleeson, J. F., Rawlings, D., Jackson, H. J., & McGorry, P. D. (2005). Early warning signs of relapse following a first episode of psychosis. *Schizophrenia Research, 80*(1), 107–111. https://doi.org/10.1016/j.schres.2005.07.019

Huber, M., Knottnerus, J. A., Green, L., van der Horst, H., Jadad, A. R., Kromhout, D., Leonard, B., Lorig, K., Loureiro, M. I., van der Meer, J. W. M., Schnabel, P., Smith, R., van Weel, C., & Smid, H. (2011). How should we define health? *BMJ, 343*, Article d4163. https://doi.org/10.1136/bmj.d4163

Johannsen, M., Damholdt, M. F., Zachariae, R., Lundorff, M., Farver-Vestergaard, I., & O'Connor, M. (2019). Psychological interventions for grief in adults: A systematic review and meta-analysis of randomized controlled trials. *Journal of Affective Disorders, 253*, 69–86. https://doi.org/10.1016/j.jad.2019.04.065

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), Article e67. https://doi.org/10.2196/jmir.1858

Kavanagh, D. J. (1990). Towards a cognitive-behavioural intervention for adult grief reactions. *The British Journal of Psychiatry, 157*(3), 373–383. https://doi.org/10.1192/bjp.157.3.373

Keeley, J. W., Reed, G. M., Roberts, M. C., Evans, S. C., Robles, R., Matsumoto, C., Brewin, C. R., Cloitre, M., Perkonigg, A., Rousseau, C., Gureje, O., Lovell, A. M., Sharan, P., & Maercker, A. (2016). Disorders specifically associated with stress: A case-controlled field study for ICD-11 mental and behavioural disorders. *International Journal of Clinical and Health Psychology, 16*(2), 109–127. https://doi.org/10.1016/j.ijchp.2015.09.002

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

Killikelly, C., & Maercker, A. (2017). Prolonged grief disorder for ICD-11: The primacy of clinical utility and international applicability. *European Journal of Psychotraumatology, 8*(sup6), Article 1476441. https://doi.org/10.1080/20008198.2018.1476441

Killikelly, C., Zhou, N., Merzhvynska, M., Stelzer, E.-M., Dotschung, T., Rohner, S., Sun, L. H., & Maercker, A. (2020). Development of the International Prolonged Grief Disorder Scale for the ICD-11: Measurement of core symptoms and culture items adapted for Chinese and German-speaking samples. *Journal of Affective Disorders, 277*, 568–576. https://doi.org/10.1016/j.jad.2020.08.057

Kohn, R., Saxena, S., Levav, I., & Saraceno, B. (2004). The treatment gap in mental health care. *Bulletin of the World Health Organization, 82*, 858–866.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16*(9), 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x

Lenferink, L. I. M., Boelen, P. A., Smid, G. E., & Paap, M. C. S. (2021). The importance of harmonising diagnostic criteria sets for pathological grief. *The British Journal of Psychiatry, 219*(3), 473–476. https://doi.org/10.1192/bjp.2019.240

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*, Article 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

Maciejewski, P. K., Maercker, A., Boelen, P. A., & Prigerson, H. (2016). "Prolonged grief disorder" and "persistent complex bereavement disorder", but not "complicated grief", are one and the same diagnostic entity: An analysis of data from the Yale Bereavement Study. *World Psychiatry, 15*(3), 266–275. https://doi.org/10.1002/wps.20348

Maercker, A., Brewin, C., & Bryant, R. (2013). Proposals for mental disorders specifically associated with stress in the International Classification of Diseases-11. *Lancet, 381*(9878), 1683–1685. https://doi.org/10.1016/S0140-6736(12)62191-6

Maercker, A., Forstmeier, S., Enzler, A., Krüsi, G., Hörler, E., Maier, C., & Ehlert, U. (2008). Adjustment disorders, posttraumatic stress disorder, and depressive disorders in old age: Findings from a community survey. *Comprehensive Psychiatry, 49*(2), 113–120. https://doi.org/10.1016/j.comppsych.2007.07.002

Mintz, E., Toner, E. R., Skolnik, A., Pan, A. Y., Frumkin, M., Baker, A., Simon, N., & Robinaugh, D. (2023). *Ecological momentary assessment with bereaved adults: Feasibility, acceptability, and measurement reactivity and their relationship to prolonged grief symptom severity.* PsyArXiv. https://doi.org/10.31234/osf.io/hw7jm

Mohr, D. C., Cuijpers, P., & Lehman, K. (2011). Supportive accountability: A model for providing human support to enhance adherence to eHealth interventions. *Journal of Medical Internet Research, 13*(1), Article e30. https://doi.org/10.2196/jmir.1602

Myin-Germeys, I., Birchwood, M., & Kwapil, T. (2011). From environment to therapy in psychosis: A real-world momentary assessment approach. *Schizophrenia Bulletin, 37*(2), 244–247. https://doi.org/10.1093/schbul/sbq164

Neale, J. (2016). Iterative categorization (IC): A systematic technique for analysing qualitative data. *Addiction, 111*(6), 1096–1106. https://doi.org/10.1111/add.13314

Olff, M. (2015). Mobile mental health: A challenging research agenda. *European Journal of Psychotraumatology, 6*(s4), Article 27882. https://doi.org/10.3402/ejpt.v6.27882

O'Toole, M. S., Jensen, M. B., Fentz, H. N., Zachariae, R., & Hougaard, E. (2014). Emotion differentiation and emotion regulation in high and low socially anxious individuals: An experience-sampling study. *Cognitive Therapy and Research, 38*(4), 428–438. https://doi.org/10.1007/s10608-014-9611-2

Palmier-Claus, J. (2011). The clinical uses of momentary assessment. *Acta Psychiatrica Scandinavica, 124*(4), 241–242. https://doi.org/10.1111/j.1600-0447.2011.01761.x

Phillips, M. M., Phillips, K. T., Lalonde, T. L., & Dykema, K. R. (2014). Feasibility of text messaging for Ecological Momentary Assessment of marijuana use in college students. *Psychological Assessment, 26*(3), 947–957. https://doi.org/10.1037/a0036612

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., 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), Article e1000121. https://doi.org/10.1371/journal.pmed.1000121

Putnam, K. M., & McSweeney, L. B. (2008). Depressive symptoms and baseline prefrontal EEG alpha activity: A study utilizing Ecological Momentary Assessment. *Biological Psychology, 77*(2), 237–240. https://doi.org/10.1016/j.biopsycho.2007.10.010

Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Dieguez, A., Carry, M., Gelaude, D., Mosley, J. P., & Travers, J. (2016). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *Journal of Biomedical Informatics, 60*, 243–251. https://doi.org/10.1016/j.jbi.2016.02.002

Shear, K., Frank, E., Houck, P. R., & Reynolds, C. F. (2005). Treatment of complicated grief: A randomized controlled trial. *Journal of the American Medical Association, 293*(21), 2601–2608. https://doi.org/10.1001/jama.293.21.2601

Shear, M. K., Reynolds, C. F., Simon, N. M., Zisook, S., Wang, Y., Mauro, C., Duan, N., Lebowitz, B., & Skritskaya, N. (2016). Optimizing treatment of complicated grief. *JAMA Psychiatry, 73*(7), 685–694. https://doi.org/10.1001/jamapsychiatry.2016.0892

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine, 166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Spuij, M., Dekovic, M., & Boelen, P. A. (2015). An open trial of 'grief-help': A cognitive-behavioural treatment for prolonged grief in children and adolescents. *Clinical Psychology & Psychotherapy, 22*(2), 185–192. https://doi.org/10.1002/cpp.1877

Stelzer, E.-M., Zhou, N., Maercker, A., O'Connor, M.-F., & Killikelly, C. (2020). Prolonged grief disorder and the cultural crisis. *Frontiers in Psychology, 10*, Article 2982. https://doi.org/10.3389/fpsyg.2019.02982

Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile App Rating Scale: A new tool for assessing the quality of health mobile apps. *JMIR mHealth and uHealth, 3*(1), Article e27. https://doi.org/10.2196/mhealth.3422

Sveen, J., Johannesson, K. B., Cernvall, M., & Arnberg, F. K. (2018). Trajectories of prolonged grief one to six years after a natural disaster. *PLoS One, 13*(12), Article e0209757. https://doi.org/10.1371/journal.pone.0209757

Torous, J., Andersson, G., Bertagnoli, A., Christensen, H., Cuijpers, P., Firth, J., Haim, A., Hsin, H., Hollis, C., Lewis, S., Mohr, D. C., Pratap, A., Roux, S., Sherrill, J., & Arean, P. A. (2019). Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry, 18*(1), 97–98. https://doi.org/10.1002/wps.20592

Torous, J., & Roux, S. (2017). Patient-driven innovation for mobile mental health technology: Case report of symptom tracking in schizophrenia. *JMIR Mental Health, 4*(3), Article e7911. https://doi.org/10.2196/mental.7911

van Os, J., Verhagen, S., Marsman, A., Peeters, F., Bak, M., Marcelis, M., Drukker, M., Reininghaus, U., Jacobs, N., Lataster, T., Simons, C., 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

Wagner, B. (2013). *Komplizierte Trauer*. Springer.

Wagner, B., Knaevelsrud, C., & Maercker, A. (2006). Internet-based cognitive-behavioral therapy for complicated grief: A randomized controlled trial. *Death Studies, 30*(5), 429–453. https://doi.org/10.1080/07481180600614385

Wakefield, J. C. (2012). Should prolonged grief be reclassified as a mental disorder in DSM-5? Reconsidering the empirical and conceptual arguments for complicated grief disorder. *The Journal of Nervous and Mental Disease, 200*(6), 499–511. https://doi.org/10.1097/NMD.0b013e3182482155

Wang, K., Varma, D. S., & Prosperi, M. (2018). A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *Journal of Psychiatric Research, 107*, 73–78. https://doi.org/10.1016/j.jpsychires.2018.10.006

Whelan, P., Machin, M., Lewis, S., Buchan, I., Sanders, C., Applegate, E., Stockton, C., Preston, S., Bowen, R. A., Ze, Z., Roberts, C., Davies, L., Wykes, T., Tarrier, N., Kapur, S., & Ainsworth, J. (2015). Mobile early detection and connected intervention to coproduce better care in severe mental illness. In I. N. Sarkar, A. Georgiou, & P. M. Azevedo Marques (Eds.), *MEDINFO 2015: eHealth-enabled Health* (pp. 123-126). https://doi.org/10.3233/978-1-61499-564-7-123

Wichers, M., Simons, C. J. P., Kramer, I. M. A., Hartmann, J. A., Lothmann, C., Myin-Germeys, I., van Bemmel, A. L., Peeters, F., Delespaul, P., & van Os, J. (2011). Momentary assessment technology as a tool to help patients with depression help themselves. *Acta Psychiatrica Scandinavica, 124*(4), 262–272. https://doi.org/10.1111/j.1600-0447.2011.01749.x

Wykes, T., & Brown, M. (2016). Over promised, over-sold and underperforming? – E-health in mental health. *Journal of Mental Health, 25*(1), 1–4. https://doi.org/10.3109/09638237.2015.1124406

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

# Development and Psychometric Evaluation of the Hope in Medicine Scale

Lea Balthasar[1], Anne-Kathrin Bräscher[2], Ted J. Kaptchuk[3], Sarah K. Ballou[3,4], Tobias Kube[1,3]

**[1]** *Pain and Psychotherapy Research Lab, RPTU Kaiserslautern-Landau, Landau, Germany.* **[2]** *Department of Clinical Psychology, Psychotherapy, and Experimental Psychopathology, Johannes Gutenberg University of Mainz, Mainz, Germany.* **[3]** *Program in Placebo Studies, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA.* **[4]** *Division of Gastroenterology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA.*

Clinical Psychology in Europe, 2024, Vol. 6(1), Article e12001, https://doi.org/10.32872/cpe.12001

**Received:** 2023-05-21 • **Accepted:** 2023-10-30 • **Published (VoR):** 2024-03-28

**Handling Editor:** Winfried Rief, Philipps-University of Marburg, Marburg, Germany

**Corresponding Author:** Tobias Kube, Pain and Psychotherapy Research Lab, RPTU Kaiserslautern-Landau, Ostbahnstr. 10, 76829 Landau, Germany. Phone: +49 (0) 6341 280 35652. E-mail: tobias.kube@rptu.de

**Supplementary Materials:** Materials, Preregistration [see Index of Supplementary Materials]

## Abstract

**Background:** Hope is an integral, multi-dimensional part of seeking medical treatment. The aim of this study was to develop a self-report scale, the Hope in Medicine (HIM) scale, to measure different modes of hoping in relation to the course of symptoms, the effects of treatment, and supporting medical research.

**Method:** We examined the psychometric properties of the scale in a sample of 74 allergic rhinitis patients participating in a 2-week randomized-controlled trial comparing open-label placebos (OLP) with treatment as usual (TAU).

**Results:** The HIM scale had a Cronbach's α of .78. An exploratory factor analysis revealed four factors: realistic hope (i.e., hoping for specific positive outcomes such as improvement in symptoms), transcendent hope (i.e., non-directed hoping that things will turn out positively), utopian hope (i.e., hoping to contribute to greater knowledge), and technoscience hope (i.e., hoping for scientific breakthroughs). Speaking to the convergent validity of the scale, realistic hope was moderately related to treatment expectancies (r = .54); transcendent hope was related to optimism (r = .50), treatment expectancies (r = .37), self-efficacy (r = .36), and inversely correlated with pessimism (r = -.43). Hope subscales predicted neither course of symptoms nor impairment.

**Conclusion:** The HIM scale is a questionnaire with adequate internal consistency allowing to assess four modes of hoping. Preliminary results for its convergent validity are promising. Yet, further validation is needed.

## Keywords

hope, placebo, questionnaire, self-report, allergic rhinitis

### Highlights

- The HIM scale was developed to assess hope specifically in relation to treatment.
- The HIM scale allows to assess several modes of hoping.
- The study shows promising results concerning internal consistency and convergent validity of the HIM scale.

Hope has been the subject of scrutiny across academic disciplines, including philosophy, psychology, and medicine. It is an integral aspect of human life (Webb, 2007) and is often present in everyday life (e.g., hoping for a promotion or a general hopefulness for a bright future). When facing a serious illness, patients may hope for a remission of their symptoms and/or successful treatment outcomes. Even terminally ill, patients often maintain hope, e.g., hoping to preserve a good quality of life (e.g. Hagerty et al., 2005).

## Defining Hope

Although hope is a nearly ever-present phenomenon in human life, defining the construct is difficult. In several scholarly disciplines, many theories and definitions of hope have been proposed (for an overview see Kube et al., 2019; Webb, 2007). According to most definitions, hope involves desiring a future event or outcome with a low or unknown probability of fulfillment (Kube et al., 2019). Although there is a certain overlap between hope and expectations, people can differentiate between these constructs (Montgomery et al., 2003), with expectations relating to subjectively higher certainty of the desired outcome (Kube et al., 2019; Leung et al., 2009). Hope also resembles optimism defined as "generalized expectations of the occurrence of good outcomes in one's life" (Scheier & Carver, 1985, p. 239) and both involve positive affect towards the future (Bruininks & Malle, 2005). In contrast, pessimism describes anticipating bad outcomes (Scheier & Carver, 1985, p. 219) and is negatively correlated with optimism. However, people distinguish between hope and optimism: Compared to optimism, hope is directed at more important outcomes, with a smaller subjective likelihood of occurring, and less perceived personal control over the obtaining of the outcome (Bruininks & Malle, 2005).

Webb (2007) integrated hope theories and definitions and developed a model with five modes of hoping assigned to the two superordinate dimensions "goal-directed hope" and "open-ended hope". In a qualitative study, Eaves et al. (2014) applied Webb's frame-

work to a medical context. Chronic pain patients participating in a randomized-controlled trial (RCT) to evaluate Traditional Chinese Medicine were interviewed before treatment started and over the 18-month course of the RCT. Five modes of hoping emerged from the patients' answers: *realistic hope*, *wishful hope*, *utopian hope*, *technoscience hope,* and *transcendent hope*. This modes-of-hoping framework will be the theoretical basis of our hope scale.

### Definitions of the Modes of Hoping

Realistic hope describes "any hope that would be considered reasonable or probable based on current medical knowledge" (Eaves et al., 2014, p. 228). It includes, for example, hopes for minor symptom reductions, needing less medication or finding new techniques to manage pain (p. 229). There is a certain overlap between realistic hope and expectations, with realistic hope resembling the definition of the term "hope" in everyday language (e.g. "desire accompanied by expectation of or belief in fulfillment", Merriam-Webster, n.d.). Wishful hope comprises very high hopes which still can be fulfilled. For example, when patients expressed "hope for a cure" or hope "related to hearsay about miraculous outcomes experienced by others" (Eaves et al., 2014, p. 229). Although patients often considered these hopes to be unrealistic, they are in the realm of possibility and motivate chronically ill patients to seek further treatment. Utopian hope contains hoping that collective action might lead to a better future. In the context of medical and psychological research, utopian hope means that patients hoped that their participation in a research study would contribute to greater overall knowledge about the disease and would help others in the future (Eaves et al., 2014, p. 229). Especially utopian hope and realistic hope show a certain overlap with self-efficacy, i.e., the belief that a certain behavior will produce the desired outcome (i.e., outcome expectancies) combined with the confidence in one's ability to perform the required behavior, i.e., efficacy expectancies (Bandura, 1977). However, in utopian hope it is a desire rather than an expectancy. Realistic hope also includes outcomes independent from one's own actions.

Technoscience hope refers to hope for unforeseeable medical or scientific breakthroughs concerning treatment or cure. It also includes faith in science and medicine (Eaves et al., 2014, p. 229). An open, hopeful attitude not directed to a specific outcome or goal is classified as transcendent hope (Eaves et al., 2014, p. 230). Transcendent hope may also contain religious faith and openness to the future.

## Measuring Hope

Due to the variety of definitions, more than 30 measures exist to assess hope (Schrank et al., 2008). They differ in the number of assessed dimensions, whether they assess hope as a trait vs. as a state or globally vs. in a specific context. Although some widely used questionnaires have been developed for clinical settings and used in medical and

nursing research (e.g. the Herth Hope Index by Herth, 1992), none of them contains items to assess hope concerning the course of an illness, treatment success or quality of life. Instead, they assess hope more globally, such as having goals or plans for the future, feeling connected to others, and spirituality. Although these questionnaires might be valuable to assess a general hopefulness, possibly linked to positive health outcomes, they do not cover concrete hopes regarding illness or treatment. Additionally, they do not account for hopes concerning participating in a research study. Covering these aspects of hope is the main goal of the newly developed hope scale presented in this article, the Hope in Medicine (HIM) scale.

## Aims of the Present Study

In the present study, we aimed to preliminarily validate the HIM scale by examining its psychometric properties, i.e., its factorial structure, internal consistency and correlations with related constructs such as treatment expectancies, optimism, pessimism, and self-efficacy. We predicted the HIM scale to have an internal consistency of Cronbach's $\alpha \geq .70$. In terms of convergent validity, we predicted a moderate relationship ($.3 \geq r \leq .7$) of hope as assessed with the HIM scale with related constructs.

We examined these aspects in a RCT comparing the effects of open-label placebos + treatment as usual (subsequently referred to as "OLP") vs. treatment as usual (TAU) in allergic rhinitis patients. The main results of this RCT are reported elsewhere (Kube et al., 2022). Here, we focus on the validation of the scale that – in the context of the specific aforementioned RCT – assessed hope concerning the effects of placebo treatment and the course of allergic symptoms. In terms of the modes-of-hoping framework, we assessed hope regarding symptom improvement (realistic hope), hope for full remission of symptoms and/or being cured from allergic rhinitis in the future (wishful hope), and hope that taking part in a research study would contribute to greater knowledge about allergic rhinitis and its treatment (utopian hope). Furthermore, we assessed an open, hopeful attitude towards the future in general (transcendent hope) and hope for unforeseeable scientific breakthroughs concerning novel treatment options for allergic rhinitis (technoscience hope).

In the OLP literature there is a recent discussion whether hope might be a better explanatory mechanism for OLP effects than expectations (e.g. Kaptchuk, 2018). While positive expectations robustly predict effects in deceptive placebos (e.g. Enck et al., 2013), expectations do not predict OLP effects in most studies (e.g. Kleine-Borgmann et al., 2019; Pan et al., 2020). In RCTs, many participants do not report positive treatment expectations; instead, they often report hope (e.g. Eaves et al., 2014; Haas et al., 2022). Therefore, we tested whether hope predicted course of symptoms and quality of life in OLP and TAU to examine criterion validity.

# Materials and Method

## Scale Development

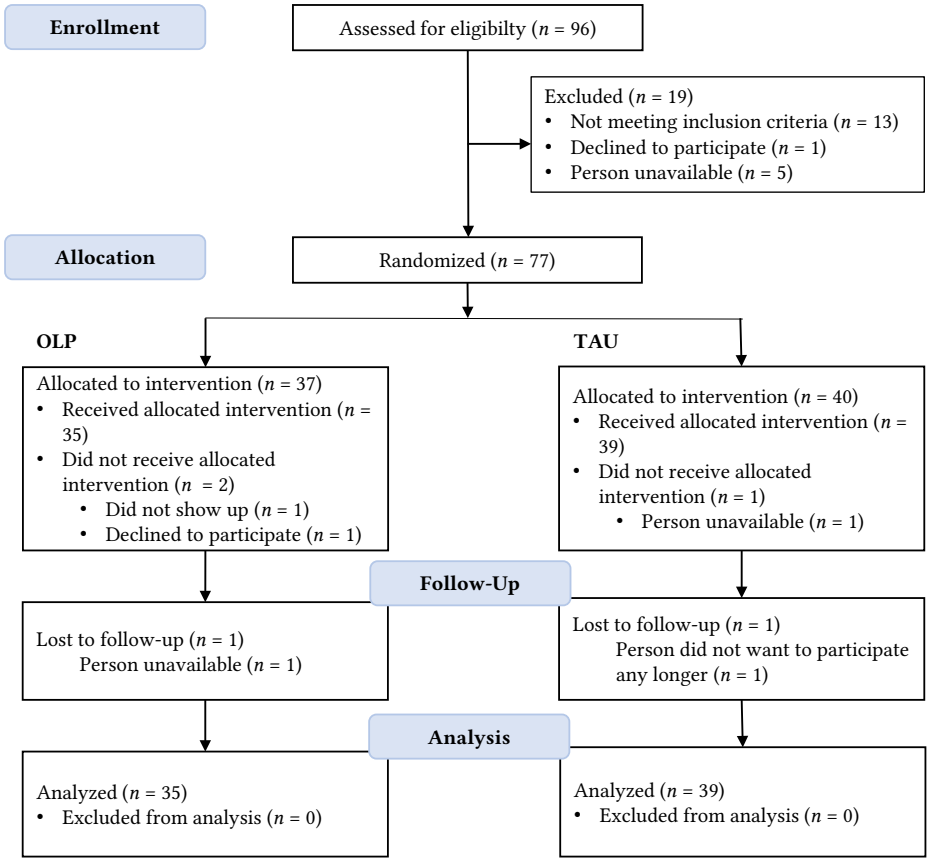
Items were generated by reviewing literature, especially the framework by Eaves et al. (2014, 2016), and by reviewing existing scales. Kube et al. (2019) stated that participating in a research study to evaluate novel treatments could include utopian hope (increasing knowledge), transcendent hope (being open to see what happens), and technoscience hope (hoping for unforeseen medical/scientific breakthroughs). These considerations were also taken into account when developing the items. Reviewing existing questionnaires assessing hope, we included two items (items no. 20, 21) of the Perceived Hope Scale by Krafft et al. (2019) in our questionnaire. Additionally, our scale development was guided by participants' answers in qualitative studies in which they were asked what they expected or hoped for prior to a new medical treatment (Di Blasi et al., 2005; Eaves et al., 2014, 2015, 2016; Kaptchuk et al., 2009). An initial item-pool of 22 items was developed by one of the authors (LB) in consultation with a second author (TK). Based on the discussion with two further authors (TJK, SKB), who have extensively addressed the concept of hope in both their scientific and clinical work, the wording of six items was slightly revised and five items were replaced entirely. As a result, the HIM scale consisted of 22 items (see Table A1 in Appendix A, Supplementary Materials) that were rated on a 6-point Likert-type scale ranging from 1 = *do not agree* to 6 = *completely agree*. Lower sum scores of the HIM scale indicate less hope. The scale was developed and administered in German (see Table A2 in Appendix A, Supplementary Materials), and it was translated into English for the present article.

## Participants

For the RCT, we aimed to reach a sample size of 90 participants, $f = .30$; $\alpha = .05$; $1\text{-}\beta = .80$, as pre-registered: https://aspredicted.org/ss6ag.pdf (see Kube et al., 2022). 96 participants were screened for study participation. Inclusion criteria were: diagnosed allergic rhinitis, at least 18 years old, and sufficient German language skills. Exclusion criteria were: diabetes, pregnancy, mental or neurological illnesses, and lactose intolerance (as the placebo tablets contained lactose). The final study sample consisted of 74 participants ($n = 54$ female, 73%; $M = 32.4$, $SD = 13.0$ years) as detailed in the CONSORT diagram (see Figure 1). The sociodemographic characteristics are presented for the two treatment conditions separately in Appendix B, Supplementary Materials. Participants were recruited via email lists, social media, and newspaper announcements. Data was collected between April and August 2021. Participants received either 10 € or course credit for their participation.

**Figure 1**

*CONSORT Diagram*




## Procedure

The RCT included a pretest (t1) and a posttest (t2) with a virtual clinical encounter each time. At t1, a psychology master student spoke to the participants about their allergic rhinitis and informed them about potentially positive effects of placebos. Afterwards, participants completed several questionnaires including the HIM scale. At the end of the pretest (i.e., after completing the questionnaires), participants were informed about their randomized treatment allocation to OLP vs. TAU. In the following 2 weeks, they took either OLP (two placebo tablets per day) + TAU or TAU alone. Participants in the TAU group only took their regular antiallergic medication (if there was any). After 2

weeks, there was a second clinical encounter (t2), in which the same psychology master student asked the participants about the course of their allergic symptoms and potential treatment effects, in addition to the second completion of questionnaires. All data was collected online via the survey platform SoSci Survey (Leiner, 2021). The study was approved by the local ethics committees of the University of Koblenz-Landau and the Johannes Gutenberg University of Mainz. All participants gave informed consent.

## Additional Measures

Severity and frequency of allergic rhinitis symptoms were assessed with the Combined Symptom Medication Score (CSMS; Pfaar et al., 2014) and a questionnaire by Schaefer et al. (2016, 2018). Allergy-related impairment of quality of life was assessed with the German version of the Mini Rhinoconjunctivitis Quality of Life Questionnaire (MiniRQLQ; Juniper et al., 2000). Treatment expectations were measured with the adapted version (Kube et al., 2021) of the Treatment Expectancy Scale by Kube et al. (2020). Self-efficacy, optimism, and pessimism were assessed with the Fragebogen zu Selbstwirksamkeit, Optimismus, Pessimismus Kurzform (SWOP-K9; Questionnaire for Self-Efficacy, Optimism, and Pessimism; Scholler et al., 1999). The instructions were adapted where necessary to refer to the last 2 weeks instead of the last week. These measures are detailed in Appendix C, Supplementary Materials.

## Statistical Analyses

Two participants dropped out between pretest and posttest. Therefore, we conducted an intention to treat analysis with expectation maximization using IBM SPSS Statistics (version 27) to estimate missing values concerning symptom severity, symptom frequency, and quality of life at t2 of those two participants. Power analyses were conducted in G*Power (version 3.1.9.6, Faul et al., 2007), and all other statistical analyses were performed in R (R Core Team, 2020). Alpha error levels were set at 5%.

We conducted an item analysis of the HIM scale and excluded items with a popularity index > 95 according to Dahl (1971; Kelava & Moosbrugger, 2020; see Appendix D, Supplementary Materials). To examine the factorial structure, an EFA was performed with the remaining items. The number of empirically relevant factors was determined with a parallel analysis according to Horn (1965) and an oblique rotation (promax) with these factors was performed as they were expected to be correlated. Items either loading > .30 on more than one factor or not loading at least .30 on any of the extracted factors were excluded (Boateng et al., 2018).

Internal consistency was determined by computing Cronbach's alpha. To determine the convergent validity, we computed correlations between hope and treatment expectancies, optimism, pessimism, and self-efficacy. Evaluating the criterion validity, three hierarchical regression analyses were conducted to test whether higher hopes at t1 were

associated with less symptom severity, frequency of symptoms, and impairment of quality of life after the 2-week intake of OLP or TAU. In the first step, the four hope subscales were included as predictors. In the second step, treatment condition (OLP vs. TAU) was added as an additional predictor. Residuals were plotted to examine whether the preconditions of homoscedasticity, normal distribution of residuals, and correct model specification were met.

# Results

## Item Analysis and Exploratory Factor Analysis

Means, standard deviations, and item popularity according to Dahl (1971) for all items are presented in Appendix E, Supplementary Materials. Items no. 4, 5, 10, and 11 were excluded from further analyses as they had a popularity index > 95. With the remaining 18 items, we performed an EFA. The Kaiser-Meyer-Olkin criterion was .65 and thus above the cutoff of .50 (Kaiser & Rice, 1974) and the Bartlett test was significant ($p < .001$), both suggesting that conducting an EFA is appropriate. A parallel analysis according to Horn (1965) yielded a four-factor solution, explaining 49% of the variance. Table 1 shows the factor loadings and communalities after oblique rotation.

Item no. 6 was excluded from further analyses as it loaded > .30 on more than one factor. Items no. 14 and 22 did not load on any of the four extracted factors, thus they were also excluded. Hence, the final HIM scale contains 15 items comprising four factors: *realistic hope*, *transcendent hope*, *utopian hope*, and *technoscience hope*. The assumed fifth factor *wishful hope* could not be confirmed. Table 2 shows the intercorrelations of these factors.

## Internal Consistency and Validity Analyses

Cronbach's α for the final 15-item HIM scale was .78.

### Convergent Validity

Table 3 shows the correlations of the four factors of the HIM scale with treatment expectancies, optimism, pessimism, and self-efficacy. Means and standard deviations of the hope subscales and the scales used for validation can be found in Appendix F, Table F1, Supplementary Materials.

### Criterion Validity

Contrary to our assumptions, none of the hope subscales predicted symptom severity (see Appendix F, Table F2, Supplementary Materials) or symptom frequency at t2 (see Appendix F, Table F3, Supplementary Materials). Taking Bonferroni correction into account, none of the hope subscales predicted impairment of quality of life at t2 (see

Table 4). The power to detect a small effect of $f^2 = 0.02$ ($\alpha = .05$, $N = 74$) was very low, though, 1-$\beta = .12$.

**Table 1**

*Results From the Exploratory Factor Analysis of the Hope in Medicine Scale*

| | Factor Loadings | | | | |
|---|---|---|---|---|---|
| **Item** | **1** | **2** | **3** | **4** | $h^2$ |
| **Factor 1: realistic hope** | | | | | |
| 1. I hope I will have less symptoms after taking the pills. | **.93** | -.10 | -.10 | .19 | .86 |
| 2. I hope that taking the pills will improve my quality of life. | **.86** | -.05 | -.09 | .16 | .74 |
| 3. I have hope that the pills will help me. | **.78** | .06 | -.08 | .11 | .63 |
| 7. I believe there is a small chance the placebos will make my symptoms go away completely. | **.39** | .10 | .09 | -.06 | .21 |
| **Factor 2: transcendent hope** | | | | | |
| 21. I am hopeful with regard to my life. | -.02 | **.89** | -.05 | -.04 | .75 |
| 19. I have the feeling that a lot of positive things await me in my future life. | -.01 | **.85** | .08 | .00 | .76 |
| 18. I have the feeling that my life will develop positively in the future. | .10 | **.76** | .00 | -.03 | .62 |
| 20. In my life hope outweighs anxiety. | .05 | **.44** | .23 | -.17 | .31 |
| **Factor 3: utopian hope** | | | | | |
| 8. I have hope that my participation in this study will contribute to a greater overall knowledge about the treatment of allergic rhinitis. | -.29 | .07 | **.96** | .04 | .89 |
| 9. I hope that my participation in this study will contribute to helping other people with allergic rhinitis in the future. | -.15 | .13 | **.82** | -.21 | .69 |
| 12. I do not believe that I will make an important contribution to the investigation of treatments for allergic rhinitis by participating in this study. (R) | .14 | -.09 | **.40** | .07 | .21 |
| 13. I think that studies like this can help us learn more about allergic rhinitis and its treatment. | .03 | -.05 | **.30** | .26 | .18 |
| **Factor 4: technoscience hope** | | | | | |
| 14. I hope that sooner or later an effective treatment for allergic rhinitis will be developed. | -.02 | -.04 | .21 | **.82** | .75 |
| 16. I hope for a scientific breakthrough in the treatment of allergic rhinitis. | -.02 | -.09 | .02 | **.70** | .48 |
| 17. I have hope that my allergic rhinitis will suddenly be cured someday. | .13 | .04 | -.13 | **.43** | .22 |
| **Excluded items** | | | | | |
| 6. I do not have specific expectations for the treatment with placebos but it is worth trying. | **-.55** | -.01 | .10 | **.32** | .34 |
| 15. I believe science will be able to find a treatment for almost every illness. | -.08 | .23 | -.12 | .17 | .09 |
| 22. I look hopelessly into the future. (R) | .06 | .08 | .08 | .06 | .04 |

*Note.* $N = 74$. Extraction method: principal factor analysis with oblique rotation. Factor loadings ≥ .30 are in bold. $h^2$ = communalities. Eigenvalue of factor 1 realistic hope = 3.20; eigenvalue of factor 2 transcendent hope = 1.91; eigenvalue of factor 3 utopian hope = 1.17; eigenvalue of factor 4 technoscience hope = 0.83.

**Table 2**

*Intercorrelations of the Four Factors*

| Factor | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1. Realistic hope | – | .24<br>[-0.11, 0.54] | .18<br>[-0.16, 0.49] | .23<br>[-0.11, 0.52] |
| 2. Transcendent hope | .24<br>[-0.11, 0.54] | – | .29<br>[-0.06, 0.58] | .09<br>[-0.21, 0.38] |
| 3. Utopian hope | .18<br>[-0.16, 0.49] | .29<br>[-0.06, 0.58] | – | .12<br>[-0.20, 0.43] |
| 4. Technoscience hope | .23<br>[-0.11, 0.52] | .09<br>[-0.21, 0.38] | .12<br>[-0.20, 0.43] | – |

*Note.* $N = 74$. 95% confidence intervals in square brackets using Bonferroni-Holm correction.

**Table 3**

*Correlations of the Four Factors of the Hope in Medicine Scale With Treatment Expectancies, Optimism, Pessimism, and Self-Efficacy*

| Hope Subscales | Treatment Expectancies | Optimism | Pessimism | Self-Efficacy |
|---|---|---|---|---|
| Realistic hope | .54*** | .13 | -.03 | .07 |
| | [0.23, 0.75] | [-0.20, 0.44] | [-0.29, 0.23] | [-0.23, 0.35] |
| Transcendent hope | .37* | .50*** | -.43*** | .36* |
| | [0.03, 0.64] | [0.18, 0.73] | [-0.68, -0.09] | [0.02, 0.63] |
| Utopian hope | .25 | .10 | .03 | -.01 |
| | [-0.10, 0.55] | [-0.21, 0.39] | [-0.25, 0.31] | [-0.24, 0.22] |
| Technoscience hope | .17 | .13 | .11 | .25 |
| | [-0.16, 0.47] | [-0.20, 0.43] | [-0.21, 0.41] | [-0.10, 0.54] |

*Note.* $N = 74$.

*$p < .05$. **$p < .01$. ***$p < .001$. $p$ values and 95% confidence intervals in square brackets using Bonferroni-Holm correction.

**Table 4**

*Hierarchical Regression Analysis With Impairment of Quality of Life at t2 as Dependent Variable*

| Predictor | *B* | *SE(B)* | β | *t* | *p* |
|---|---|---|---|---|---|
| **Model 1** | | | | | |
| Realistic hope | -0.08 | 0.12 | -.08 | -0.66 | .51 |
| Transcendent hope | 0.01 | 0.15 | .01 | 0.09 | .93 |
| Utopian hope | 0.43 | 0.18 | .30 | 2.43 | .02 |
| Technoscience hope | 0.06 | 0.13 | .06 | 0.49 | .62 |
| **Model 2** | | | | | |
| Realistic hope | -0.07 | 0.12 | -.07 | -0.59 | .56 |
| Transcendent hope | 0.01 | 0.15 | .01 | 0.09 | .93 |
| Utopian hope | 0.41 | 0.18 | .28 | 2.33 | .02 |
| Technoscience hope | 0.08 | 0.13 | .08 | 0.64 | .52 |
| Treatment condition | 0.42 | 0.23 | .21 | 1.84 | .07 |

*Note.* Model 1: $R^2 = .09$, $F(4, 69) = 1.74$, $p = .15$. Model 2: $R^2 = .14$, $F(1, 68) = 2.12$, $p = .07$.

## Discussion

We developed the Hope in Medicine (HIM) scale to assess hope specifically in a medical context and examined its psychometric properties in a 2-week RCT comparing the effects of OLP vs. TAU on symptoms of allergic rhinitis.

An exploratory factor analysis of the HIM scale yielded a four-factor solution with the factors "realistic hope", "transcendent hope", "utopian hope", and "technoscience hope". Thus, we could extract 4 of the 5 modes of hoping suggested by Eaves et al. (2014, 2016). However, we did not find a fifth factor relating to the mode "wishful hope". It might be difficult to assess "wishful hope" in allergic rhinitis patients in general as there is a variety of promising treatment options and a wide range of treatment outcomes which can be considered realistic, including full remission. Wishful hope might be more important in more desperate chronically ill patients with a lower likelihood of experiencing full remission. Nonetheless, based on the current data, the HIM scale allows to assess four of the intended modes of hoping that are relevant especially in medical settings and prior to starting a new treatment, speaking to its validity. Furthermore, the scale shows good internal consistency given the heterogeneity of the construct.

Analyses regarding the convergent validity of the HIM scale provided a mixed pattern of results. Realistic hope correlated significantly with treatment expectancies, speaking to its convergent validity as the items assessing realistic hope also relate to desired improvements following placebo treatment which could be considered probable. Transcendent hope was significantly correlated with treatment expectancies, optimism, and self-efficacy and inversely correlated with pessimism, suggesting convergent validity

as well. In contrast, the factors utopian hope and technoscience hope did not correlate with any of the assumed related constructs. However, this is not surprising because both utopian hope and technoscience hope are relatively specific aspects of hope. They neither refer to the general attitude towards the future (i.e., optimism, pessimism) nor to specific treatments outcomes (i.e., treatment expectancies). The SWOP-K9 assesses self-efficacy concerning mastering difficulties instead of a general confidence in being able to show a certain behavior to reach a specific outcome (Bandura, 1977). This might explain why utopian hope did not correlate with self-efficacy in the present study, although it shares a certain overlap with self-efficacy defined by Bandura (1977). Future research may examine whether a more substantial association between utopian hope and self-efficacy can be found with other measures of self-efficacy.

None of the hope subscales predicted symptom severity, symptom frequency or impairment of quality of life after the 2-week intake of OLP or TAU, questioning the predictive validity of the HIM scale. However, the statistical power in the present study was low due to the small sample size, possibly explaining the nonsignificant results. It is worth noting, though, that it is unclear so far whether hope is an explanatory mechanism for OLP effects. Only few studies, which show limitations concerning the assessment of hope, have examined the role of hope in OLP RCTs so far (Haas et al., 2022; Kube et al., 2020; Pan et al., 2020). Possibly, measurable hope does not matter in OLP effects or the course of symptoms. Instead, it might mainly instill the motivation to seek treatment.

## Limitations and Future Directions

The present study has several limitations: 1) The same sample was used for developing and validating the HIM scale. Thus, further validation in another independent sample is recommended. 2) Most items showed ceiling effects leading to limited variance which might explain some of the unexpected non-significant results concerning validation. The ceiling effects might be due to social desirability or a self-selection bias with only those patients expressing interest in the study who hoped to benefit from it. Alternatively, giving information about possible positive effects of placebos during the clinical encounter might have instilled hope. 3) The present study focused on allergic rhinitis and baseline allergic symptoms and baseline impairment were rather low. We assume that the HIM scale can be applied to other medical conditions except for life threatening illnesses. However, the external validity of the present RCT is rather limited. Therefore, it would be valuable to examine the HIM scale's validity in more severe or chronic diseases as hope might be more important in those cases, possibly leading to higher variance, increased explained variance, and higher correlations with measures of convergent validity. In future studies, criterion validity could be addressed by testing whether treatment conditions differ in certain hope subscales after treatment allocation. Content validity could be tested by examining the relation of certain hope subscales and the BIG-5 traits

(e.g., transcendent hope with openness). 4) The psychometric properties of the English version of the HIM scale should be tested, and larger sample sizes are recommended for future studies to increase statistical power.

## Conclusions

Since there has been a lack of measures assessing hope specifically in relation to medical treatment and symptom course, the HIM scale may fill this gap as it covers several modes of hoping in the context of starting a new treatment and participating in a research study. As validated in a sample of patients with allergic rhinitis, the scale shows good internal consistency and the preliminary results for its convergent validity are promising. Contrary to our hypotheses, however, hope was not related to greater symptom improvement following treatment. The current study is just a very first step into more systematically investigating the role of hope, which allows only some very cautious conclusions due to the small sample size and some other limitations.

**Funding:** The authors have no funding to report.

**Acknowledgments:** The authors have no additional (i.e., non-financial) support to report.

**Competing Interests:** The authors have declared that no competing interests exist.

**Author Contributions:** **Lea Balthasar**, conceptualization, methodology, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, visualization. **Anne-Kathrin Bräscher**, resources, writing – review and editing, supervision, project administration. **Ted J. Kaptchuk**, writing – review and editing. **Sarah K. Ballou**, writing – review and editing. **Tobias Kube**, conceptualization, methodology, resources, writing – review and editing, supervision, project administration.

**Ethics Statement:** This research was approved by the local ethics committees of the University of Koblenz-Landau and the Johannes Gutenberg University of Mainz. Informed consent was obtained from all participants included in the study.

**Twitter Accounts:** @DrSarahBallou

**Data Availability:** **Open Data:** The information needed to reproduce all of the reported results are not openly accessible. The data is available on request from the authors. **Code:** Code is not openly accessible. **Open Materials:** The information needed to reproduce all of the reported methodology is not openly accessible.

## Supplementary Materials

The Supplementary Materials include the following items:

- **The pre-registration protocol for the study.** The RCT was preregistered at AsPredicted (see Kube et al., 2021). The development and psychometric evaluation of the HIM scale was not preregistered.
- **Online appendices** (see Balthasar et al., 2024)
  - ***Appendix A.*** *Table A1-2.* Table A1 shows the English translation of the instructions and items of the HIM scale. Table A2 shows the German instructions and items of the HIM scale.
  - ***Appendix B.*** *Table B1.* Table B1 shows sociodemographic characteristics of the two treatment groups separately.
  - ***Appendix C.*** Detailed description of additional measures.
  - ***Appendix D.*** Formula to calculate the popularity index according to Dahl (1971).
  - ***Appendix E.*** *Table E1.* Table E1 shows the results of the item analysis.
  - ***Appendix F.*** *Table F1-3.* Table F1 shows means and standard deviations of the hope scales, treatment expectancies, self-efficacy, optimism, and pessimism. Table F2 shows the results of the hierarchical regression analysis with symptom severity as dependent variable. Table F3 shows the results of the hierarchical regression analysis with symptom frequency as dependent variable.

### Index of Supplementary Materials

Kube, T., Kirsch, I., Glombiewski, J. A., Witthöft, M., & Bräscher, A. (2021). *Effects of remotely provided open-label placebos on allergic symptoms* (ID #63631) [Pre-registration protocol]. AsPredicted. https://aspredicted.org/ss6ag.pdf

Balthasar, L., Bräscher, A., Kaptchuk, T. J., Ballou, S. K., & Kube, T. (2024). *Supplementary materials to "Development and psychometric evaluation of the Hope in Medicine scale"* [Online appendices]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.14058

## References

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Bruininks, P., & Malle, B. F. (2005). Distinguishing hope from optimism and related affective states. *Motivation and Emotion, 29*(4), 327–355. https://doi.org/10.1007/s11031-006-9010-4

Dahl, G. (1971). Zur Berechnung des Schwierigkeitsindex bei quantitativ abgestufter Aufgabenbewertung [Calculating the popularity index in quantitatively graded task assessment]. *Diagnostica, 17*(3), 139–142.

Di Blasi, Z., Crawford, F., Bradley, C., & Kleijnen, J. (2005). Reactions to treatment debriefing among the participants of a placebo controlled trial. *BMC Health Services Research, 5*(1), Article 30. https://doi.org/10.1186/1472-6963-5-30

Eaves, E. R., Nichter, M., & Ritenbaugh, C. (2016). Ways of hoping: Navigating the paradox of hope and despair in chronic pain. *Culture, Medicine and Psychiatry, 40*(1), 35–58. https://doi.org/10.1007/s11013-015-9465-4

Eaves, E. R., Ritenbaugh, C., Nichter, M., Hopkins, A. L., & Sherman, K. J. (2014). Modes of hoping: Understanding hope and expectation in the context of a clinical trial of complementary and alternative medicine for chronic pain. *EXPLORE, 10*(4), 225–232. https://doi.org/10.1016/j.explore.2014.04.004

Eaves, E. R., Sherman, K. J., Ritenbaugh, C., Hsu, C., Nichter, M., Turner, J. A., & Cherkin, D. C. (2015). A qualitative study of changes in expectations over time among patients with chronic low back pain seeking four CAM therapies. *BMC Complementary and Alternative Medicine, 15*(1), Article 12. https://doi.org/10.1186/s12906-015-0531-9

Enck, P., Bingel, U., Schedlowski, M., & Rief, W. (2013). The placebo response in medicine: Minimize, maximize or personalize? *Nature Reviews: Drug Discovery, 12*(3), 191–204. https://doi.org/10.1038/nrd3923

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Haas, J. W., Ongaro, G., Jacobson, E., Conboy, L. A., Nee, J., Iturrino, J., Rangan, V., Lembo, A., Kaptchuk, T. J., & Ballou, S. (2022). Patients' experiences treated with open-label placebo versus double-blind placebo: A mixed methods qualitative study. *BMC Psychology, 10*(1), Article 20. https://doi.org/10.1186/s40359-022-00731-w

Hagerty, R. G., Butow, P. N., Ellis, P. M., Lobb, E. A., Pendlebury, S. C., Leighl, N., Mac Leod, C., & Tattersall, M. H. N. (2005). Communicating with realism and hope: Incurable cancer patients' views on the disclosure of prognosis. *Journal of Clinical Oncology, 23*(6), 1278–1288. https://doi.org/10.1200/JCO.2005.11.138

Herth, K. (1992). Abbreviated instrument to measure hope: Development and psychometric evaluation. *Journal of Advanced Nursing, 17*(10), 1251–1259. https://doi.org/10.1111/j.1365-2648.1992.tb01843.x

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179–185. https://doi.org/10.1007/BF02289447

Juniper, E. F., Thompson, A. K., Ferrie, P. J., & Roberts, J. N. (2000). Development and validation of the Mini Rhinoconjunctivitis Quality of Life Questionnaire. *Clinical and Experimental Allergy, 30*(1), 132–140. https://doi.org/10.1046/j.1365-2222.2000.00668.x

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement, 34*(1), 111–117. https://doi.org/10.1177/001316447403400115

Kaptchuk, T. J. (2018). Open-label placebo: Reflections on a research agenda. *Perspectives in Biology and Medicine, 61*(3), 311–334. https://doi.org/10.1353/pbm.2018.0045

Kaptchuk, T. J., Shaw, J., Kerr, C. E., Conboy, L. A., Kelley, J. M., Csordas, T. J., Lembo, A. J., & Jacobson, E. E. (2009). "Maybe I made up the whole thing": Placebos and patients' experiences

in a randomized controlled trial. *Culture, Medicine and Psychiatry, 33*, 382–411. https://doi.org/10.1007/s11013-009-9141-7

Kelava, A., & Moosbrugger, H. (2020). Deskriptivstatistische Itemanalyse und Testwertbestimmung [Descriptive statistical item analysis and test score determination]. In H. Moosbrugger & A. Kelava (Eds.), *Testtheorie und Fragebogenkonstruktion* (pp. 143–158). Springer. https://doi.org/10.1007/978-3-662-61532-4_7

Kleine-Borgmann, J., Schmidt, K., Hellmann, A., & Bingel, U. (2019). Effects of open-label placebo on pain, functional disability, and spine mobility in patients with chronic back pain: A randomized controlled trial. *Pain, 160*(12), 2891–2897. https://doi.org/10.1097/j.pain.0000000000001683

Krafft, A. M., Martin-Krumm, C., & Fenouillet, F. (2019). Adaptation, further elaboration, and validation of a scale to measure hope as perceived by people: Discriminant value and predictive utility vis-à-vis dispositional hope. *Assessment, 26*(8), 1594–1609. https://doi.org/10.1177/1073191117700724

Kube, T., Blease, C., Ballou, S. K., & Kaptchuk, T. J. (2019). Hope in medicine: Applying multidisciplinary insights. *Perspectives in Biology and Medicine, 62*(4), 591–616. https://doi.org/10.1353/pbm.2019.0035

Kube, T., Hofmann, V. E., Glombiewski, J. A., & Kirsch, I. (2021). Providing open-label placebos remotely—A randomized controlled trial in allergic rhinitis. *PLoS ONE, 16*(3), Article e0248367. https://doi.org/10.1371/journal.pone.0248367

Kube, T., Kirsch, I., Glombiewski, J. A., Witthöft, M., & Bräscher, A.-K. (2022). Remotely provided open-label placebo reduces frequency of and impairment by allergic symptoms. *Psychosomatic Medicine, 84*(9), 997–1005. https://doi.org/10.1097/PSY.0000000000001110

Kube, T., Rief, W., Vivell, M.-B., Schäfer, N. L., Vermillion, T., Körfer, K., & Glombiewski, J. A. (2020). Deceptive and nondeceptive placebos to reduce pain: An experimental study in healthy individuals. *The Clinical Journal of Pain, 36*(2), 68–79. https://doi.org/10.1097/AJP.0000000000000781

Leiner, D. J. (2021). *SoSci Survey (Version 3.2.29)*. https://www.soscisurvey.de

Leung, K. K., Silvius, J. L., Pimlott, N., Dalziel, W., & Drummond, N. (2009). Why health expectations and hopes are different: The development of a conceptual model. *Health Expectations, 12*(4), 347–360. https://doi.org/10.1111/j.1369-7625.2009.00570.x

Merriam-Webster. (n.d.). Hope. In *Merriam-Webster.com Dictionary*. Retrieved July 16, 2022, from https://www.merriam-webster.com/dictionary/hope

Montgomery, G. H., David, D., DiLorenzo, T., & Erblich, J. (2003). Is hoping the same as expecting? Discrimination between hopes and response expectancies for nonvolitional outcomes. *Personality and Individual Differences, 35*(2), 399–409. https://doi.org/10.1016/S0191-8869(02)00202-7

Pan, Y., Meister, R., Löwe, B., Kaptchuk, T. J., Buhling, K. J., & Nestoriuc, Y. (2020). Open-label placebos for menopausal hot flushes: A randomized controlled trial. *Scientific Reports, 10*, Article 20090. https://doi.org/10.1038/s41598-020-77255-z

Pfaar, O., Demoly, P., Gerth van Wijk, R., Bonini, S., Bousquet, J., Canonica, G. W., Durham, S. R., Jacobsen, L., Malling, H. J., Mösges, R., Papadopoulos, N. G., Rak, S., Rodriguez del Rio, P., Valovirta, E., Wahn, U., & Calderon, M. A. (2014). Recommendations for the standardization of clinical outcomes used in allergen immunotherapy trials for allergic rhinoconjunctivitis: An EAACI Position Paper. *Allergy, 69*(7), 854–867. https://doi.org/10.1111/all.12383

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Schaefer, M., Harke, R., & Denke, C. (2016). Open-label placebos improve symptoms in allergic rhinitis: A randomized controlled trial. *Psychotherapy and Psychosomatics, 85*(6), 373–374. https://doi.org/10.1159/000447242

Schaefer, M., Sahin, T., & Berstecher, B. (2018). Why do open-label placebos work? A randomized controlled trial of an open-label placebo induction with and without extended information about the placebo effect in allergic rhinitis. *PLoS One, 13*(3), Article e0192758. https://doi.org/10.1371/journal.pone.0192758

Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4*(3), 219–247. https://doi.org/10.1037/0278-6133.4.3.219

Scholler, G., Fliege, H., & Klapp, B. F. (1999). SWOP-K9 – Fragebogen zu Selbstwirksamkeit-Optimismus-Pessimismus Kurzform [Questionnaire for self-efficacy, optimism, and pessimism]. *Leibniz-Institut für Psychologie (ZPID), Open Test Archive*. https://doi.org/10.23668/psycharchives.337

Schrank, B., Stanghellini, G., & Slade, M. (2008). Hope in psychiatry: A review of the literature. *Acta Psychiatrica Scandinavica, 118*(6), 421–433. https://doi.org/10.1111/j.1600-0447.2008.01271.x

Webb, D. (2007). Modes of hoping. *History of the Human Sciences, 20*(3), 65–83. https://doi.org/10.1177/0952695107079335

Clinical Psychology in Europe

PsychOpen GOLD

Letter to the Editor, Commentary

# Process-Based Therapy as a Novel Treatment Approach and Framework for Classifying Psychopathology

Stefan G. Hofmann[1], Steven C. Hayes[2]

**[1]** *Translational Clinical Psychology Group, Department of Psychology, Philipps University of Marburg, Marburg, Germany.* **[2]** *Department of Psychology, University of Nevada, Reno, NV, USA.*

Clinical Psychology in Europe, 2024, Vol. 6(1), Article e13727, https://doi.org/10.32872/cpe.13727

**Published (VoR):** 2024-03-28

**Corresponding Author:** Stefan G. Hofmann, Philipps-University of Marburg, Department of Psychology, Translational Clinical Psychology Group, Schulstraße 12, 35032 Marburg, Germany. E-mail: stefan.hofmann@uni-marburg.de

In a recent article, one of us co-authored a discussion paper comparing prominent classification frameworks (Rief et al., 2023). In the discussion, the article noted the following:

> "PBT is primarily a treatment approach, while the systems perspective is a broader framework for understanding mental disorders. While PBT draws on the systems perspective to inform its understanding of mental disorders, it is primarily focused on developing and implementing novel interventions. The systems perspective, on the other hand, seeks to provide a comprehensive understanding of mental disorders that can inform the development of a wide range of future treatments" (p. 27).

We would like to correct and clarify these statements. In fact, PBT is *not* primarily a specific treatment approach, and it *does* seek a broader understanding of mental and behavioral health. In essence, PBT provides a different and more idiographic perspective on systems approaches to clinical science. It begins with an idiographic focus on how processes of change combine in complex networks and can best be altered case by case, which is extended to nomothetic principles if and only if doing so maintains or increases idiographic fit: what we term an "idionomic" approach.

As we noted in one of our first publications introducing PBT (Hofmann & Hayes, 2019), we contend that modern clinical science needs to focus on the following question: "What core biopsychosocial processes should be targeted with this client given this goal in this situation, and how can they most efficiently and effectively be changed?" (p. 38).

Our proposed answer was an idionomic understanding of "the contextually specific use of evidence-based processes linked to evidence-based procedures to help solve the problems and promote the prosperity of particular people" (p. 38).

In the context of evolutionary science, adaptation or maladaptation is a function of variation, selection, and retention of biopsychosocial processes in given contexts. Any process can be helpful or hurtful depending on the person's history, goals, or circumstances. Processes are often functionally interconnected, forming a complex network that may differ in degree of abstraction and complexity.

We contend that a broader and more functional approach to mental health will come by viewing psychopathology as a complex system – evolution gone awry within networks of biopsychosocial processes in the life trajectories of individuals, that may then be corrected with intervention. When such knowledge is extended in an idionomic fashion PBT argues it will provide a comprehensive understanding of mental disorders that can inform the development of a wide range of future treatments.

We hope this clarifies the distinguishing features of PBT and other frameworks discussed in the article by Rief et al. (2023).

**Funding:** The authors have no funding to report.

**Acknowledgments:** The authors have no additional (i.e., non-financial) support to report.

**Competing Interests:** Dr. Hofmann receives financial support by the Alexander von Humboldt Foundation (as part of the Alexander von Humboldt Professur) and the Hessische Ministerium für Wissenschaft und Kunst (as part of the LOEWE Spitzenprofessur). He also receives compensation for his work as editor from SpringerNature and royalties and payments for his work from various publishers.

## References

Hofmann, S. G., & Hayes, S. C. (2019). The future of intervention science: Process-based therapy. *Clinical Psychological Science, 7*(1), 37–50. https://doi.org/10.1177/2167702618772296

Rief, W., Hofmann, S. G., Berg, M., Forbes, M. K., Pizzagalli, D. A., Zimmermann, J., Fried, E., & Reed, G. M. (2023). Do we need a novel framework for classifying psychopathology? A discussion paper. *Clinical Psychology in Europe, 5*(4), Article e11699. https://doi.org/10.32872/cpe.11699

EACLIPT

*Clinical Psychology in Europe* (CPE) is the official journal of the European Association of Clinical Psychology and Psychological Treatment (EACLIPT).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.