cpe.psychopen.eu | ISSN 2625-3410 | May 2026 | Volume 8 | Issue 2 |

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

May 2026 | Volume 8 | Issue 2

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

How to Define Psychological Therapy and Psychotherapy? – An Interdisciplinary Proposal

Winfried Rief[1], Julia Asbrand[2], Luisa Baumgärtner[3], Ulrike Dinger[4], Christoph Flückiger[5], Nina Heinrichs[6], Sabine C. Herpertz[7,8], Tania Lincoln[9], Silke Lipinski[10,11], Wolfgang Lutz[12], Bernhard Strauss[13], Svenja Taubner[14], Oliver Vorthmann[10,15], Jan Philipp Klein[16]

[1] *Division of Clinical Psychology and Psychotherapy, University of Marburg, Marburg, Germany.* **[2]** *Clinical Psychology in Childhood and Adolescence, University of Jena, Jena, Germany.* **[3]** *Wilhelm Wundt Institute of Psychology, University of Leipzig, Leipzig, Germany.* **[4]** *Clinical Institute for Psychosomatic Medicine and Psychotherapy, Medical Faculty and University Hospital Düsseldorf, Düsseldorf, Germany.* **[5]** *Clinical Psychology II, University of Kassel, Kassel, Germany.* **[6]** *Clinical Psychology in Children and Adolescents, Bielefeld University, Bielefeld, Germany.* **[7]** *Department of Psychiatry and Psychotherapy, University of Heidelberg, Heidelberg, Germany.* **[8]** *German Center for Mental Health (DZPG), Site Mannheim/Heidelberg/Ulm, Germany.* **[9]** *Clinical Psychology and Psychotherapy, University of Hamburg, Hamburg, Germany.* **[10]** *Trialogical Board of the German Center for Mental Health (DZPG), Berlin-Potsdam Site, Germany.* **[11]** *Clinical Psychology of Social Interaction, Humboldt-Universität zu Berlin, Berlin, Germany.* **[12]** *Clinical Psychology and Psychotherapy, Trier University, Trier, Germany.* **[13]** *Institute of Psychosocial Medicine, Psychotherapy and Psychooncology, University Hospital Jena, Jena, Germany.* **[14]** *Institute for Psychosocial Prevention and Psychotherapy, University Hospital of Heidelberg, Heidelberg, Germany.* **[15]** *DepressionsLiga, Berlin, Germany.* **[16]** *Department of Psychiatry, Psychosomatics and Psychotherapy, University of Lübeck, Lübeck, Germany.*

Clinical Psychology in Europe, 2026, Vol. 8(2), Article e22829, https://doi.org/10.32872/cpe.22829

Published (VoR): 2026-05-29

Corresponding Author: Winfried Rief, University of Marburg, Clinical Psychology and Psychotherapy, Gutenbergstraße 18, 35032 Marburg, Germany. E-mail: rief@uni-marburg.de

Supplementary Materials: Materials [see Index of Supplementary Materials]

Abstract

Background: The definition of psychological treatments and psychotherapy has various implications for communication within research and clinical care, as well as for legal issues — particularly in countries with psychotherapy acts. It can define what should be covered by public healthcare systems and who should be permitted to provide these services.

Method: We assembled a group of psychotherapists in Germany to develop an inclusive, scientific definition of psychotherapy. This definition should serve as an umbrella that is not limited to traditional treatment schools and is suitable for science, clinical care and legal acts. The group comprised people with different therapeutic backgrounds (e.g., CBT, psychodynamic and systemic therapy), experts from various disciplines (e.g., psychology, psychiatry and psychosomatic medicine), people focusing on different age groups (e.g., adults and youth), individuals with lived experience of psychotherapy and early career clinicians.
Results: A literature review led us to consider four aspects of a definition. What are the treatment means or 'tools'? (modality of treatment); how does the treatment work? (Proposed change mechanisms, assumptions about causal and evidence-based processes and concepts); who receives psychotherapy? (target group); and who provides psychotherapy? (Provider). Based on these aspects, we offer a definition of psychological therapy.
Discussion: Our suggested definition is not bound to specific treatment orientations, but is intended to encompass both traditional approaches and innovative developments. Our aim is to promote the evidence-based provision of psychological therapy within healthcare systems and legal frameworks.

Keywords

psychotherapy, definition, evidence-based treatments, transtheoretical approach, integration

Highlights

- A definition of psychological therapy/psychotherapy should address aspects such as treatment modality, processes and mechanisms of change, receiver and provider of treatments.
- We provide a suggestion for a definition of psychotherapy that is also suitable for legal regulations.

The definition of psychological treatments and psychotherapy has various implications, and should offer a framework for categorization within research and clinical care. Even more importantly, the definition of psychotherapy can have legal relevance, in particular in countries with psychotherapy acts. A psychotherapy act defines the role of psychological therapy in healthcare systems (including reimbursement) and who is permitted to provide them. In many countries, a need for legal regulations of psychological therapy was expressed (e.g., Flora, 2024). Therefore, the definition of psychological therapy has an important practical value, influencing the allocation of healthcare resources, regulations for professional training and access to treatment (e.g., Barkham et al., 2010). For the purpose of this editorial, we will use the terms 'psychological therapy' and 'psychotherapy' synonymously. They are tools used to provide treatment in a clinical context to clinical groups.

Interestingly, proposals for the definition of psychological therapies are less frequent than expected. Hodges et al. (2011) summarized reviews for definitions and concluded: "Surprisingly, we were not able to find an explicit definition of what constituted a 'psychological' intervention in any of the reviews". Although there are some attempts to define psychological therapy (see examples in the Supplement, e.g., Tolin et al., 2025), many of them show some weaknesses, are overly complex, not suitable for legal purposes, or are not feasible for a clear definition of boundaries.

In many countries, the definition of psychotherapy was based on the approval of a few general approaches ("schools") of psychotherapy (often a selection of cognitive behavioral therapy [CBT], psychodynamic treatments, sometimes also family-systemic interventions, or person-centered interventions). The approval of psychotherapeutic schools and traditions can be more restrictive (e.g., CBT, psychodynamic and interpersonal psychotherapy [IPT] in Sweden; www.socialstyrelsen.se), more liberal (e.g., 23 treatment approaches in Austria; www.psychotherapie.at/patientinnen/ueber-psychotherapie), with very different concepts about "evidence based", and different implications for public funding. Some experts criticize the strong orientation on traditional psychotherapy schools as "pre-scientific", and argue in favor of more general and transtheoretical concepts of psychotherapy (Goldfried, 2020). Other definitions are based on an understanding of a psychotherapeutic situation as a one-on-one in-person setting, which challenges the inclusion of some modern developments such as videotherapy, blended therapy or even group therapy. Furthermore, psychological therapy as an overall scientific and clinical construct emphasizes a commitment to dialogue across different psychological therapies and the involvement of networks of scientists and practitioners (e.g., Castonguay et al., 2019). This requires a process of exchange and dynamic developments across professions and orientations, instead of establishing different schools as parallel structures (Goldfried et al., 2019). In the pursuit of personalized, evidence-based healthcare, a focus on few treatment traditions can hinder dynamic developments and the implementation of new, evidence-based approaches. For example, a recent survey showed that psychotherapists hold several distinct prejudices towards those from other theoretical orientations, which makes cooperation in research and clinical practice difficult (Schröder et al., 2026). This also leads to fragmentation within the field. Furthermore, many modern developments in psychotherapy do not adhere to the categories of a few traditional schools, but instead integrate treatment procedures from different traditions (e.g., schema therapy, mentalization-based therapy MBT or CBASP) or describe innovative approaches that are evidence-based with an own theoretical framework beyond the traditional schools. In many settings, treatments with different theoretical backgrounds are integrated into one comprehensive treatment program (Bohus et al., 2012; Venturo-Conerly et al., 2023). Some new developments, such as transtheoretical approaches (Lutz & Rief, 2024), process-based psychotherapy (PBT; Hofmann & Hayes, 2019, 2024), competence-based

psychotherapy (Rief, 2021), or modular, mechanism-based psychotherapy (Herpertz & Schramm, 2022; Schramm et al., 2024) aim to overcome the traditional boundaries between theoretical frameworks. Cuijpers et al. (2025) outline that improvement of mental healthcare for depression needs to be done by multiple, incremental innovations, which requires an open and flexible system. Many of these innovations do not result from changes within specific school theories, but are motivated by better respecting other processes and context factors (e.g., societal and cultural factors; Wainberg et al., 2026).

These are aspects indicating the need for an umbrella definition of psychotherapy that is not based on single traditions, but that encompasses all interventions based on scientific principles including testable theories and empirical support for their effect. Furthermore, the definition should be aligned with the development of a measurement-based approach to psychological therapy (e.g., Delgadillo & Lutz, 2020). It should not be limited to a few theories, but should allow the consideration of all relevant theoretical and empirical approaches of modern psychological science and neuroscience. This definition should also facilitate innovation and help develop a psychotherapeutic system that allows for optimal personalization, or 'precision psychotherapy'.

The German Interdisciplinary Group of Psychotherapists

In Germany, we have assembled a group of psychotherapists to develop an inclusive definition of psychotherapy rooted in scientific principles. The group was supposed to comprise people with different therapeutic backgrounds (CBT, psychodynamic, systemic therapy), experts from various disciplines (psychology, psychiatry, psychosomatic medicine) and age foci (adults as well as youth), and individuals with lived experience of psychotherapy, as well as early career clinicians. The group (which is identical to the authors of this editorial) screened existing definitions, and discussed the aspects that should be covered in a modern definition of psychotherapy. In contrast to other definition proposals (see Supplementary Table 1), this definition should be suitable for legal purposes, and should focus on the relevance of psychotherapy in health care systems; therefore, a focus on prevention and treatment of clinical conditions was assumed.

The group agreed that a definition of psychotherapy should include at least four components:

1. What are treatment means, 'tools'? (Modality of treatment)
2. How does the treatment work? (Proposed change mechanisms; assumptions on causal and evidence-based processes and concepts)
3. Who is the receiver of psychotherapy? (Target group)
4. Who provides psychotherapy? (Provider)

Having defined these categories, we propose the definition shown in Table 1. This should not be considered a final definition suitable for all purposes, but rather a basis for discussion and national adaptations. In the following section, we will discuss the proposal in more detail, summarizing the reasons behind it and exploring potential variations.

Table 1

Proposal for a Definition of Psychotherapy

Category	*Psychotherapy is*
Modality	*the evidence-based application of psychological procedures,*
Mechanisms	*based on scientifically sound assumptions about psychological processes*
Target	*to classify, diagnose, prevent, treat, and rehabilitate clinically relevant syndromes, disorders, and illnesses*
By whom?	*by specifically educated and trained people.*

Ad 1) **Modality**: Psychotherapy is the evidence-based application of psychological procedures: this defines the tool used by psychotherapists. While other clinicians may use pharmacological, surgical, electrostimulation or other procedures, psychotherapy is characterized by the systematic use of psychological procedures. 'Psychological' is not intended as a description of professions, but as a description of the intervention modality.

Ad 2) **Mechanisms**: Psychotherapy is based on scientifically sound assumptions about psychological processes relevant for problem definitions and change. This part emphasizes that professional psychotherapists should apply existing scientific knowledge about psychological processes not only originating from one single theory, but from the whole richness of knowledge and theories of empirical psychology, neuroscience and psychotherapy research. It also specifies that merely demonstrating positive pre- and post-intervention effects is insufficient; professional psychotherapy must be grounded in the best current knowledge of the underlying processes, theories of disorders and changes (Borsboom et al., 2021). Of note, while psychological therapy necessarily requires the definition of relevant psychological mechanisms, this should also be embedded in a broad biopsychosocial understanding of disorders and change processes.

Ad 3) **Target Group:** When defining the scope of psychotherapy, most people primarily consider mental and behavioral disorders as target groups of psychological interventions. Indeed, psychotherapy has become the recommended first-line treatment for most of these conditions (as single treatment or in combination with other treatments) (Rief et al., 2024). However, psychotherapy is also indicated for other conditions that are not classified as mental or behavioral disorders, but can be found under other categories of classification systems. Examples include insomnia and chronic pain conditions, with psychological therapies being among the first line treatments, although the main treatment focus can be on physical functioning. Psychotherapy is further indicated for

several medical conditions typically considered "biomedical", ranging from behavioral change programs to illness coping in diseases such as diabetes, cardiovascular conditions, or functional somatic disorders. Therefore, many clinically relevant syndromes, disorders and illnesses are subject to psychotherapy. Nevertheless, psychotherapy is not indicated for every medical problem and every patient; therefore, specifying that psychotherapy is an evidence-based application limits its use to clinical conditions for which there is evidence for the effect of applying psychotherapy. Thus, also sub-threshold conditions (e.g., high risk profile for psychosis) can justify psychological interventions.

Ad 4) **By whom?** Psychotherapy is provided by people who have been educated and trained to do so. In many countries, this is further specified by legal acts, defining who is allowed to provide psychological treatments in the health care system. In Germany, this is now mainly limited to psychologists and physicians who have received training in psychotherapy as part of their university education, and who have undergone further post-graduate training/ specialization. According to the suggested definition, psychotherapy can be delivered as videotherapy and also be blended with digital interventions. In contrast, task shared interventions can be delivered by not specifically trained psychotherapists (Karyotaki et al., 2022), e.g., under the supervision of approved psychotherapists. Such a definition also excludes mere digital interventions or AI based interventions (chatbots etc.) not embedded in a person-guided therapeutic process.

Conclusion

In this editorial, we set out a framework for defining psychotherapy and make a proposal for a definition. We are aware that this is just a proposal and a starting point, and should be subject to further discussions. This definition is open to be modified according to national needs, future developments and other purposes, but we recommend considering the four criteria for definitions and adaptations. Our suggested definition is not bound to specific treatment orientations, but is intended to encompass traditional treatment approaches as well as new developments beyond these traditions. We aim to promote the evidence-based provision of psychological treatments in healthcare systems and legal acts at a national level that help provide this useful intervention to all patients in need of it.

Funding: WR receives research funding from the German Research Foundation DFG (CRC Treatment Expectation; Project-ID: 422744262; The Adaptive Mind, and for the excellence center LOEWE DYNAMIC from the Hessian Ministry of Science and Arts, grant No. LOEWE1/16/519/03/09.001[0009]/98)). **UD** receives funding from the German Research Foundation DFG (451696278), the Federal Ministry of Research, Technology and Space BMFTR (16SV9620), and the Heigl-Stiftung, Düsseldorf. **CF** receives funding from the Swiss National Science Foundation (ID: 205047). **WL** receives research funding from the German Research Foundation DFG (525286173 & 493169211 & 504507043), and via the German Center for Mental Health (DZPG, Tübingen). **TL** receives funding from the German Research Foundation DFG (RU 5389, RTG 2753). **NH** receives research funding from the European Union. **SCH** receives research funding from the German Research Foundation DFG (Project No 462340798 & 499083365 & 512007073) and via the German Center for Mental Health (DZPG, Site Mannheim/Heidelberg, Ulm). **BS** currently receives funding from the German Research Foundation (DFG), the Federal Ministry of Research, Technology and Space (BMFTR), and the German Innovation Fund. **JA** receives funding from the German Research Foundation DFG, Federal Ministry of Research, Technology and Space BMFTR and via the German Center for Mental Health (DZPG; Site Halle/Jena/Magdeburg). **ST** receives funding from the German Research Foundation (DFG), the Federal Ministry of Research, Technology and Space (BMFTR) and the German Center for Mental Health (DZPG, Site Mannheim/Heidelberg/Ulm).

Acknowledgments: We are grateful for comments of Andrea Benecke to former versions of this discussion.

Competing Interests: Winfried Rief is Editor-in-Chief of *Clinical Psychology in Europe*. Tania M. Lincoln is Subject Editor of *Clinical Psychology in Europe*.

Use of AI Tools: AI tools have been used for initial literature research; however, all references were checked for correctness. The authors take full responsibility for the text.

Supplementary Materials

The Supplementary Material includes examples of definitions of psychological treatments / psychotherapy (for access, see Rief et al., 2026S).

Index of Supplementary Materials

Rief, W., Asbrand, J., Baumgärtner, L., Dinger, U., Flückiger, C., Heinrichs, N., Herpertz, S. C., Lincoln, T., Lipinski, S., Lutz, W., Strauss, B., Taubner, S., Vorthmann, O., & Klein, J. P. (2026S). *Supplementary materials to "How to define psychological therapy and psychotherapy? – An interdisciplinary proposal"* [Examples of definitions of psychological treatments / psychotherapy]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.21895

References

Barkham, M., Hardy, G. E., & Mellor-Clark, J. (2010). *Developing and delivering practice-based evidence: A guide for the psychological therapies*. John Wiley & Sons.

Bohus, M., Herpertz, S. C., & Falkai, P. (2012). Modulare Psychotherapie – Rationale und Grundprinzipien. *Die Psychiatrie, 9*(2), 89–97. https://doi.org/10.1055/s-0038-1671776

Borsboom, D., Van Der Maas, H. L. J., Dalege, J., Kievit, R. A., & Haig, B. D. (2021). Theory construction methodology: A practical framework for building theories in psychology. *Perspectives on Psychological Science, 16*(4), 756–766. https://doi.org/10.1177/1745691620969647

Castonguay, L. G., Constantino, M. J., & Beutler, L. E. E. (2019). *Principles of change: How psychotherapists implement research in practice*. Oxford University Press.

Cuijpers, P., Harrer, M., & Furukawa, T. (2025). Assessing the strength of innovations in the treatment of depression. *The British Journal of Psychiatry, 227*(5), 810–813. https://doi.org/10.1192/bjp.2025.98

Delgadillo, J., & Lutz, W. (2020). A development pathway towards precision mental health care. *JAMA Psychiatry, 77*(9), 889–890. https://doi.org/10.1001/jamapsychiatry.2020.1048

Flora, K. (2024). The history of clinical psychology in Greece: A brief review – Legal deficiencies, practical dimensions and challenges for the future. *Clinical Psychology in Europe, 6*(4), Article e12515. https://doi.org/10.32872/cpe.12515

Goldfried, M. R. (2020). The field of psychotherapy: Over 100 years old and still an infant science. *Clinical Psychology in Europe, 2*(1), Article e2753. https://doi.org/10.32872/cpe.v2i1.2753

Goldfried, M. R., Pachankis, E., & Goodwin, B. J. (2019). A history of psychotherapy integration. In J. C. Norcross & M. R. Goldfried (Eds.), *Handbook of psychotherapy integration* (3rd ed., pp. 28-63). Oxford University Press. https://doi.org/10.1093/med-psych/9780190690465.003.0002

Herpertz, S. C., & Schramm, E. (2022). *Modulare Psychotherapie: Ein mechanismus-basiertes, personalisiertes Vorgehen*. Schattauer.

Hodges, L. J., Walker, J., Kleiboer, A. M., Ramirez, A. J., Richardson, A., Velikova, G., & Sharpe, M. (2011). What is a psychological intervention? A metareview and practical proposal. *Psycho-Oncology, 20*(5), 470–478. https://doi.org/10.1002/pon.1780

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

Hofmann, S. G., & Hayes, S. C. (2024). A process-based approach to transtheoretical clinical research and training. *Clinical Psychology in Europe, 6*(Special Issue), Article e11987. https://doi.org/10.32872/cpe.11987

Karyotaki, E., Araya, R., Kessler, R. C., Waqas, A., Bhana, A., Rahman, A., Matsuzaka, C. T., Miguel, C., Lund, C., & Garman, E. C. (2022). Association of task-shared psychological interventions with depression outcomes in low-and middle-income countries: A systematic review and individual patient data meta-analysis. *JAMA Psychiatry, 79*(5), 430–443. https://doi.org/10.1001/jamapsychiatry.2022.0301

Lutz, W., & Rief, W. (2024). Transtheoretical psychological therapy – New perspectives for clinical training and practice. *Clinical Psychology in Europe, 6*(Special Issue), Article e13891. https://doi.org/10.32872/cpe.13891

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

Rief, W., Asmundson, G. J. G., Bryant, R. A., Clark, D. M., Ehlers, A., Holmes, E. A., McNally, R. J., Neufeld, C. B., Wilhelm, S., Jaroszewski, A. C., Berg, M., Haberkamp, A., & Hofmann, S. G. (2024). The future of psychological treatments: The Marburg Declaration. *Clinical Psychology Review, 110*, Article 102417. https://doi.org/10.1016/j.cpr.2024.102417

Schramm, E., Elsaesser, M., Jenkner, C., Hautzinger, M., & Herpertz, S. C. (2024). Algorithm-based modular psychotherapy vs. cognitive-behavioral therapy for patients with depression, psychiatric comorbidities and early trauma: A proof-of-concept randomized controlled trial. *World Psychiatry, 23*(2), 257–266. https://doi.org/10.1002/wps.21204

Schröder, J., Trautmann, S., Töpfer, N. F., Rubel, J. A., Schweizer, K., Hermans, B. E., Mora, M. S., & Kauff, M. (2026). Theoretical orientations and the stereotype content model – Are prejudices barriers to psychotherapy integration? *Clinical Psychology in Europe, 8*(1), Article e17975. https://doi.org/10.32872/cpe.17975

Tolin, D. F., Grasso, D., Boness, C. L., Beck, J. G., Keane, T. M., Leichsenring, F., Olatunji, B. O., Otto, M. W., & Weinand, J. (2025). A proposed definition of psychological treatment and its relation to empirically supported treatments. *Clinical Psychology, 32*(3), 213–225. https://doi.org/10.1037/cps0000220

Venturo-Conerly, K. E., Reynolds, R., Clark, M., Fitzpatrick, O. M., & Weisz, J. R. (2023). Personalizing youth psychotherapy: A scoping review of decision-making in modular treatments. *Clinical Psychology, 30*(1), 45–62. https://doi.org/10.1037/cps0000130

Wainberg, M. L., Alpert, J. E., Benton, T. D., Clarke, D. E., Drexler, K., Fung, K. P., Gogtay, N., Malaspina, D., O'Keefe, V. M., & Oquendo, M. A. (2026). The future of DSM: A strategic vision for incorporating socioeconomic, cultural, and environmental determinants and intersectionality. *The American Journal of Psychiatry*. Advance online publication. https://doi.org/10.1176/appi.ajp.20250875

EACLIPT

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

zpid Leibniz-Institut für Psychologie

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

Clinical Psychology in Europe

PsychOpen GOLD

Scientific Update and Overview

Check for updates

A Scoping Review of Amenable Patient-Specific Predictors of Treatment Failure in the Treatment of Anxiety and Depressive Disorders

Vivian Peerbooms[1,2], Th. Michael van den Boogaard[3], Matthias J. Wieser[2], Colin van der Heiden[1,2]

[1] *Department of Anxiety Disorders, Outpatient Treatment Center PsyQ, Rotterdam, The Netherlands.* **[2]** *Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, The Netherlands.* **[3]** *Department of Affective Disorders, Outpatient Treatment Center PsyQ, The Hague, The Netherlands.*

Clinical Psychology in Europe, 2026, Vol. 8(2), Article e17335, https://doi.org/10.32872/cpe.17335

Received: 2025-03-14 • **Accepted:** 2026-01-06 • **Published (VoR):** 2026-05-29

Handling Editor: Cornelia Weise, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Corresponding Author: Vivian Peerbooms, Max Euwelaan 70, 3062 MA Rotterdam, The Netherlands. Telephone: +31883574960. E-mail: vivian.peerbooms@psyq.nl

Abstract

Background: By identifying predictors of treatment failure that are susceptible to change (amenable), we can move towards studying ways to decrease the odds of treatment failure, e.g., by targeting these predictors before treatment, adapting interventions, accordingly, choosing more suitable treatments, or preparing patients better for psychotherapy. While treatment success, within anxiety and depressive disorders, has been studied extensively, it seems that treatment failure is overlooked, even while we know that about one third of the treatment population shows no benefit in treatment.

Method: In order to review the available body of knowledge concerning amenable patient-specific predictors for treatment failure, we conducted a literature search in PubMed, PsycInfo, Embase, and Medline, following the Prisma-ScR guidelines. Thirty articles met the inclusion criteria and are summarized in this review. Conclusions were drawn for scientific and clinical implications.

Results: Predictors of treatment failure that are replicated or are significant in multiple studies are low treatment expectancy, high neuroticism, low use of social support, low outcome expectancy, and low perceived social support. Treatment failure is hard to define, and very few studies are replicated. There are predictors that are studied in multiple articles, but they are measured with different instruments, or in very small or specific patient samples, therefore it is difficult to compare findings from different studies.

Conclusions: There are no predictors that stand out as overall strong amenable predictors of treatment failure. Possible predictors are high neuroticism, low treatment expectancies, and low use of social support. Future research should focus on replicating studies to confirm these predictors of treatment failure.

Keywords

treatment failure, predictors, scoping review, anxiety disorders, depressive disorders

Highlights

- Reducing treatment failure might improve treatment quality and shorten duration.
- Possible predictors are high neuroticism, low treatment expectancies, and low use of social support.
- There are no predictors that stand out as overall strong amenable predictors of treatment failure.
- Replication studies are necessary to identify predictors of treatment failure.

With a worldwide prevalence of 3.8% and 4%, anxiety and depressive disorders are the most common mental health disorders (World Health Organization [WHO], 2023a, 2023b). Despite the availability of effective treatments, treatment failure remains a major problem. Average response rates are still unsatisfactory, with 48% response for patients with major depressive disorder (Cuijpers et al., 2014) and 50% response of CBT in anxiety disorders (Loerinc et al., 2015). These figures emphasize the need to identify factors that predict the (in)effectiveness of treatment. Understanding these predictors can help reduce treatment failure by adapting interventions, selecting more suitable treatments, or better preparing patients for psychotherapy.

Remarkably, despite the abundance of studies reporting on predictors of treatment success (protective factors), relatively few studies report on treatment failure (risk factors; Dandachi-FitzGerald et al., 2023). However, treatment failure is often not well operationalized or clearly defined and tends to be used as an umbrella term for unwanted outcomes such as drop-out, non-response, attrition, deterioration, or poor treatment outcome (Oasi & Werbart, 2020). This lack of consensus complicates research into this subject. Similarly, the term response rates, which is often used to describe clinical effectiveness, lacks a consistent definition and calculation method across studies.

So far, studies into predictors of poor treatment outcome or non-response mainly confine themselves to demographic or disorder-specific features (e.g., duration or severity of symptoms) associated with treatment failure (Edmonds et al., 2018; McDevitt-Petrovic et al., 2020). Most of these predictors are the focus of the therapy itself and not changeable with other interventions (e.g., decreasing severity of symptoms is one important goal of therapy). However, amenable predictors may be more useful in clinical practice, as they can be influenced before therapy starts to prevent treatment failure.

Eilertsen and Eilertsen (2023) suggest that focusing on changeable predictors is preferable, also for ethical reasons, because such knowledge can help to adapt treatments proactively. For example, van den Boogaard (2012) showed that drop-out rates in interpersonal therapy for depression decreased when patient-treatment compatibility was improved through a short intervention before therapy. This may as well be the case for more predictors like motivation, treatment attitude, or expectancies about treatment.

Although little is known regarding the mechanism of change in psychotherapy, a key to the black box of psychotherapy might be the distinction of trait-like and state-like components of the mechanisms of change (Zilcha-Mano, 2021). Mechanisms of change, or process variables, are events or constructs that change during therapy and whose change can lead to subsequent changes in outcome. Focusing on trait-like components as baseline for therapy and state-like components that are needed to influence the course of treatment can lead to personalised treatment recommendations. Knowing how treatment works and what mechanisms of change work for whom, can improve our treatments and eliminate ingredients of treatment that do not work (Zilcha-Mano, 2021). Therefore, finding out if there are also these mechanisms of change in play for treatment failure can also enhance the personalisation of therapy.

This scoping review provides an overview of the outcomes of studies on predictors of treatment failure (defined as treatment failure, deterioration, non-response, or drop-out for psychotherapy) in the treatment of adult patients with anxiety and depressive disorders. We chose to use this broad definition of treatment failure to capture all relevant studies addressing this topic as comprehensively as possible. Narrowing the definition too strictly carries the risk of excluding well-conducted research that could contribute valuable insights into the predictors of poor treatment outcomes. Next to that, we specifically focused on predictors that are patient-related and patient-specific, as such features are expected to be more amenable to change before starting therapy than demographic or disorder-specific features, and (thus) may be useful targets to reduce treatment failure.

Method

Literature Search

In order to review the available body of knowledge concerning patient-specific predictors for treatment failure, a literature search in PubMed, PsycInfo (OVID interface, 1806 onwards), Embase (OVID interface, 1974 onwards), and Medline, and Pre-Medline (OVID interface, 1946 onwards) was conducted with help from a research librarian. This search was aimed at articles that combined the following search criteria and terms using subject headings (MeSH for Medline, Emtree for Embase, and APA thesaurus for PsycInfo), if available: predictors, treatment failure, depressive or anxiety disorders, and psychotherapy. Treatment failure was further defined as poor outcome, non-response, drop-out,

deterioration, or not successful as the search criteria. The literature search was confined to English, German, and Dutch. A draft of the search strategy in Embase is included in Table 1.

Table 1

Draft Search in Embase on November 14, 2017

#[a]	Keywords used in the search	Results[b]
1	exp *anxiety disorder/	93842
2	*anxiety/	48269
3	1 or 2	137512
4	*depression/ or *agitated depression/ or *atypical depression/ or *dysphoria/ or *dysthymia/ or *endogenous depression/ or *involutional depression/ or *late life depression/ or *major depression/ or *masked depression/ or *melancholia/ or *"mixed anxiety and depression"/ or *organic depression/ or *perinatal depression/ or *postoperative depression/ or *premenstrual dysphoric disorder/ or *puerperal depression/ or *reactive depression/ or *recurrent brief depression/ or *seasonal affective disorder/ or *treatment resistant depression/	167039
5	3 or 4	281358
6	treatment failure/	91108
7	("poor outcome" or failure or failing or deteriorat* or worse* or unsucces* or "nonsucces*" "not succesful" or "no success" or drop-out* or drop out* or regress* or quit* or non-respons* or "non repsons*").mp.	2684336
8	deterioration/	34929
9	6 or 7 or 8	2684336
10	5 and 9	36039
11	prediction/ or "prediction and forecasting"/ or adverse outcome/ or forecasting/ or prediction/ or predictive validity/ or predictive value/ or prognosis/	1004236
12	10 and 11	3518
13	exp *psychotherapy/	116310
14	12 and 13	144
15	limit 14 to (dutch or english or german)	143

Note. Search words ending with / used the thesaurus (subject headings) of Embase, using "" forces the search to use words like "and" or "no" as a search term instead of a command. * Before the word is a focus command, which indicates that the search focuses on articles where the search word is the main topic. *Behind the word is a truncation command and indicates that it searches for variations on a word with different suffixes. Exp = explodes, meaning that it expands the search results of terms entered and includes more specific related topics. .mp = multiple fields, indicating these search terms are searched only in the most useful fields (e.g., title, abstract, keywords).

[a]number referring to a single search. [b]number of articles found with this specific search term on that specific date.

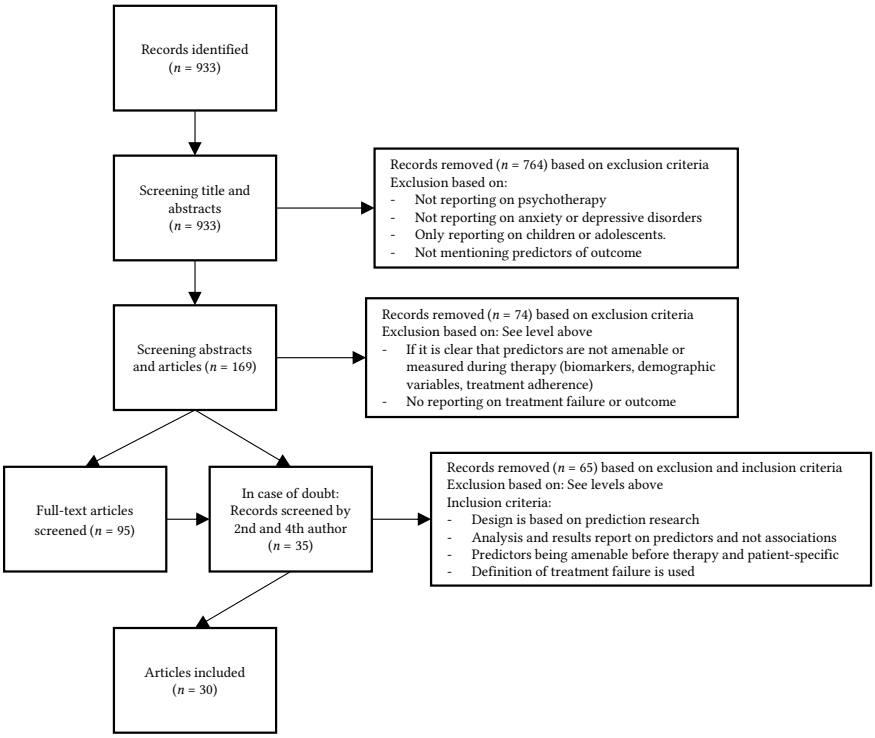
The last search was conducted on October 2, 2025, earlier searches were conducted in March 2022 and November 2017. We included studies examining a form of psychotherapy among adult populations (older than 18 years) with anxiety and depressive disorders.

Selection Process

We found 933 articles within this search. Figure 1 shows the PRISMA flow chart. Out of all these articles, the first author (VP) made a further selection. First, all titles and abstracts were checked for exclusion criteria. They were excluded if there was no mentioning of psychotherapy, anxiety, or depressive disorders, or prediction of outcome. They were also excluded if they only reported on child and adolescent studies.

Figure 1

Prisma Flow Chart



Next, the remaining 169 articles were screened by reading the abstracts and the conclusions for meeting exclusion criteria. Articles were excluded based on the above-described exclusion criteria, and when predictors were not amenable or were measured during

therapy and not at the start of therapy. Articles were also excluded when there was no mentioning of treatment failure or outcome. The remaining 95 papers were inspected more precisely by reading them fully, in order to check for reported amenable predictors of treatment failure. If it was not clear what kind of predictors were used and how they were measured, the methods and results sections were read carefully to inspect the analyses in order to find out if the variables could be labeled as predictors. Articles were included when they reported on treatment failure, predictors being amenable before therapy, and patient-specific, and when studies were based on prediction analyses. To decrease the risk of bias, in case of doubt, the second and fourth co-author also read the articles before including or excluding them. This was the case in 35 articles. In total, 30 papers met the inclusion criteria and were included in this review. Table 2 summarizes these articles.

Primary Outcomes

The main outcome of this review will be the predictors of treatment failure. A predictor was included if the variable was amenable to change and patient-specific, such as motivation or treatment attitude. As such, biological markers, demographic variables, comorbidity, or severity of complaints were excluded. We included only predictors of treatment failure, not predictors of follow-up or risk of relapse.

To improve readability, we categorized the predictors into categories based on both their definition and the instrument used to measure them in the studies included in the review (e.g., Motivation measured with the URICA). For instance, seeking social or emotional support, as mentioned by Grilo et al. (1998) and Marquett et al. (2013), was defined as a coping style and measured with a coping questionnaire. In contrast, perceived social support, mentioned by Marquett et al. (2013), Schilling et al. (2021), and Solomonov et al. (2021), was measured with instruments focusing on how patients perceive social support. This construct was therefore considered distinct from coping. Because of this conceptual distinction, we placed these seemingly similar constructs in different categories to ensure that each predictor reflected the specific theoretical framework and measurement approach used in the studies. Through this process, we specified seven predictor categories and one category "other" for predictors that did not fit elsewhere. The identified predictor categories are personality factors, coping strategies, motivational aspects, attributions, treatment attitude and expectancies, interpersonal relationships, and others.

Table 2

Overview of Included Articles

Article	Measurement of treatment failure	Sample size[a]	Treatment	Target group[b]	Predictors and measurement
Arndt et al. (2020)	Drop-out: Termination of treatment before week 12 of the intervention	Total: 1013 Intervention: 509 Control: 504	12-week internet-based CBT intervention and Control group	Adults (18-65) with depressive symptoms	Treatment Attitude (Attitudes Towards Psychological Online Interventions Questionnaire), Physical health (Short-Form Health Survey)
Bélanger et al. (2017)	Drop-out: discontinuation of therapy. Difference made between drop-out before treatment, in the first 7 sessions and after 8 sessions	Total: 77	CBT	Adults (18-65) Panic disorder with agoraphobia	Treatment expectations (Process Expectations Questionnaire), Dyadic adjustment (Dyadic Adjustment Scale)
Blom et al. (2007)	Drop-out. (Outcome: residual gain & norm score)	Total: 193	12 sessions IPT, medication and minimal contact, 12 sessions IPT and medication, 12 sessions IPT and pill-placebo	Age >18 Non-psychotic, non-bipolar major depressive disorder	Personality traits (NEO-Five Factor Inventory)
Chambless & Steketee (1999)	Drop-out: Leaving treatment before receiving at least 10 sessions of treatment	Total: 101	22 sessions in 16 weeks	Adults (18-65) with obsessive compulsive disorder or panic disorder with agoraphobia	Perceived criticism (Perceived Criticism Measure), hostility (Camberwell Family Interview), expressed emotion (Relatives Reactions Questionnaire, Composite Measure of Emotional Overinvolvement)
Chambless et al. (1997)	Residual gain: change with treatment adjusted for pretreatment severity levels	Total: 64	Group CBT	Social Phobia	Treatment expectancy (Treatment Expectancy Scale)

Article	Measurement of treatment failure	Sample size[a]	Treatment	Target group[b]	Predictors and measurement
Critchfield et al. (2007)	End state functioning: normal range score on zero to three of the six measures	Total: 24 Cognitive therapy: 10 Applied relaxation: 6 Combination: 8	14 sessions CBT, three different variants of CBT: cognitive therapy, applied relaxation, and combination	Generalized anxiety disorder	Interpersonal process between therapist and client. (Structural Analysis of Social Behavior coding system)
Grilo et al. (1998)	Drop-out: not completing the 11-session therapy	Total: 162	Individual 11- session CBT, pharmacotherapy or combined treatment	Panic disorder uncomplicated or with mild agoraphobia	Illness/treatment attributions (Treatment Attitude Measure; Etiological Model Questionnaire) Coping styles (Ways of Coping Checklist) Personality styles (Wisconsin Personality Inventory)
Hoyer et al. (2016)	Non-response: Less than 31% reduction of anxiety symptoms Drop-out: all patients who stopped treatment or assessments	Total: 244	Up to 30 individual sessions cognitive therapy	Age: 18-70 Primary diagnosis social anxiety disorder	Personality dimensions (Tri-Dimensional Personality Questionnaire) Self-esteem (subscale of the Frankfurt Self Concept Scales) Shame (subscale of Test of Self-conscious Affects) Interpersonal problems and attachment style (Inventory of Interpersonal Problems)
Johnson et al. (2014)	Drop-out: Missing more than one or two sessions (depending on the treatment)	Total: 74 Virtual reality exposure: 32 Exposure group therapy: 42	8 sessions virtual reality exposure therapy or exposure group therapy	Social anxiety disorder with a primary fear of public speaking	Stereotype confirmation concerns (Stereotype Confirmation Concerns Scale)
Keefe et al. (2021)	Drop-out: not completing at least 16 sessions in 12 weeks	Total: 200 Psychodynamic therapy: 80 CBT: 80 Applied relaxation therapy: 40	CBT & Psychodynamic therapy & applied relaxation training total of 24 sessions in 12 weeks	Adults (18-70) Panic disorder with or without agoraphobia	Treatment expectancies (Expectancy Scale)

Clinical Psychology in Europe
2026, Vol. 8(2), Article e17335
https://doi.org/10.32872/cpe.17335

Article	Measurement of treatment failure	Sample size[a]	Treatment	Target group[b]	Predictors and measurement
Keijsers et al. (1994)	Treatment failure: improvement equal or less than 30%	Total: 51	18 sessions exposure in vivo and exposure with response prevention	Obsessive compulsive disorder	Motivation for treatment (Willingness to Participate Scale of the Nijmegen Motivation List)
LeBeau et al. (2013)	Improvement of clinical severity ratings	Total: 84 CBT: 48 Acceptance and commitment therapy: 36	12 sessions CBT or acceptance and commitment therapy	Anxiety disorder Age 18-60	Treatment expectations (modified form of The Credibility/Expectancy Questionnaire)
Lin & Farber (2021)	Latent growth mixture modelling Treatment Outcome Package depression score	Total: 63	Different psychotherapies for over 9 months (i.e. psychodynamic, CBT, dialectical behavior therapy)	Depression	Self-concealment (Self-Concealment Scale)
Lutz et al. (2019)	Drop-out: non-consensual and non-recommended termination of therapy	Total: 1,234	Average 30.85 sessions individual psychotherapy	Age: 14-76 Different DSM-IV diagnoses with SCID-I	Personality style (Persönlichkeits-Stil- und Störungs-Inventar (PSSI-K)) Treatment expectations by patient and therapist (one-item question) Interpersonal problems (Outcome Questionnaire 30, Questionnaire for the Evaluation of Psychotherapeutic Progress)
Marker et al. (2019)	Drop-out: if session 10, 11 and 12 were not attended	Total: 58	12 sessions transdiagnostic Group-CBT for anxiety	Anxiety disorders Age: 19-58	Readiness to change (University of Rhode Island Change Assessment Scale) Motivation/change talk (Client Language Easy Rating Coding System)

Article	Measurement of treatment failure	Sample size[a]	Treatment	Target group[b]	Predictors and measurement
Marquett et al. (2013)	Significant clinical improvement: no diagnosis (MINI) or non-depressed range on the questionnaires	Total: 60	12 sessions individual CBT over 3-4 months	Age: 60 and above Major or minor depression and dysthymic disorder	Impact of stressful events (Elders Life Stress Inventory, Impact of Events Scale -6, The Integration of Stressful Life Event Scales), Social support (Abbreviated Duke Social Support Index), Locus of control (Rotter's Locus of Control theory), Personality (Big Five Inventory), Coping style (The Brief Cope)
Miller et al. (1996)	Non-response: score of 11 or more on the Hamilton Depression Rating Scale after 26 weeks of treatment	Total: 61	Up to 26 weeks of treatment, medication, IPT, or a combination	Elderly with recurrent depressive disorder	Perception of illness. (Perception of Illness Scale)
Moggia et al. (2020)	Growth mixture modelling (with Clinical Outcome in Routine Evaluation—Short Form B & Beck Depression Inventory-II)	Total: 108	7 sessions CBT group therapy and after that Individual CBT or Individual dilemma focused therapy	Age 18-70, Major depressive disorder or dysthymic disorder	Self-ideal discrepancy (Repertory Grid Technique: system to interpret the self, others and the world)
Moradveisi et al. (2014)	Drop-out: not completing therapy	Total: 100 Behavioral activation: 50 Anti-depressant medication: 50	16 sessions in 12 weeks behavioral activation or 12 weekly sessions anti-depressant medication	Major depressive disorder Age 18-60	Preference/attitudes towards therapy (Preference-Attitude Questionnaire)
Parker et al. (1986)	Improvement of depression	Total: 91 Used for predictive value of coping questionnaire: 48	Psychiatric consults	Depressive disorder	Coping behavior (Coping Questionnaire)

Article	Measurement of treatment failure	Sample size[a]	Treatment	Target group[b]	Predictors and measurement
Renaud et al. (2014)	Change in clinical global impression score	Total: 256	Average of 19 CBT sessions	Anxiety or depressive disorder	Treatment expectancy (Suitability for Short-Term Cognitive Therapy)
Safren et al. (1997)	Improvement after treatment Drop-out	Total: 113	Group CBT	Social phobia	Expectations of treatment (Reaction to Treatment Questionnaire)
Schilling et al. (2021)	Not on track: Failure boundary per session, based on Hopkins-Symptom-Checklist-11 scores	Total: 413 Control group: 157 Feedback group: 256	CBT	Adults with different diagnoses	Motivation (Assessment for Signal Clients), social support (Assessment for Signal Clients), emotion regulation (Affective Style Questionnaire)
Schindler et al. (2013)	RCI & quality associated drop-out, not completed number of allowed sessions	Total: 193	CBT	Major depressive disorder or dysthymic disorder	Treatment expectancies (one-item question)
Solomonov et al. (2021)	Latent Growth Mixture Models Early non-response: minimal change in depression severity	Total: 221 Problem-solving therapy: 107 Supportive therapy: 111	12 weeks problem-solving therapy or supportive therapy	Adults >60 years Non-psychotic major depression disorder	Neuroticism (NEO-Personality Inventory) Treatment expectancy (4-item Treatment Rationale Scale) Perceived social support (four subscales of the Duke Social Support Index)
Steketee et al. (2011)	Drop-out: completed less than 18 sessions	Total: 39 Uncontrolled pilot trail: 10 Waitlist controlled trial: 29	22 sessions cognitive therapy	Obsessive compulsive disorder	Personality traits (Personality Diagnostic Questionnaire-4) Motivation (University of Rhode Island Change Assessment Scale) Treatment expectancy (Expectancy Rating)
Strauss et al. (2017)	Drop-out: premature treatment discontinuation	Total: 412 CBT: 213 Psychodynamic therapy: 199	25 individual sessions psychodynamic therapy 30 sessions CBT	Age 18-70 Primary diagnosis: social anxiety disorder	Experiences in close relationships (Experiences in Close Relationships-Revised Questionnaire)

Article	Measurement of treatment failure	Sample size[a]	Treatment	Target group[b]	Predictors and measurement
Vogel et al. (2006)	Improvement between pre- and posttreatment measures	Total: 37	Twice weekly individual exposure and response prevention sessions for 6 weeks	Obsessive compulsive disorder with overt compulsions	Treatment expectancy (one-item question) Motivation to change (The University of Rhode Island Change Assessment Scale)
Westra (2011)	Improvement between pre- and posttreatment measures on Penn State Worry Questionnaire	Total: 38 Motivational Interviewing: 19 No Motivational interviewing: 18	14 hours CBT group 6 weekly 2-hour sessions followed by 2 1-hour sessions Motivational Interviewing pretreatment condition: 4 individual weekly sessions	Generalized anxiety disorder Age 18-66	Motivation (Change Questionnaire), Motivation to change (Client Motivation for Psychotherapy Scale), Resistance to therapy (Client Resistant Code)
Zilcha-Mano et al. (2016)	Drop-out: failure to complete the 16-week treatment protocol	Total: 156 Short-term Psychodynamic therapy: 51 Medication: 55 Placebo: 50	16 weeks 20 sessions short-term psychodynamic therapy, medication (sertraline) or placebo.	Major depressive disorder	Working alliance expectations (Working alliance inventory), Interpersonal Problems (Inventory of Interpersonal Problems-Circumplex)

[a]Sample size is total number of included patients, some articles specify these in different subsamples. [b]Age is always 18 plus, if not otherwise described.

Clinical Psychology in Europe
2026, Vol. 8(2), Article e17335
https://doi.org/10.32872/cpe.17335

Results

An important problem in the literature on treatment failure is the lack of consensus on its definition, which makes comparisons of studies difficult. This problem is reflected in the studies included in this review. Eleven of the included articles (37%) defined treatment failure as insignificant improvement in complaints, four of them used an additional component to define non-response, such as norm scores or a decrease percentage. Four studies (13%) defined treatment failure as residual gain, i.e., the individual difference in improvement of complaints controlled for the expected difference based on the pre-test score, indicating that treatment failure is viewed as a continuous variable in these studies. Two of these studies used an additional component to define treatment failure, such as norm scores or a decrease in percentage. Another three studies (10%) used norm scores of their primary outcome measure to determine treatment failure, one study (3%) used the reliable change index as formulated by Jacobson and Truax (1991) to define treatment failure, whereas one study (3%) stated that treatment failure was less than 50% decrease in complaints. In the remaining ten of the studies (34%), treatment failure was defined as drop-out from treatment.

In this results section, we will describe the findings of all studies per predictor category. In Table 3, we specify for each predictor, the papers that report them as predictors of treatment failure.

Personality Factors

Six studies examined personality factors in relation to treatment failure. Three of them used large samples of patients with depressive disorders (Blom et al., 2007; Marquett et al., 2013; Solomonov et al., 2021). Solomonov et al. (2021) studied the Big Five personality traits in problem-solving therapy and supportive therapy for late-life depression (age > 60) with executive dysfunction. Blom et al. (2007) used a sample of depressive adults following interpersonal psychotherapy. Both Blom et al. (2007) and Solomonov et al. (2021) found that only higher scores on neuroticism are a predictor for treatment failure. The other personality traits of the Big Five (i.e., extraversion, openness to experiences, conscientiousness, and agreeableness) were not found to be predictive of treatment failure (defined as drop-out and early non-response; Solomonov et al., 2021). Marquett et al. (2013) found that a lower score on openness to experience is a predictor of treatment failure (no significant clinical improvement) in a small sample of late-life depression (age > 60) treated with CBT.

Table 3

Summary of Predictors of Treatment Failure

Predictor	Deterioration	Drop-Out	Non-Response
Personality traits			
Histrionic personality style		+ Lutz et al. (2019)	
Neuroticism		+ Blom et al. (2007)	+ Solomonov et al. (2021)
Openness to experiences			- Marquett et al. (2013)
Obsessive personality style		- Lutz et al. (2019)	
Coping strategies			
Use and seeking social or emotional support		- Grilo et al. (1998)	- Marquett et al. (2013)
Self-consolation			+ Parker et al. (1986)
Motivation			
Motivation	- Schilling et al. (2021)		- Westra (2011) - Keijsers et al. (1994)
Resistance to therapy			+ Westra (2011)
Change talk		- Marker et al. (2019)	
Attributions			
To life stressors		+ Grilo et al. (1998)	

Predictor	Deterioration	Drop-Out	Non-Response
Treatment attitude and expectancies			
Treatment attitudes		- Grilo et al. (1998)	
		- Arndt et al. (2020)	
Outcome expectancy		- Schindler et al. (2013)	
		- Solomonov et al. (2021)	
		- Keefe et al. (2021)	
Expectations of a strong alliance		- Zilcha-Mano et al. (2016)	
Process expectations		+ Bélanger et al. (2017)	
Interpersonal relationships			
Emotional overinvolvement and hostility		+ Chambless & Steketee (1999)	
Impairment of interpersonal relationships		+ Lutz et al. (2019)	
Perceived social support			- Solomonov et al. (2021)
Vindictive tendencies			+ Zilcha-Mano et al. (2016)
Perceptions of acceptation by the therapist			- Solomonov et al. (2021)
Other			
Negative impact of stressful events			+ Marquett et al. (2013)
External locus of control			+ Marquett et al. (2013)
Assign blame for stressful events to others			- Marquett et al. (2013)
Physical health		- Arndt et al. (2020)	
Stereotype confirmation concerns		+ Johnson et al. (2014)	

Note. - = low score on the predictor, + = high score on the predictor.

Three studies focused on personality styles or indicators instead of traits (Grilo et al., 1998; Hoyer et al., 2016; Lutz et al., 2019). Only a more histrionic personality style and a less obsessive personality style predicted a higher probability of treatment failure (drop-out) in CBT for several diagnoses, with mostly affective, personality, and anxiety disorders (Lutz et al., 2019). Other predictors like harm-avoidance, self-esteem, shame, novelty seeking, dominance, and reward dependence were not found as predictors of treatment failure (treatment response and drop-out) in cognitive therapy for social phobia (Hoyer et al., 2016). Neither were personality styles, self-reported features from personality disorders from an internal perspective, such as avoidant or narcissistic personality styles, predictors of treatment failure in cognitive behavioral therapy, medication therapy, or combination treatment for panic disorder (Grilo et al., 1998).

Coping Strategies

Several coping strategies were studied as predictors of outcome. It is remarkable that all studies used different instruments to measure coping styles or strategies.

The use of seeking social or emotional support is the only coping strategy that was found to be predictive of treatment failure (drop-out and no clinically significant improvement) in more than one study (Grilo et al., 1998; Marquett et al., 2013). Samples used in these studies were a small sample of depressive older adults (age >60) following CBT (Marquett et al., 2013) and patients with uncomplicated panic disorder or with mild agoraphobia following CBT, pharmacotherapy, or combined treatment.

Other coping strategies were only studied as a predictor in one article and in different patient samples. Of those strategies, only higher self-consolation (with behavior like spending money, eating more, and drinking more alcohol) was associated with treatment failure (a high likelihood of not improving) after psychotherapy in depressed patients (Parker et al., 1986). Other coping strategies turned out to be not predictive of treatment failure. The Grilo et al. (1998) study revealed that confrontive coping, distancing, self-control, acceptance of responsibility, escape avoidance, problem solving, positive reappraisal, problem focusing, self-blame, wishful thinking and avoidance were not associated with treatment failure (drop-out) in the treatment of patients with a panic disorder. Finally, Lin and Farber (2021) found that self-concealment (the tendency to conceal distressing information aiming at self-protection and avoidance of stigma) measured at the baseline of psychotherapy, was not predictive of treatment failure (non-improvement) in depressive complaints.

Motivational Aspects

Findings about motivation as a predictor of treatment failure are contradictory. For instance, Keijsers et al. (1994) found that poor motivation predicted treatment failure (no improvement) in obsessive compulsive disorder (OCD) treatment. In line with these

findings Schilling et al. (2021) found that a sudden drop in therapy motivation was predictive for treatment failure (deterioration of complaints) in CBT for different diagnosis, although therapy motivation did not differ between on-track and not on-track patients. However, motivation was found not to be predictive of treatment failure (no improvement) of CBT worry treatment (Westra, 2011). All three studies used different questionnaires for measuring motivation.

Readiness for change, measured with the URICA, is seen as a different conceptualization of motivation. However, readiness for change was not predictive of treatment failure (no improvement and drop-out) in small samples of CBT for OCD and anxiety disorders (Marker et al., 2019; Steketee et al., 2011; Vogel et al., 2006).

Finally, change talk and counter change talk i.e., patient statements indicating support or opposition to change, an observational measure of motivation (Marker et al., 2019), did not predict treatment failure (drop-out) during early sessions of therapy in which psychoeducation and cognitive restructuring were offered. However, low change talk did predict treatment failure (drop-out) during exposure sessions later in CBT for anxiety disorders. Next to that, Westra (2011) found that low motivation to change and higher resistance to therapy, considered by Westra (2011) to be an indirect measure of motivation, were predictive of treatment failure (no improvement) in a small GAD-sample treated with CBT.

Attributions

There are only two studies on attributions as predictor of treatment failure. They both study a different sort of attribution. Attributing their panic disorder to life stressors predicted treatment failure (drop-out) for patients receiving treatment for panic disorder (Grilo et al., 1998). Perception of health, defined as the way people rate their overall physical health, did not predict treatment failure (non-response) in a group of late-life depressive patients who received IPT (Miller et al., 1996).

Treatment Attitude and Expectancies

Several studies found that treatment expectancy did not predict treatment failure (Chambless et al., 1997; LeBeau et al., 2013; Lutz et al., 2019; Renaud et al., 2014; Safren et al., 1997; Steketee et al., 2011). These outcomes are replicated in several studies with big sample sizes and different diagnosis.

On the other hand, outcome expectancy is found as a predictor of treatment failure in multiple studies (Keefe et al., 2021; Schindler et al., 2013; Solomonov et al., 2021). Patients with a more negative outcome expectancy were found to be more likely to have treatment failure (drop-out) for depression (Schindler et al., 2013; Solomonov et al., 2021). In a study comparing three treatments for panic disorder (CBT, focused psychodynamic therapy and applied relaxation), Keefe et al. (2021) found that outcome expectancies

measured at session two were not predictive of treatment failure (drop-out) in the overall study. However, they did find that in the CBT group, lower outcome expectancies at session two were predictive of treatment failure (drop-out), which was not found for the other treatment conditions. Contrary to these findings, Vogel et al. (2006) did not find outcome expectancy as a predictor of treatment failure in a small sample of OCD patients with overt compulsions.

There are also predictors related to treatment expectancies that were only found in one study but not replicated in other studies. Patients with lower expectations of a strong alliance had a higher risk at treatment failure (drop-out) during supportive-expressive therapy than during medication or placebo therapy (Zilcha-Mano et al., 2016). In the early phase of panic disorder treatment, in which psychoeducation and cognitive restructuring were offered, higher process expectations (i.e., expectations on the therapy process and role expectations) were predictive of treatment failure (drop-out), but anxiety expectations (i.e., expectations on having panic-related symptoms) were not. In the behavioral phase of treatment, where exposure and anxiety-provoking exercises took place, anxiety and process expectations were not predictive of treatment failure (drop-out) (Bélanger et al., 2017).

Treatment attitude implies the way people think of certain therapies or views they have on psychotherapy in general. Treatment attitude is studied in depressive and panic disorder patient populations with large sample sizes (Arndt et al., 2020; Grilo et al., 1998; Moradveisi et al., 2014). Moradveisi et al. (2014) found that treatment attitudes are no predictor of treatment failure. Contrary to that, treatment failure (attrition, or the likelihood of drop-out) could be predicted by less favorable or negative treatment attitudes towards the offered treatment, found in an internet-based intervention for depression and panic disorder treatment (Arndt et al., 2020; Grilo et al., 1998).

Interpersonal Relationships

Interpersonal relationships are a broad concept and include for example social support and attachment characteristics. Several studies found aspects of interpersonal relationships to be predictive of treatment failure. The only predictor within this category studied in multiple studies is perceived social support (Marquett et al., 2013; Schilling et al., 2021; Solomonov et al., 2021). Whereas Schilling et al. (2021) and Marquett et al. (2013) did not find perceived social support as a predictor of treatment failure in depression and other diagnosis, Solomonov et al. (2021) found low perceived social support to be predictive of treatment failure (early non-response) in psychotherapy for late-life depression.

Other interpersonal factors that were found as predictors for treatment failure are emotional overinvolvement from relatives and hostility coming from relatives (Chambless & Steketee, 1999) as well as higher impairment of interpersonal relationships (Lutz et al., 2019). Both were found to predict higher rates of drop-out in CBT for several

diagnosis. Higher vindictive tendencies in interpersonal relationships were found to be predictive of treatment failure (drop-out) in pharmacotherapy for depressive disorder, but not in other treatment groups (Zilcha-Mano et al., 2016). Lastly, Solomonov et al. (2021) found that perceptions of the therapist as less accepting are predictive of treatment failure (early non-response) in psychotherapy for late-life depression.

However, not all studies found interpersonal relationships to be predictive of treatment failure. Although low end state functioning groups had higher levels of interpersonal hostility, Critchfield et al. (2007) found no predictors in interpersonal process behaviors in a small sample of patients with generalized anxiety disorder (GAD) receiving CBT. Further, dyadic adjustment, which focuses on consensus, satisfaction, cohesion, and expression of affection within a romantic relationship, did not predict treatment failure (dropout) in CBT for panic disorder (Bélanger et al., 2017). Perceived criticism was no predictor of treatment failure in treatment for patients with OCD or panic disorder (Chambless & Steketee, 1999). Hoyer et al. (2016) did not find interpersonal problems and attachment style as predictors of treatment failure in cognitive therapy for patients with social anxiety disorder. Finally, Strauss et al. (2017) found that partner-related attachment anxiety and avoidance in romantic relationships are no predictors of treatment failure (drop-out) in both CBT and psychodynamic therapy for patients with social anxiety disorder.

Other Categories

Several predictors that were studied in the included papers could not be classified under one of the categories of predictors discussed so far. Notably, none of those predictors were studied in multiple studies, whereas a number of the target groups in these studies were very specific, for example late life depression, or social anxiety disorder with a primary fear of public speaking.

First of all, patients who experience a high negative impact of stressful events are more likely to experience treatment failure (non-response) to CBT for late-life depression (Marquett et al., 2013). The same was true for an external locus of control. Interestingly, also patients who do *not* tend to assign blame for stressful events to others (decreased external blame) turned out to be more likely to experience treatment failure (non-response) (Marquett et al., 2013). Further, Arndt et al. (2020) studied different drop-out patterns in an internet-based depression intervention and found that low physical health increases the risk to belong to the drop-out group, whereas Johnson et al. (2014) found that stereotype confirmation concerns, defined as being afraid of confirming certain stereotypes leading to negative evaluations, are associated with higher risk of treatment failure (drop-out) in social anxiety disorder treatment. Moggia et al. (2020) measured self-ideal discrepancy, which is the difference between the rating of the ideal self and the self as it is now, self-others discrepancy, explained as how someone views him/herself in respect to others, and a dilemmatic construction of the self, which means that a person desires a

change in one construct of the self that correlates with an undesirable change in another construct of the self. All of these measures appeared not to be predictive of treatment failure (non-improvement) in psychotherapy for depression. Finally, Schilling et al. (2021) found that emotion regulation was no predictor of treatment failure (deterioration of complaints) after CBT for different diagnosis.

Discussion

Summary of Results

In this paper, we reviewed studies on predictors of treatment failure (also defined as non-response, deterioration, and drop-out) for depressive and anxiety disorders. We specifically focused on amenable patient-specific predictors, such as motivation or coping styles. Knowledge of such factors might help us reduce treatment failure and thus improve treatment outcomes. Like Steketee and Chambless (1992) and Eilertsen and Eilertsen (2023), we encountered several problems within the existing literature. One important problem is the fact that both treatment failure and predictors are measured differently, which makes studies difficult to compare. Another problem is a lack of replication studies, and if predictors were replicated, this was done in studies using small or specific patient samples (e.g., elderly, social phobia with fear of speaking in public, or OCS with overt compulsions). Consequently, it is difficult to conclude the overall impact of these predictors.

However, a few predictors that stand out are treatment expectancy (replicated in three studies with a big sample size and different patient samples), neuroticism (replicated in two studies with the same instrument in different patient samples), the use of social support (replicated in two studies with different measuring instruments), readiness to change (replicated in three studies, in very small patient samples), outcome expectancy (two studies with different measuring instruments), and perceived social support (three studies, but only in late-life depression found as a predictor).

Focusing more on the different definitions of treatment failure, we see that about half of the included articles focus on drop-out and the other half on non-response. There are two articles that mention deterioration. Because drop-out, non-response, and deterioration are different constructs, we also looked at the difference in predictors per treatment failure definition. We see that almost none of the predictors are replicated within these treatment failure categories. The only predictors that remain are readiness to change and treatment expectancy. Readiness to change was studied in three different studies, two of which focused on drop-out, and both studies found that it was no predictor of drop-out. Treatment expectancy was studied in six different studies, all of which found that it was not a predictor of drop-out (three studies) or non-response (three studies). However,

drop-out and non-response in all these studies were measured differently, which still makes it hard to draw firm conclusions about the predictive validity of these measures.

Clinical Implications

Targeting the identified predictors before or early in the treatment of anxiety and depressive disorders might help reduce treatment failure, which in turn might increase the cost-effectiveness of treatments. For example, in accordance with Constantino et al. (2018), assessing the outcome expectancy of patients at the start of treatment gives insight into the probability of a poor treatment outcome or even treatment failure. It offers an opportunity to assess whether interventions targeting the expectations of treatment should be integrated in the treatment in order to reduce the risk of treatment failure. Also, by assessing the use of social support early in treatment, interventions can be focused more on involving the social support system in therapy or enhancing coping strategies for asking for help, which might prevent treatment failure.

Unfortunately, as there are very few to no amenable factors that show strong evidence of being a predictor of treatment failure, clinicians need to be careful to use variables of which they assume they influence a poor therapy outcome. With the current knowledge, it is not possible to withhold patients from specific treatments or adapt treatments based on amenable predictors of treatment failure.

Research Implications and Suggestions

One important limitation is that for each predictor, only a limited number of studies have been carried out, often for only one specific treatment and/or disorder. For instance, CBT was the treatment of choice in five of the six studies into treatment expectancy as a predictor, whereas only two of them studied a sample of patients with the same disorder. As we cannot rule out the possibility that predictors of treatment failure differ for different disorders as well as for different treatments, predictors preferably should be studied in multiple treatment and patient groups to reach more definitive conclusions on overall predictors of treatment failure (Dalgleish et al., 2020). A further limitation is the absence of standardized measures for response rates (Loerinc et al., 2015), which makes it difficult to draw definitive conclusions about treatment outcome and -thus- about treatment failure. The reliable change index (RCI), as defined by Jacobson and Truax (1991), could address this problem (Loerinc et al., 2015), as by using this index treatment failure can be defined as meeting the criteria for no clinically significant change or deterioration at the end of treatment. The fact that the vast majority of studies reporting on predictors of treatment outcome only studied (or reported) predictors of treatment success or reported on poor(er) outcome is a third limitation. As we could not assume that poor(er) outcome is identical to treatment failure, or that results of treatment success can be inverted to treatment failure, most studies had to be excluded. Future

studies would benefit from also studying and reporting on predictors of treatment failure instead of treatment success only. Another limitation is that the terms 'association' and 'prediction' were used interchangeably in both the results and discussion sections of papers excluded from this review. Hence, it was not always clear if the studied variable was in fact a predictor, or if the predictors were just correlated with the outcome. We only included the studies that were very clear that they measured predictors of treatment failure. The field would benefit from a more consistent and precise use of these terms in presenting results on predictors of treatment outcome. A final limitation is that this review is not pre-registered and is not completely written according to current scoping review guidelines (Tricco et al., 2018). This stems from the fact that the first search of this review was conducted in 2017, before the publication of these guidelines. We recommend a future scoping or systematic review on the same topic to replicate and renew these findings.

Conclusion

By knowing which factors are predictors of treatment failure, we can move towards studying how to intervene on these factors in order to decrease the odds of treatment failure. Based on the results of this review, there are no predictors that stand out as overall strong amenable predictors of treatment failure. Future studies are needed to determine whether already found predictors are replicable with the same instruments in different patient samples and with different treatments.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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

Competing Interests: The authors report no conflict of interest.

Author Contributions: *Vivian Peerbooms:* Conceptualization, Methodology, Investigation, Data Curation, Writing-Original Draft. *Michael van den Boogaard:* Conceptualization, Methodology, Data Curation, Writing- Review and Editing. *Colin van der Heiden:* Conceptualization, Methodology, Data Curation, Writing- Review and Editing. *Matthias Wieser:* Writing-Review and Editing.

Reporting Guidelines: The Prisma ScR guidelines were followed in writing this scoping review (Tricco et al., 2018).

Preregistration: There is no preregistration of this manuscript.

Social Media Accounts: *Vivian Peerbooms:* LinkedIn

Data Availability: Data sharing is not applicable to this article as no new data were created or analysed in this study.

References

Arndt, A., Lutz, W., Rubel, J., Berger, T., Meyer, B., Schröder, J., Späth, C., Hautzinger, M., Fuhr, K., Rose, M., Hohagen, F., Klein, J. P., & Moritz, S. (2020). Identifying change-dropout patterns during an Internet-based intervention for depression by applying the Muthen-Roy model. *Cognitive Behaviour Therapy, 49*(1), 22–40. https://doi.org/10.1080/16506073.2018.1556331

Bélanger, C., Courchesne, C., Leduc, A. G., Dugal, C., El-Baalbaki, G., Marchand, A., Godbout, N., Marcaurelle, R., & Perreault, M. (2017). Predictors of dropout from cognitive-behavioral group treatment for panic disorder with agoraphobia: An exploratory study. *Behavior Modification, 41*(1), 113–140. https://doi.org/10.1177/0145445516656614

Blom, M. B. J., Spinhoven, P., Hoffman, T., Jonker, K., Hoencamp, E., Haffmans, P. M. J., & van Dyck, R. (2007). Severity and duration of depression, not personality factors, predict short term outcome in the treatment of major depression. *Journal of Affective Disorders, 104*(1-3), 119–126. https://doi.org/10.1016/j.jad.2007.03.010

Chambless, D. L., & Steketee, G. (1999). Expressed emotion and behavior therapy outcome: A prospective study with obsessive-compulsive and agoraphobic outpatients. *Journal of Consulting and Clinical Psychology, 67*(5), 658–665. https://doi.org/10.1037/0022-006X.67.5.658

Chambless, D. L., Tran, G. Q., & Glass, C. R. (1997). Predictors of response to cognitive-behavioral group therapy for social phobia. *Journal of Anxiety Disorders, 11*(3), 221–240. https://doi.org/10.1016/S0887-6185(97)00008-X

Constantino, M. J., Vîslă, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. *Psychotherapy, 55*(4), 473–485. https://doi.org/10.1037/pst0000169

Critchfield, K. L., Henry, W. P., Castonguay, L. G., & Borkovec, T. D. (2007). Interpersonal process and outcome in variants of cognitive-behavioral psychotherapy. *Journal of Clinical Psychology, 63*(1), 31–51. https://doi.org/10.1002/jclp.20329

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: A meta-analysis. *Journal of Affective Disorders, 159*, 118–126. https://doi.org/10.1016/j.jad.2014.02.026

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

Dandachi-FitzGerald, B., Otgaar, H., & Merckelbach, H. (2023). When psychotherapy fails. In C. L. Cobb, S. J. Lynn, & W. O'Donohue (Eds.), *Toward a science of clinical psychology: A tribute to the life and works of Scott O. Lilienfeld* (pp. 301-319). Springer International. https://doi.org/10.1007/978-3-031-14332-8_16

Edmonds, M., Hadjistavropoulos, H. D., Schneider, L. H., Dear, B. F., & Titov, N. (2018). Who benefits most from therapist-assisted internet-delivered cognitive behaviour therapy in clinical practice? Predictors of symptom change and dropout. *Journal of Anxiety Disorders, 54*, 24–32. https://doi.org/10.1016/j.janxdis.2018.01.003

Eilertsen, S. E. H., & Eilertsen, T. H. (2023). Why is it so hard to identify (consistent) predictors of treatment outcome in psychotherapy? – Clinical and research perspectives. *BMC Psychology, 11*(1), Article 198. https://doi.org/10.1186/s40359-023-01238-8

Grilo, C. M., Money, R., Barlow, D. H., Goddard, A. W., Gorman, J. M., Hofmann, S. G., Papp, L. A., Shear, M. K., & Woods, S. W. (1998). Pretreatment patient factors predicting attrition from a multicenter randomized controlled treatment study for panic disorder. *Comprehensive Psychiatry, 39*(6), 323–332. https://doi.org/10.1016/S0010-440X(98)90043-8

Hoyer, J., Wiltink, J., Hiller, W., Miller, R., Salzer, S., Sarnowsky, S., Stangier, U., Strauss, B., Willutzki, U., & Leibing, E. (2016). Baseline patient characteristics predicting outcome and attrition in cognitive therapy for social phobia: Results from a large multicentre trial. *Clinical Psychology & Psychotherapy, 23*(1), 35–46. https://doi.org/10.1002/cpp.1936

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. https://doi.org/10.1037/0022-006X.59.1.12

Johnson, S., Price, M., Mehta, N., & Anderson, P. L. (2014). Stereotype confirmation concerns predict dropout from cognitive behavioral therapy for social anxiety disorder. *BMC Psychiatry, 14*, Article 233. https://doi.org/10.1186/s12888-014-0233-8

Keefe, J. R., Chambless, D. L., Barber, J. P., & Milrod, B. L. (2021). Predictors and moderators of treatment dropout in cognitive-behavioral and psychodynamic therapies for panic disorder. *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 31*(4), 432–442. https://doi.org/10.1080/10503307.2020.1784487

Keijsers, G. P. J., Hoogduin, C. A. L., & Schaap, C. P. D. R. (1994). Predictors of treatment outcome in the behavioural treatment of obsessive-compulsive disorder. *British Journal of Psychiatry, 165*(6), 781–786. https://doi.org/10.1192/bjp.165.6.781

LeBeau, R. T., Davies, C. D., Culver, N. C., & Craske, M. G. (2013). Homework compliance counts in cognitive-behavioral therapy. *Cognitive Behaviour Therapy, 42*(3), 171–179. https://doi.org/10.1080/16506073.2013.763286

Lin, T., & Farber, B. A. (2021). Trajectories of depression in psychotherapy: How client characteristics predict clinical improvement. *Journal of Clinical Psychology, 77*(6), 1354–1370. https://doi.org/10.1002/jclp.23119

Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: Need for standardized criteria. *Clinical Psychology Review, 42*, 72–82. https://doi.org/10.1016/j.cpr.2015.08.004

Lutz, W., Rubel, J. A., Schwartz, B., Schilling, V., & Deisenhofer, A.-K. (2019). Towards integrating personalized feedback research into clinical practice: Development of the Trier Treatment Navigator (TTN). *Behaviour Research and Therapy, 120*, Article 103438. https://doi.org/10.1016/j.brat.2019.103438

Marker, I., Salvaris, C. A., Thompson, E. M., Tolliday, T., & Norton, P. J. (2019). Client motivation and engagement in transdiagnostic group cognitive behavioral therapy for anxiety disorders:

Predictors and outcomes. *Cognitive Therapy and Research, 43*(5), 819–833. https://doi.org/10.1007/s10608-019-10014-1

Marquett, R. M., Thompson, L. W., Reiser, R. P., Holland, J. M., O'Hara, R. M., Kesler, S. R., Stepanenko, A., Bilbrey, A., Rengifo, J., Majoros, A., & Thompson, D. G. (2013). Psychosocial predictors of treatment response to cognitive-behavior therapy for late-life depression: An exploratory study. *Aging & Mental Health, 17*(7), 830–838. https://doi.org/10.1080/13607863.2013.791661

McDevitt-Petrovic, O., Shevlin, M., & Kirby, K. (2020). Modelling changes in anxiety and depression during low-intensity cognitive behavioural therapy: An application of growth mixture models. *British Journal of Clinical Psychology, 59*(2), 169–185. https://doi.org/10.1111/bjc.12237

Miller, M. D., Schulz, R., Paradis, C., Houck, P. R., Mazumdar, S., Frank, E., Dew, M. A., & Reynolds, C. F., III. (1996). Changes in perceived health status of depressed elderly patients treated until remission. *American Journal of Psychiatry, 153*(10), 1350–1352. https://doi.org/10.1176/ajp.153.10.1350

Moggia, D., Lutz, W., Arndt, A., & Feixas, G. (2020). Patterns of change and their relationship to outcome and follow-up in group and individual psychotherapy for depression. *Journal of Consulting and Clinical Psychology, 88*(8), 757–773. https://doi.org/10.1037/ccp0000562

Moradveisi, L., Huibers, M., Renner, F., & Arntz, A. (2014). The influence of patients' preference/attitude towards psychotherapy and antidepressant medication on the treatment of major depressive disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 45*(1), 170–177. https://doi.org/10.1016/j.jbtep.2013.10.003

Oasi, O., & Werbart, A. (2020). Editorial: Unsuccessful psychotherapies: When and how do treatments fail? *Frontiers in Psychology, 11*, Article 578997. https://doi.org/10.3389/fpsyg.2020.578997

Parker, G., Brown, L., & Blignault, I. (1986). Coping behaviors as predictors of the course of clinical depression. *Archives of General Psychiatry, 43*(6), 561–565. https://doi.org/10.1001/archpsyc.1986.01800060055007

Renaud, J., Russell, J. J., & Myhr, G. (2014). Predicting who benefits most from cognitive-behavioral therapy for anxiety and depression. *Journal of Clinical Psychology, 70*(10), 924–932. https://doi.org/10.1002/jclp.22099

Safren, S. A., Heimberg, R. G., & Juster, H. R. (1997). Clients' expectancies and their relationship to pretreatment symptomatology and outcome of cognitive-behavioral group treatment for social phobia. *Journal of Consulting and Clinical Psychology, 65*(4), 694–698. https://doi.org/10.1037/0022-006X.65.4.694

Schilling, V. N. L. S., Zimmermann, D., Rubel, J. A., Boyle, K. S., & Lutz, W. (2021). Why do patients go off track? Examining potential influencing factors for being at risk of psychotherapy treatment failure. *Quality of Life Research, 30*(11), 3287–3298. https://doi.org/10.1007/s11136-020-02664-6

Schindler, A., Hiller, W., & Witthöft, M. (2013). What predicts outcome, response, and drop-out in CBT of depressive adults? A naturalistic study. *Behavioural and Cognitive Psychotherapy, 41*(3), 365–370. https://doi.org/10.1017/S1352465812001063

Solomonov, N., Lee, J., Banerjee, S., Fluckiger, C., Kanellopoulos, D., Gunning, F. M., Sirey, J. A., Liston, C., Raue, P. J., Hull, T. D., Arean, P. A., & Alexopoulos, G. S. (2021). Modifiable predictors of nonresponse to psychotherapies for late-life depression with executive dysfunction: A machine learning approach. *Molecular Psychiatry, 26*(9), 5190–5198. https://doi.org/10.1038/s41380-020-0836-z

Steketee, G., & Chambless, D. L. (1992). Methodological issues in prediction of treatment outcome. *Clinical Psychology Review, 12*(4), 387–400. https://doi.org/10.1016/0272-7358(92)90123-P

Steketee, G., Siev, J., Fama, J. M., Keshaviah, A., Chosak, A., & Wilhelm, S. (2011). Predictors of treatment outcome in modular cognitive therapy for obsessive-compulsive disorder. *Depression and Anxiety, 28*(4), 333–341. https://doi.org/10.1002/da.20785

Strauss, B., Koranyi, S., Altmann, U., Nolte, T., Beutel, M. E., Wiltink, J., Herpertz, S., Hiller, W., Hoyer, J., Joraschky, P., Nolting, B., Stangier, U., Willutzki, U., Salzer, S., Leibing, E., Leichsenring, F., & Kirchmann, H. (2017). Partner-related attachment as a moderator of outcome in patients with social anxiety disorder—A comparison between short-term cognitive-behavioral and psychodynamic therapy. *Psychotherapy, 54*(4), 339–350. https://doi.org/10.1037/pst0000129

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., & Weeks, L. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467–473. https://doi.org/10.7326/M18-0850

van den Boogaard, T. M. (2012). *The negotiated approach in the treatment of depressive disorders: The impact on patient-treatment compatibility and outcome* [PhD-Thesis – Research external, graduation internal, Vrije Universiteit Amsterdam]. VU Research Portal.

Vogel, P. A., Hansen, B., Stiles, T. C., & Gotestam, K. G. (2006). Treatment motivation, treatment expectancy, and helping alliance as predictors of outcome in cognitive behavioral treatment of OCD. *Journal of Behavior Therapy and Experimental Psychiatry, 37*(3), 247–255. https://doi.org/10.1016/j.jbtep.2005.12.001

Westra, H. A. (2011). Comparing the predictive capacity of observed in-session resistance to self-reported motivation in cognitive behavioral therapy. *Behaviour Research and Therapy, 49*(2), 106–113. https://doi.org/10.1016/j.brat.2010.11.007

World Health Organization. (2023a, March 31). *Depressive disorder (depression).* https://www.who.int/news-room/fact-sheets/detail/depression

World Health Organization. (2023b, September 27). *Anxiety disorders.* https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders

Zilcha-Mano, S. (2021). Toward personalized psychotherapy: The importance of the trait-like/state-like distinction for understanding therapeutic change. *The American Psychologist, 76*(3), 516–528. https://doi.org/10.1037/amp0000629

Zilcha-Mano, S., Keefe, J. R., Chui, H., Rubin, A., Barrett, M. S., & Barber, J. P. (2016). Reducing dropout in treatment for depression: Translating dropout predictors into individualized treatment recommendations. *The Journal of Clinical Psychiatry, 77*(12), e1584–e1590. https://doi.org/10.4088/JCP.15m10081

EACLIPT

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

zpid
Leibniz-Institut für Psychologie

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

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

Adaptive and Maladaptive Networks Using Ecological Momentary Assessment

Bettina Hufschmidt[1], Mareike Ebert[1], Arwin Nemani[1], Viktoria Kohl[1], Lucie Pahlen[1], Simon Müller[1], Stefan G. Hofmann[2], Ulrich Stangier[1]

[1] *Department of Clinical Psychology and Psychotherapy, Goethe University Frankfurt, Frankfurt am Main, Germany.*
[2] *Department of Psychology, Philipps-University of Marburg, Marburg, Germany.*

Clinical Psychology in Europe, 2026, Vol. 8(2), Article e17891, https://doi.org/10.32872/cpe.17891

Received: 2025-05-02 • **Accepted:** 2026-02-03 • **Published (VoR):** 2026-05-29

Handling Editor: Cornelia Weise, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Corresponding Author: Ulrich Stangier, Department of Clinical Psychology and Psychotherapy, Goethe University Frankfurt, Varrentrappstr. 40-42, 60486, Frankfurt am Main, Germany. Phone +49 (0)69 798 25373. E-mail: stangier@psych.uni-frankfurt.de

Supplementary Materials: Code, Materials [see Index of Supplementary Materials]

Abstract

Background: A person's ability to adapt to a given context is a critical determinant of mental health and psychopathology, which has been redefined by network approaches and Ecological Momentary Assessment (EMA). This study examined whether including adaptive processes in EMA enhances the informational value of idiographic network models.

Method: Forty-five university students participated in a multi-week EMA protocol assessing psychological dimensions using bipolar visual analogue scales ranging from maladaptive to adaptive. Participants were randomly assigned to two groups: one assessed only maladaptive processes; the other included both maladaptive and adaptive processes.

Results: Network analyses indicated higher density and connectivity in the adaptive–maladaptive group, as well as significantly reduced floor effects across all variables. Greater response dispersion was associated with more differentiated network structures. Motivation emerged as the most central node across conditions, highlighting its relevance as a transdiagnostic treatment target. Cognitive processing showed strong associations with other variables, underlining its clinical importance.

Conclusion: The findings suggest that incorporating adaptive dimensions into EMA facilitates a more comprehensive understanding of psychological functioning and improves the interpretability

of idiographic models. The study represents an initial feasibility investigation and a basis for further investigations in clinical practice.

Keywords

ecological momentary assessment, case conceptualization, maladaptive psychological processes, adaptive psychological processes, network approach

Highlights

- Adaptive processes enrich EMA data by increasing response variability and network connectivity.
- Adaptive processes enhance EMA data, reducing floor effects and revealing psychological dynamics.
- Assessing motivational processes appears beneficial for enhancing psychological functioning.
- Targeting cognitive processes in EMA seems effective in promoting adaptive functioning.

Case conceptualization plays a crucial role in treatment planning, since it offers a context-sensitive, yet concise model of the client's functioning and helps to identify relevant treatment targets along with appropriate assessment methods (Gilboa-Schechtman, 2024). Usually, retrospective reports are used as the primary source of information. However, they may be affected by a number of biases, e.g., overestimation of frequency and intensity of symptoms (Van den Bergh & Walentynowicz, 2016). Ecological Momentary Assessments (EMA) offer a potential solution to these limitations of retrospective self-reports. By collecting data in real time in people's everyday lives, EMA enables the capture of detailed information about the dynamics of psychological processes, including emotions, cognitions, and behaviors (Shiffman et al., 2008).

Recent advances in dynamic network analysis have provided a new perspective on using EMA data for treatment planning (Hayes et al., 2019). An increasing number of studies have provided evidence that dynamic network analyses may significantly enhance treatment planning by creating an idiographic dynamic model of the patient's problem based on EMA data (Burger et al., 2020; Frumkin et al., 2021; Roefs et al., 2022; Rubel et al., 2018).

However, it is not yet clear how to select variables to provide a valid idiographic conceptualization of the patient's problem. Many studies use symptom items based on diagnostic criteria or standardized psychometric instruments for EMA assessment (Cusack et al., 2024), thereby adopting a nomothetic approach, which may not be appropriate for the individual processes of the patient (Beltz et al., 2016).

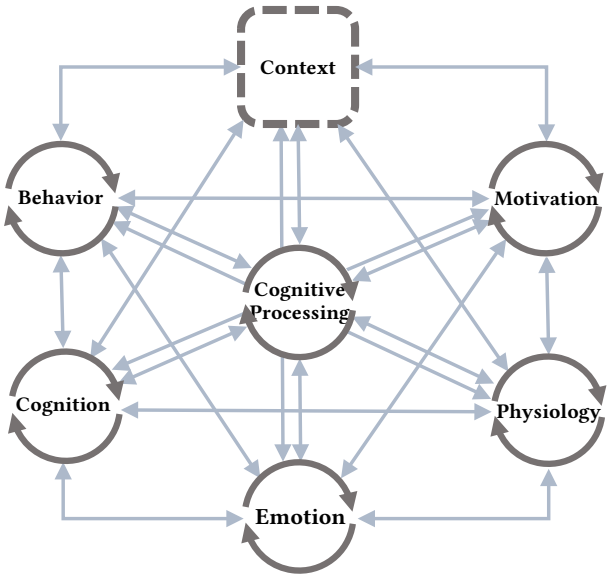
As an alternative, the dynamic network approach advocates an idiographic view of psychological processes as interacting nodes of an individual network (Borsboom, 2017).

While the dynamic network emphasizes causal relations and feedback loops within psychological systems, where certain nodes/symptoms activate others, the Extended Evolutionary Meta Model (Hayes et al., 2020) highlights the adaptiveness of human behavior and psychological functioning. An integration of both theories is provided by the Adaptive Process-Based Network Model (APNM), which postulates that adaptation results from the self-regulation and interaction of psychological processes conceptualized as nodes of an individual network. These processes can be described on the six core dimensions of cognition, emotion, physiology, behavior, cognitive processing, and motivation. The dimensions can be assessed in relation to a specific context using questions such as:

"How did I feel?" (e.g., sad vs. happy), "How was my body sensation?" (e.g., restless vs. calm), "How were my thoughts?" (e.g., „My life is hopeless." vs. „I can influence my life in a direction that I feel good about."), "How was my behavior?" (e.g., approaching vs. avoidance), "How was my inner cognitive processing?" (e.g., ruminating vs. attentive), and "What motivational goal did I have in the situation?" (e.g., avoidance of failure vs. striving for appreciation). The resulting psychological networks may be adaptive or maladaptive depending on the context (see Figure 1).

Figure 1

The Adaptive Process-Based Network Model (APNM)



In emotional disorders, maladaptation to the environment is also influenced by distorted cognitive processing. For instance, memory and attentional biases play a key role in the development and maintenance of emotional disorders (Cisler & Koster, 2010; Everaert et al., 2020; Phelps & Hofmann, 2019).

Although the reliability of measures used to assess cognitive processing in clinical practice remains questionable (Rodebaugh et al., 2016), treatment approaches targeting cognitive processes have been suggested to enhance therapeutic outcomes (e.g., Ehlers et al., 2012; Hvenegaard et al., 2020; Schreiber et al., 2015). In addition, research in neurocognitive science provides strong evidence for the substantial impact of motivational mechanisms on cognitive processing, emotions, and behavior (Braver et al., 2014). For instance, reward motivation has a robust influence on improving proactive cognitive control of goal-oriented behavior (Chiew & Braver, 2014). Furthermore, neurocognitive research has shown that motivational processes such as reduced reward responsiveness, via neural circuits, profoundly affect behavior and mental health problems like depression (Nestler, 2015).

Whereas psychological treatments have traditionally focused on maladaptive, psychopathological processes, defining "positive" and adaptive dimensions has become increasingly important in establishing treatment goals (Epton et al., 2017).

Thus, within a network approach to case conceptualization, it may be beneficial to specify not only maladaptive, but also adaptive poles of the dimensions relevant to the individual patient. Therefore, a dynamic network model may facilitate patient engagement in change processes, enhancing motivation to adopt the desired psychological processes and behaviors (Locke & Latham, 2002). So far, however, this approach has not yet been applied in the context of EMA. When applying continuous rating scales in EMA, however, it is important to consider potential psychometric issues such as floor effects.

In the evaluation of self-report measures, floor effects can limit sensitivity to detect variation at the lower end of the scale. A floor effect occurs when a substantial proportion of responses cluster at the lower bound, suggesting that the scale may not adequately capture lower levels of the construct of interest (de Vet et al., 2011; Terwee et al., 2007). This issue is particularly relevant in EMA contexts, where continuous scales (e.g., 0–100) are commonly used, and dynamic variation is critical for model interpretability. A floor effect is typically considered relevant when responses fall at or near the minimum score. For this study, we defined the floor range as the lowest 10% of the scale (i.e., values ≤ 10), allowing us to quantify and compare the prevalence of floor effects across experimental groups. This operationalization balances empirical convention with the specific characteristics of EMA data and continuous response formats.

In the present study, we investigated the feasibility of assessing maladaptive and adaptive psychological processes in an analogue sample of university students. We hypothesized that the inclusion of adaptive processes can be successfully implemented in EMA data used to calculate dynamic network models. Specifically, this approach is

expected to provide a more holistic view of an individual's experiences and associated processes compared to models based solely on maladaptive processes. A recent study by Nemani et al. (2025) provides evidence that the choice of bipolar versus unipolar scales in EMA significantly impacts the structural complexity of network models. The findings indicate that bipolar scaling captures a wider range of emotional and cognitive states, as participants are less likely to choose a neutral "0" response. This allows for a more detailed and accurate picture of psychological processes.

Previous research on emotion regulation and its impact on emotional experience suggests that the simultaneous assessment of both adaptive and maladaptive dimensions using EMA can reveal more nuanced patterns (De la Barrera et al., 2024; Short et al., 2018).

Building on these results, the present study examines how overall the variability in response distributions affects the density and connectivity of psychological networks. We hypothesize that network density will be greater when both maladaptive and adaptive processes are assessed, compared to when only maladaptive processes are considered.

Given that motivational schemas – as described above – control behavior and information processing, we hypothesize that motivational processes - ranging along a continuum from approach to avoidance – play a central role in understanding both maladaptive and adaptive psychological functioning and are characterized by a high degree of influence on other variables which is manifested by strong outgoing edges to other nodes in a dynamic network model. Furthermore, we propose that cognitive processing, due to its well-documented associations with a variety of psychological variables, represents another crucial node in dynamic networks of interconnected psychological processes, as indicated by strong associations with the other variables.

Method

Participants

Of the 50 individuals initially recruited, five participants discontinued the study before or during the EMA phase. The dropout reasons were not systematically recorded. The final sample consisted of 45 students from German universities who volunteered to participate in exchange for a small monetary compensation and the opportunity to gain experience with smartphone-assisted self-monitoring. Recruitment occurred via social media and digital flyers. By providing informed consent, participants confirmed that they were currently enrolled as university students and had access to a smartphone suitable for completing daily assessments. Participants ranged in age from 21 to 40 years (M = 25.19, SD = 4.65). The majority of the sample identified as female (73.6%), and over half (55.3%) were enrolled in psychology programs.

Procedures

Participants received an email with a QR code to set up the EMA application and a questionnaire to identify dimensions of a personally relevant problem experienced in daily life. Participants collected data between October 2023 and January 2024 using the Vacay Status-PBT app [Mobile app] for iOS and Android (Vacay GmbH, 2020).

Participants were instructed to complete three daily assessments, each lasting about two minutes. They received at fixed times three prompts per day – 6 am, 12 pm, 6 pm. Each survey consisted of seven questions, with each question relating to one of the seven domains: emotion, cognition, situation, cognitive processing, motivation, bodily response, and behavior.

Participants were randomly assigned to one of two conditions: one group (n = 23) specified maladaptive aspects and processes related to their identified problem, while the other group (n = 22) also specified corresponding adaptive, desirable processes. For the group that assessed maladaptive and adaptive dimensions, each question pertained solely to the specific process under investigation. For example: "How did I feel?" or "How was my body sensation?". The items were arranged with two opposing poles placed at either end and a movable slider in between. The left pole represented the maladaptive expression of the process, while the right pole represented its adaptive form. As the slider was adjusted, numerical values appeared: moving it toward the maladaptive pole produced values from 0 to −100, while moving it toward the adaptive pole produced values from 0 to +100. Figure 2 shows how this was displayed in the app.

For the group that assessed maladaptive dimensions only, questions contained the participant's self-defined expression of the respective process. For example, if the individually defined feeling was "*sad*," the corresponding question was: "*How sad did I feel?*" (Figure 3).

Participants were given substantial flexibility in the selection of anchors. For clarification and support in finding suitable descriptions for individual processes, they were given a questionnaire with examples of each domain available in Supplementary Material A.

Based on questionnaire responses, the research team programmed individualized EMA items in the app. Participants then began the EMA phase and were informed that they could request modifications to any item they felt did not accurately reflect their experience. This implied rewording items but also replacing predefined items with more appropriate ones. However, the project team took care that semantic-logical opposites of the condition-specific aspects were chosen, while also accounting for the subjectively perceived significance in the linguistic formulation. In cases where participants experienced difficulties, the issues were addressed through telephone consultations. The daily EMA protocol began with a situational anchoring item: "*How similar was the current, objective situation to my problem situation?*" followed by brief prompts assessing the seven identified dimensions (see Appendix A1 for the maladaptive group and Appendix

A2 for the maladaptive-adaptive group), each rated using a visual analog scale. The length of the EMA phase was approximately five weeks on average but varied between participants and depended on the number of missings. In the maladaptive group, they participated on average for 42 days (*SD* = 8.74) and missed 18% of the beeps. In the maladaptive-adaptive group, they took part on average for 41 days (*SD* = 5.19) and missed 17% of the beeps. During the EMA phase, participants completed around 100 assessments (on average, 94 assessments in the maladaptive group and, depending on the variable 98/99 in the maladaptive-adaptive group) – aligning with prior research recommendations for idiographic modeling (Epskamp et al., 2018). During the EMA phase, the study team monitored participant entries. However, as monitoring was not performed daily, minor variations in the total number of measurement points occurred. Participants did not receive feedback and had no access to self-monitoring features.

The final evaluation yielded a total of 4363 EMA observations across all participants.

Figure 2

Example of an Individualized Item Assessing the Adaptive and Maladaptive Dimension of Emotion

Figure 3

Example of the Individualized Item Assessing Only the Maladaptive Dimension of Emotion

Data Analysis

In order to evaluate the added value of including adaptive processes on a temporal group level, multilevel vector autoregressive (mlVAR) models were calculated. For the purpose of this study, we focused on the fixed-effect temporal networks (using a lag-1 approach) to model the average temporal relationships within individuals. The associations between variables ("nodes") will be referred to as "edges", with a blue edge indicating a positive association and a red edge a negative association. As a measure of relative importance, outStrength refers to the sum of the weights of all outgoing edges from a specific node. InStrength refers to the sum of the weights of all incoming edges to a specific node, representing how strongly that node is influenced by other nodes.

The main data analysis was conducted using the mlVAR package (Version 0.5.2; Epskamp et al., 2024) within RStudio (Version 2024.04.2, Posit Team). Prior to analysis, important preprocessing steps included the removal of the linear time trends through detrending and the handling of missing data using the Kalman filter (function na_kalman) from the imputeTS package (Version 3.3; Moritz & Bartz-Beielstein, 2017). For each participant and each variable, missing data were below 40%.

Results

Descriptive Statistics

Descriptive statistics for both groups are summarized in Table 1. The maladaptive group showed mean scores ranging from 38.64 (Behaviour) to 49.17 (Motivation), with standard deviations around 29 to 32, indicating moderate within-group variability. The maladaptive-adaptive group exhibited scores centered closer to zero across most domains, with notably higher variability (*SD* mostly > 50), reflecting a more diverse range of responses. The broader scale range (-100 to +100) for the maladaptive-adaptive group compared to the restricted positive range (zero to 100) for the maladaptive group further highlights the distinct response patterns between groups. Intraclass correlation coefficients (ICCs) calculated quantified the proportion of variance attributable to stable between-person differences across psychological domains (Siepe et al., 2025). The maladaptive group showed a mean ICC of .389 (*SD* = 0.05), with values ranging from .319 (Behaviour) to 0.465 (Body Response). The maladaptive-adaptive group demonstrated a slightly lower mean ICC of .340 (*SD* = 0.05), with the highest stability observed for Situation (.445) and lowest for Cognitive Processing (.289). These moderate ICC values indicate that approximately 34 to 39% of total variance stems from consistent individual differences, while the majority (666%) reflects within-person fluctuation.

Table 1

Summary of Descriptive Statistics for Items by Group

Group	Item	*M*	*SD*	*Mdn*	Min	Max	*N*
1	Behaviour	38.64	31.34	31.82	0	100	2069
1	Body Response	45.65	31.80	45.04	0	100	2068
1	Cognition	40.57	30.01	35.00	0	100	2069
1	Emotion	40.77	29.26	35.00	0	100	2068
1	Motivation	49.17	32.45	57.00	0	100	2068
1	Cognitive Processing	45.98	30.40	45.00	0	100	2069
1	Situation	49.13	30.78	57.00	0	100	2071
2	Behaviour	-0.36	53.00	0.00	-100	100	2287
2	Body Response	0.35	55.80	-6.20	-100	100	2286
2	Cognition	-5.45	54.27	-15.45	-100	100	2288
2	Emotion	-1.34	51.95	-9.70	-100	100	2288
2	Motivation	-1.02	54.92	-6.60	-100	100	2283
2	Cognitive Processing	0.54	56.14	0.00	-100	100	2287
2	Situation	45.04	36.99	50.00	-100	100	2246

Note. Group 1 = Maladaptive (n = 23) with raw scores (zero to 100); Group 2 = Maladaptive Adaptive (n = 22) with standardized scores (-100 to 100). *SD* = Standard Deviation; *N* = total number of participants. Chi-square tests were used to compare groups.

To examine potential multicollinearity and node redundancy, zero-order correlation matrices among the seven key domains were calculated for each group. In the maladaptive group, correlations ranged from r = .33 to r = .74, indicating moderate associations but also sufficient distinction among nodes. For the maladaptive-adaptive group, correlations spanned from approximately r = -.54 to r = .81.

Network Analysis of EMA Variables

The data was analyzed separately for each group: a group of participants which exclusively recorded maladaptive processes and the other group that assessed maladaptive and adaptive processes. Between-person networks, reflecting trait-like associations among participant averages, are presented in Supplementary Material B in Figure 1 and Figure 2.

Fixed Effects Temporal Networks (Maladaptive/Maladaptive-Adaptive)

Figures 4 and 5 show the results of the two fixed effects temporal networks – Figure 4 for the group with exclusively maladaptive networks and Figure 5 for the group with maladaptive and adaptive processes.

Density

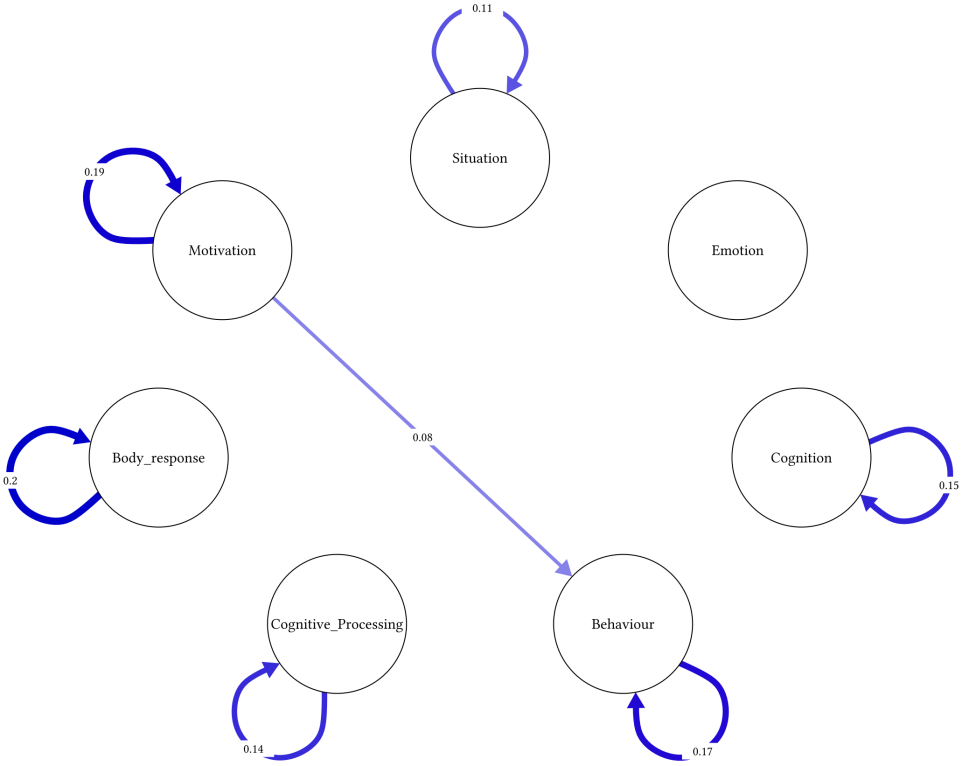
In the fixed-effect temporal maladaptive network, 14% of all possible edges were present, compared to 27% in the adaptive-maladaptive network. While some nodes in the average maladaptive network remained isolated, all nodes in the adaptive network were interconnected.

In Figure 4, which depicts the average temporal network of the maladaptive group, the *Emotion* node exhibits no temporal associations with any other variables. Most of the remaining nodes display only auto-temporal associations. Notably, the only cross-variable temporal link is a positive association from *Motivation* to *Behavior*, suggesting that, on average, motivation temporally predicts subsequent behavior across measurement points.

In Figure 5, which presents the average temporal network of the maladaptive-adaptive group, two interconnected subsystems emerge. In the first, *Cognition* is positively associated with *Emotion*. In the second, *Motivation* appears to initiate adaptive cascades, showing positive temporal associations with both *Behavior* and *Cognitive Processing*. *Cognitive Processing*, in turn, is negatively linked to *Situation* and positively associated with *Behavior*, which then shows a positive temporal connection to *Body Response*. Additionally, all nodes in the network display autoregressive loops. These reflect that each construct's state at one time point significantly predicts its own state at the next, demonstrating temporal consistency within that psychological dimension. Notably, in an idiographic network model with bipolar scaling, the network structure alone does not reveal which variable manifestations influence each other over time. To determine this, the individual values must be examined.

Figure 4

Fixed Effects Temporal Network, Maladaptive Processes

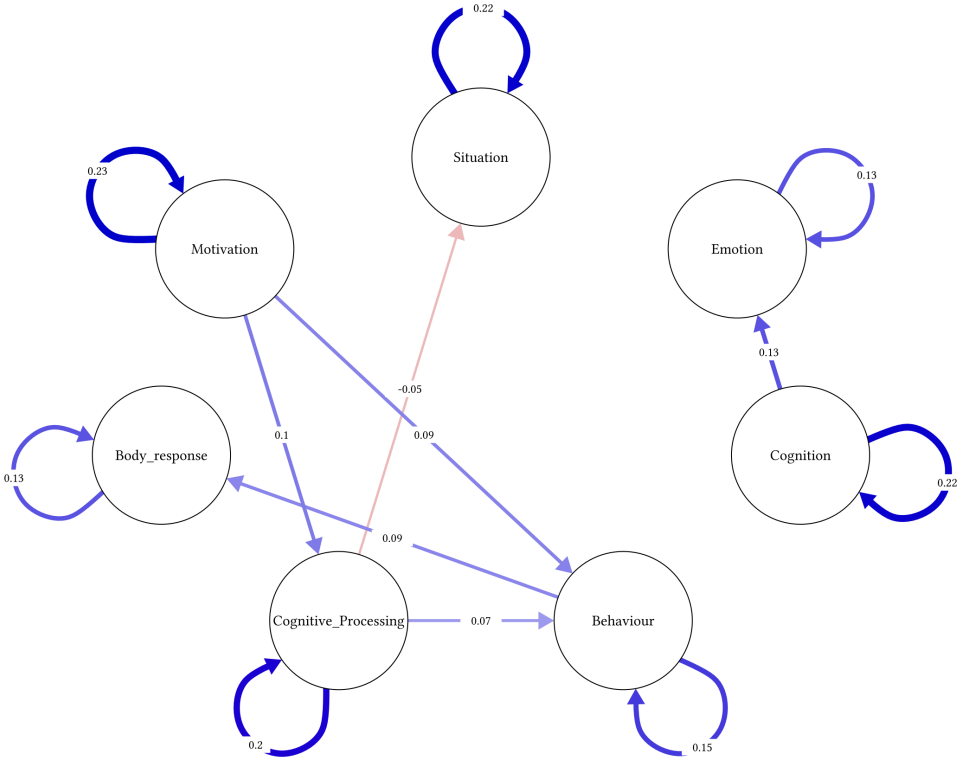


Note. Thickness of the autoregressive loops indicates self-influence strength. Thickness of the arrow indicates association strength. Arrowheads indicate the direction of the association between nodes (outStrength or inStrength). Arrow colour indicates whether the association is positive (blue) or negative (red). Higher colour saturation indicates a stronger association.

The estimation of the fixed effects temporal networks showed *Motivation* had the highest outStrength in both groups, which agrees with our hypothesis that motivation strongly impacts other psychological processes and should be targeted. The highest inStrength is in *Behavior* (see Table 2).

Figure 5

Fixed Effects Temporal Network, Maladaptive – Adaptive Processes



Note. Thickness of the autoregressive loops indicates self-influence strength. Thickness of the arrow indicates association strength. Arrowheads indicate the direction of the association between nodes (outStrength or inStrength). Arrow colour indicates whether the association is positive (blue) or negative (red). Higher colour saturation indicates a stronger association.

Response Distribution and Its Impact on the Network Structure

To ensure comparability across groups and to examine potential differences in floor effects, all bipolar scales ranging from −100 to +100 (used in the maladaptive-adaptive group) were linearly rescaled to match the unipolar 0 to 100 formats employed in the maladaptive only group. The transformation was linear, mapping −100 to zero, zero to 50, and +100 to 100. Thus, both groups were expressed on the same 0 to 100 metrics, despite the original difference in raw scale range. This transformation allowed for a uniform definition of the lower 10% of the scale (≤ 10) as the operational threshold for floor effects across all variables (Figure 6). We then analyzed response distributions across all scales to assess whether more evenly distributed responses across the scale increased network

density. Results revealed that participants in the adaptive-maladaptive group exhibited a significantly more balanced response distribution compared to the maladaptive-only group ($p < .05$) suggesting that the inclusion of adaptive processes encourages a broader range of responses across the scale, rather than clustering at extreme values.

Table 2

OutStrength and inStrength of the Fixed Effects Temporal Network for Both Groups (Maladaptive; Maladaptive-Adaptive)

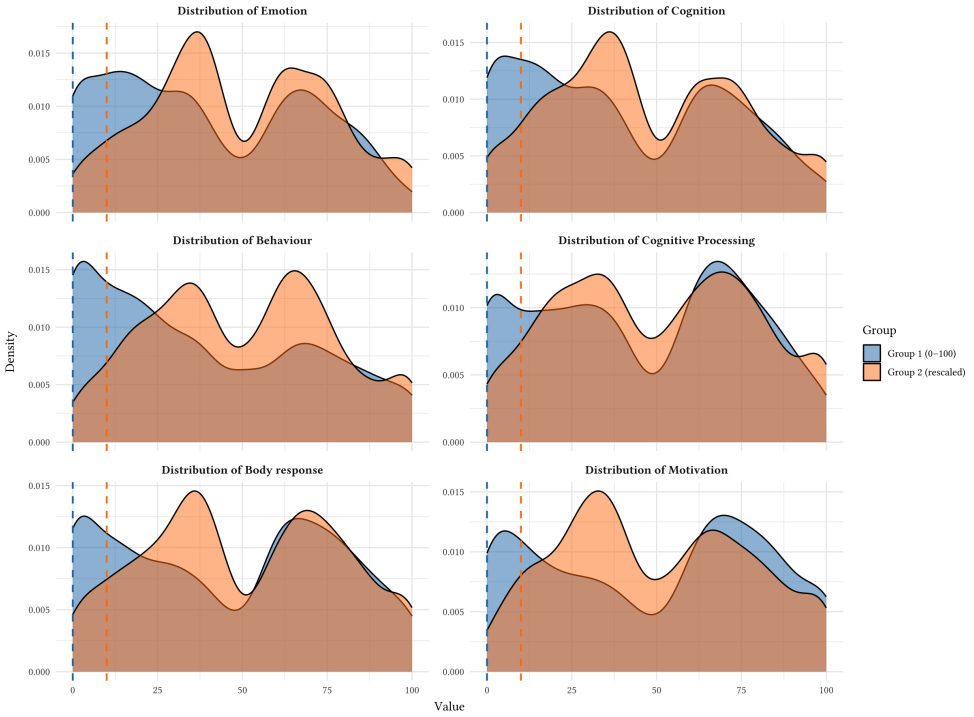
	Maladaptive (n = 22)		Maladaptive – Adaptive (n = 23)	
Variable	**outStrength total**	**inStrength total**	**outStrength total**	**inStrength total**
Situation	0.110	0.110	0.222	0.273
Emotion	0.000	0.000	0.134	0.267
Cognition	0.151	0.151	0.351	0.218
Behaviour	0.165	0.247	0.245	0.321
Cognitive Processing	0.145	0.145	0.323	0.297
Body Response	0.200	0.200	0.130	0.222
Motivation	0.266	0.185	0.423	0.230

Note. OutStrength = sum of the weights of all outgoing edges from a node; InStrength = sum of the weights of all incoming edges to a node. Higher values indicate stronger connections.

Individuals with greater response dispersion exhibited higher network connectivity, indicating that broader variability enables more nuanced psychological interconnections. By contrast, those with narrow response distributions (i.e., excessive clustering in specific scale regions) demonstrated lower network density, reinforcing the notion that response flexibility enhances the structural integrity of network models. Adaptive–maladaptive networks consistently exhibited significantly fewer floor responses across all variables compared to maladaptive-only networks (Table 3). This indicates enhanced sensitivity and greater response variability when adaptive processes are included in EMA assessments. Chi-square tests revealed highly significant differences between groups (all $p < .001$).

Figure 6

Distribution of Data in the Two Groups Maladaptive and Adaptive-Maladaptive



Note. Group 1 = maladaptive-only group; Group 2 = maladaptive–adaptive group. Values in Group 2 were linearly transformed from −100–0, 0–50, and 100–100 to a unipolar 0–100 scale to match the maladaptive-only group. The dotted lines indicate the floor range (lowest 10% of the scale).

Table 3

Comparison of Floor Effects of the Maladaptive (n = 2,070) and Maladaptive-Adaptive (n = 2,287) Group

	Maladaptive		Maladaptive-Adaptive			
Variable	**Floor count**	**Floor percent**	**Floor count**	**Floor percent**	χ^2	*p*
Behavior	513	24.8%	145	6.34%	287.05	< .001
Body Response	407	19.7%	186	8.14%	122.02	< .001
Cognition	435	21.0%	179	7.82%	155.31	< .001
Emotion	402	19.4%	151	6.60%	160.40	< .001
Motivation	379	18.3%	155	6.79%	133.09	< .001
Cognitive Processing	348	16.8%	175	7.65%	85.55	< .001

Note. Chi-square tests were used to compare groups.

Discussion

Building on previous studies on EMA (Mink et al., 2025; Nemani et al., 2025; Yim et al., 2020), the present study demonstrates the feasibility of incorporating both maladaptive and adaptive processes into the collection of EMA data for the calculation of idiographic dynamic network models. The present study advances existing research on EMA as a methodological approach to data collection and provides a foundation for subsequent investigations in clinical populations. Future work could specifically evaluate its applicability and utility for individualized case conceptualization as the implementation in psychotherapeutic settings is theoretically useful but still lacking (Mink et al., 2025).

The findings provide evidence that including adaptive processes enhances the richness of EMA data, used in network analysis. For example, when assessing emotion exclusively on maladaptive dimensions, emotional states lacked temporal stability and associations with other psychological processes, indicating that maladaptive emotion alone may not sufficiently capture the dynamic interplay within individuals. Furthermore, increased richness of EMA data is reflected by greater response dispersion and increased connectivity between variables in the dynamic network model. This in turn contributes to a deeper understanding of interrelationships among psychological processes and supports our hypothesis that density is stronger when both maladaptive and adaptive processes are assed.

It should be noted that in the present study, network density is operationalized as the average absolute strength of temporal (lagged) connections and thus primarily reflects the magnitude of direct interdependencies rather than network topology or functional optimality. When networks predominantly capture maladaptive processes, higher density may indicate increased vulnerability, as strongly coupled systems can be more sensitive to external perturbations and show reduced flexibility (Borsboom, 2017). However, in the present context, increased density is primarily interpreted as reflecting enhanced response variability and reduced floor effects due to the inclusion of adaptive dimensions, which facilitates the detection and interpretation of temporal associations. Because density is based on absolute edge weights, it does not distinguish between stabilizing and destabilizing dynamics, underscoring the need for cautious interpretation (Fried & Cramer, 2017).

By incorporating adaptive processes, the narrow definition of well-being as merely the absence of suffering is challenged, allowing for a more holistic understanding of an individual's experiences (Joseph & Wood, 2010). This aligns with the growing acceptance of a dimensional view on psychological processes and the idea that the severity of psychopathology varies along a continuous spectrum (Clark et al., 2017; van Os et al., 2009). Practically, adding adaptive processes in EMA, broadens the perspective on processes relevant to everyday life. In our study, motivation emerged as the central node in both groups, supporting our hypothesis of its high impact on other variables in a dynamic network model.

Thus, assessing approach motivation seems to be particularly important, as it may enhance intrinsic motivation, which mediates links between clinical symptoms and social functioning (Yamada et al., 2010) and drives behavior aimed at fulfilling psychological needs (Ryan & Deci, 2000).

Memory, attention, and consciousness have been proposed as crucial components of cognitive processing that significantly influence psychological functioning and mental health (Becker et al., 2021; Knowles et al., 2016; Phelps & Hofmann, 2019). Several dimensions of cognitive processing, such as concrete versus abstract thinking modes and rumination (Watkins & Moulds, 2005), as well as emotion regulation strategies like attentional deployment (McRae & Gross, 2020), have been linked to the development of emotional disorders. In contrast, cognitive processing styles such as decentering and mindful attention (Bernstein et al., 2015) have been proposed as adaptive responses to distress and are considered key targets in various psychological treatment approaches, including the Third Wave of CBT. Given that cognitive processes are fundamental to the mechanisms maintaining psychological problems and symptoms, it has been proposed that personalized clinical case formulation should address not only dysfunctional cognitive content (e.g., beliefs and automatic thoughts), but also dysfunctional cognitive processes (e.g. attentional biases; Dennis-Tiwary et al., 2019). However, due to their implicit nature, the identification of memory and attentional processes often requires substantial support, such as guidance from a therapist and intensive self-monitoring to recognize automatic patterns (Stefana et al., 2024). Therefore, the results of the present study support efforts to incorporate cognitive processing as a potential target for enhancing adaptive functioning within ecological momentary assessment frameworks by showing that cognitive processing has connections to other variables in a dynamic network model. In the present study, the connections are not particularly strong compared to others. Further research is needed to explore potential underlying mechanisms. Incorporating measures such as reflective functioning (Fonagy et al., 2016) could reveal whether interindividual differences in data quality are moderated by the capacity to reflect on internal states (Fonagy et al., 2016). In general, we recommend considering motivational processes as well as cognitive processing when designing future research.

For decades, different therapy schools used different approaches to case conceptualization. While in psychodynamic therapy the assessment of central psychodynamic constructs is crucial (Luborsky et al., 1994), the SORCK model (Kanfer & Saslow, 1965) is an established model for case conceptualization in behavior therapy. The transdiagnostic approach is gaining increasing popularity, because it pushes traditional diagnostic boundaries and offers a deeper understanding of the structure of psychological problems (Dalgleish et al., 2020). The findings of our study, which support a more in-depth exploration using bipolar scales, align with the broader shift to a more comprehensive understanding of interacting psychological processes. This approach can serve as a

fundament for case conceptualization in the context of personalized psychotherapy, an area of growing importance (Hayes et al., 2020; Huibers et al., 2021).

Nemani et al. (2025) demonstrated that bipolar scaling reduces zero inflation, our findings suggest that overall response variability across the entire scale plays a critical role in determining network connectivity. Specifically, the concentration of responses in extreme boundary regions (e.g., zero to 10, 90 to 100, -10 to 0, -100 to -90) may constrain network formation, limiting meaningful interconnections between psychological variables.

Our findings extend previous research by demonstrating that the additional assessment of adaptive processes can reduce floor effects which are associated with restrictions in the possibility to capture fluctuations in symptom ratings (Mestdagh et al., 2018) and limited variance (Šimkovic & Träuble, 2019). Since values at the lower bound cannot decrease further (e.g., Fries et al., 2014; Šimkovic & Träuble, 2019), floor effects constrain the ability to detect longitudinal changes in the lower range of the scale. By incorporating adaptive processes into the scale, overall variability is increased, thereby reducing constraints on the detection of interconnections among psychological variables. Our findings suggest examining both adaptive and maladaptive dimensions to gain a comprehensive understanding of the interplay among psychological processes.

This study has several limitations that should be acknowledged. First of all, temporal associations and node centrality metrics derived from network analyses represent statistical dependencies rather than causal effects. Therefore, conclusions about mechanisms should be made with caution. Network analysis conceptualizes psychological phenomena as systems of interacting symptoms or variables rather than manifestations of a single latent construct (Bastiaansen et al., 2020). One the one hand, this approach allows researchers to explore direct relationships among variables, identify central nodes, and detect possible pathways of change, offering insights for both theory and intervention (Bringmann et al., 2019). On the other hand, the method faces important limitations, particularly regarding the stability and replicability of estimated networks, which may vary with sample size, measurement error, and analytic choices (Hevey, 2018). As the robustness of the estimated networks was not systematically evaluated, the interpretation of the findings should pay attention to this fact.

In addition, adaptive and maladaptive processes were assessed in relation to the specific context in which the problem typically arises. This approach aligns with research showing coping effectiveness depends strongly on context (Cheng, 2001; Cole et al., 1994; Gross, 1998, Kalokerinos et al., 2017). However, this context-specific focus also limits the generalizability of our findings. Bonanno (2021) argues that psychological flexibility – the ability to select the most appropriate coping strategy based on situational demands – is a key factor in successfully coping with stressful events. This perspective may be especially relevant for EMA, where focusing on flexibility over rigid maladaptive–adaptive distinctions could offer more nuanced insights.

Furthermore, the sample primarily consisted of female university students, resulting in both a gender imbalance and limited demographic diversity.

Additionally, the relatively small sample size ($N = 45$), although comparable to feasibility samples in EMA research, may limit statistical power to detect subtle group differences in multilevel vector autoregressive modeling. Future research should aim to replicate these findings in larger samples, with a more balanced gender distribution and diverse educational backgrounds.

Another limitation concerns the method of participant recruitment: individuals volunteered to participate and were given the opportunity to engage in smartphone-based self observation which may have introduced a bias, as participants were likely already interested in using ecological momentary assessment (EMA) tools which may not represent clinical populations.

Furthermore, an important limitation of the study is the reliance on self-report measures, which may be influenced by social desirability or lack of awareness.

Additionally, given the relatively young and tech-savvy nature of the sample, the ease of smartphone use observed in this study may not be representative of other populations, potentially limiting the applicability of these findings to broader groups. A further limitation is the lack of procedures to identify or address careless responding in EMA, which may add noise and reduce the accuracy and interpretability of the dynamic network models.

Another limitation is the uncertainty associated with centrality estimates derived from mlVAR models. The complexity of multilevel temporal modeling combined with our relatively small sample size and the variability inherent in EMA data contribute to potential imprecision. As a result, while motivation appeared as a central node, the robustness of this finding remains uncertain. Future research employing bootstrap-based stability analyses is needed to strengthen conclusions about key intervention targets.

Moreover, as our sample consisted of non-clinical university students, reduced floor effects may partly reflect lower baseline symptom severity and greater variability in adaptive functioning in this population. Replication in clinical samples is needed to determine whether including adaptive processes in EMA yields the same benefits under higher symptom burden and lower functioning.

Finally, the use of different scale formats: zero to 100 in the maladaptive group and −100 to +100 in the maladaptive–adaptive group is another limitation. Although responses were linearly rescaled for comparability, the interpretation of the anchors and the understanding of the midpoint of the scale, which may not consistently reflect a true "neutral" state, may vary. This interpretative variability could introduce noise or bias in the data, potentially affecting the observed differences between the two groups. Future studies should standardize formats to rule out this source of bias.

Funding: This work was supported by the Hessische Ministerium für Wissenschaft und Kunst (as part of the LOEWE Spitzenprofessur), and 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 authors have no additional (i.e., non-financial) support to report.

Competing Interests: The authors declare that there is no conflict of interest regarding the publication of this article. None of the authors have any financial, personal, or professional relationships that could be perceived as influencing the research presented in this study. Dr. Hofmann receives financial support from the Alexander von Humboldt Foundation (as part of the Alexander von Humboldt Professur), and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 521379614 – TRR 393SFB/Transregio 393.

Ethics Statement: Informed consent was obtained from all participants included in the study. All procedures in studies involving human participants were performed in accordance with the ethical standards of the institution's Human Research Ethics Committee of the Faculty of Psychology and Sports Science at the Johann-Wolfgang-Goethe University Frankfurt.

Preregistration: Due to the small size the study was not submitted to a preregistration platform.

Reporting Guidelines: The study adheres to the JARS-Quant Guidelines of the American Psychological Association.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process: During the preparation of this work, the authors used ChatGPT-4o and DeepL Write in order to improve the style and grammar of the final manuscript. After using this tool/service, the authors reviewed and edited the content as needed and took full responsibility for the content of the publication.

Social Media Accounts: *Bettina Hufschmidt:* LinkedIn

Data Availability: The data underlying the results presented in this study are available from the Department of Clinical Psychology and Psychotherapy at Goethe University Frankfurt (contact via stangier@psych.uni-frankfurt.de). The data cannot be shared publicly due to ethical restrictions and privacy concerns regarding participant confidentiality. Access to these data may be granted by the Goethe University Institutional Data Access/Ethics Committee for researchers who meet the criteria for access to confidential data.
The R code used for the main analyses is available in OSF at https://osf.io/7amzu

Supplementary Materials

The Supplementary Materials contain the following items:

- **R code** used for the main analyses (Hufschmidt, 2026S)
- **Appendix** (Hufschmidt et al., 2026S): The appendix associated with this article provides the EMA-items used for assessing the maladaptive dimensions and the maladaptive and adaptive dimensions of individual psychological processes.
- **Additional information, tables, and figures** (Hufschmidt et al., 2026S): The supplementary material contains an additional information sheet provided to participants to help them define their own psychological processes (*Supplementary Material A*) and a comprehensive set of tables and figures that exceeded the scope of the main article (*Supplementary Material B*).

Index of Supplementary Materials

Hufschmidt, B. (2026S). *Adaptive and maladaptive networks using ecological momentary assessment* [R code]. OSF. https://osf.io/7amzu

Hufschmidt, B., Ebert, M., Nemani, A., Kohl, V., Pahlen, L., Müller, S., Hofmann, S. G., & Stangier, U. (2026S). *Supplementary materials to "Adaptive and maladaptive networks using ecological momentary assessment"* [Appendix; Additional information, tables, and figures]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.21827

References

Bastiaansen, J. A., Kunkels, Y. K., Blaauw, F. J., Boker, S. M., Ceulemans, E., Chen, M., Chow, S.-M., de Jonge, P., Emerencia, A. C., Epskamp, S., Fisher, A. J., Hamaker, E. L., Kuppens, P., Lutz, W., Meyer, M. J., Moulder, R., Oravecz, Z., Riese, H., Rubel, J., . . . Bringmann, L. F. (2020). Time to get personal? The impact of researchers choices on the selection of treatment targets using the experience sampling methodology. *Journal of Psychosomatic Research, 137*, Article 110211. https://doi.org/10.1016/j.jpsychores.2020.110211

Becker, E. S., Vanderhasselt, M.-A., Koster, E. H. W., & Vrijsen, J. N. (2021). Memories warm you up from the inside. But they also tear you apart: Editorial for the special issue on memory training for emotional disorders. *Cognitive Therapy and Research, 45*(5), 841–847. https://doi.org/10.1007/s10608-021-10249-x

Beltz, A. M., Wright, A. G., Sprague, B. N., & Molenaar, P. C. (2016). Bridging the nomothetic and idiographic approaches to the analysis of clinical data. *Assessment, 23*(4), 447–458. https://doi.org/10.1177/1073191116648209

Bernstein, R. E., Delker, B. C., Knight, J. A., & Freyd, J. J. (2015). Hypervigilance in college students: Associations with betrayal and dissociation and psychometric properties in a Brief Hypervigilance Scale. *Psychological Trauma: Theory, Research, Practice and Policy, 7*(5), 448–455. https://doi.org/10.1037/tra0000070

Bonanno, G. A. (2021). The resilience paradox. *European Journal of Psychotraumatology, 12*(1), Article 1942642. https://doi.org/10.1080/20008198.2021.1942642

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry, 16*(1), 5–13. https://doi.org/10.1002/wps.20375

Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Westbrook, J. A., Clement, N. J., Adcock, R. A., Barch, D. M., Botvinick, M. M., Carver, C. S., Cools, R., Custers, R., Dickinson, A., Dweck, C. S., Fishbach, A., Gollwitzer, P. M., Hess, T. M., Isaacowitz, D. M., Mather, M., . . . Somerville, L. H. (2014). Mechanisms of motivation-cognition interaction: Challenges and opportunities. *Cognitive, Affective & Behavioral Neuroscience, 14*(2), 443–472. https://doi.org/10.3758/s13415-014-0300-0

Bringmann, L. F., Elmer, T., Epskamp, S., Krause, R. W., Schoch, D., Wichers, M., Wigman, J. T. W., & Snippe, E. (2019). What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology, 128*(8), 892–903. https://doi.org/10.1037/abn0000446

Burger, J., van der Veen, D. C., Robinaugh, D. J., Quax, R., Riese, H., Schoevers, R. A., & Epskamp, S. (2020). Bridging the gap between complexity science and clinical practice by formalizing idiographic theories: A computational model of functional analysis. *BMC Medicine, 18*(1), Article 99. https://doi.org/10.1186/s12916-020-01558-1

Cheng, C. (2001). Assessing coping flexibility in real-life and laboratory settings: A multimethod approach. *Journal of Personality and Social Psychology, 80*(5), 814–833. https://doi.org/10.1037/0022-3514.80.5.814

Chiew, K. S., & Braver, T. S. (2014). Dissociable influences of reward motivation and positive emotion on cognitive control. *Cognitive, Affective & Behavioral Neuroscience, 14*(2), 509–529. https://doi.org/10.3758/s13415-014-0280-0

Cisler, J. M., & Koster, E. H. W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clinical Psychology Review, 30*(2), 203–216. https://doi.org/10.1016/j.cpr.2009.11.003

Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the National Institute of Mental Health's Research Domain Criteria (RDoC). *Psychological Science in the Public Interest, 18*(2), 72–145. https://doi.org/10.1177/1529100617727266

Cole, P. M., Michel, M. K., & Teti, L. O. (1994). The development of emotion regulation and dysregulation: A clinical perspective. *Monographs of the Society for Research in Child Development, 59*(2–3), 73–102. https://doi.org/10.2307/1166139

Cusack, C. E., Ralph-Nearman, C., Christian, C., Fisher, A. J., & Levinson, C. A. (2024). Understanding heterogeneity, comorbidity, and variability in depression: Idiographic models and depression outcomes. *Journal of Affective Disorders, 356*, 248–256. https://doi.org/10.1016/j.jad.2024.04.034

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

De la Barrera, U., Arrigoni, F., Monserrat, C., Montoya-Castilla, I., & Gil-Gómez, J.-A. (2024). Using ecological momentary assessment and machine learning techniques to predict depressive

symptoms in emerging adults. *Psychiatry Research, 332*, Article 115710. https://doi.org/10.1016/j.psychres.2023.115710

Dennis-Tiwary, T. A., Roy, A. K., Denefrio, S., & Myruski, S. (2019). Heterogeneity of the anxiety-related attention bias: A review and working model for future research. *Clinical Psychological Science, 7*(5), 879–899. https://doi.org/10.1177/2167702619838474

de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: A practical guide*. Cambridge University Press. https://doi.org/10.1017/CBO9780511996214

Ehlers, A., Mauchnik, J., & Handley, R. (2012). Reducing unwanted trauma memories by imaginal exposure or autobiographical memory elaboration: An analogue study of memory processes. *Journal of Behavior Therapy and Experimental Psychiatry, 43*(Suppl 1), S67–S75. https://doi.org/10.1016/j.jbtep.2010.12.009

Epskamp, S., Deserno, M., & Bringmann, L. (2024). *mlVAR: Multi-level vector autoregression* (Version 0.5.2) [R package]. CRAN. https://CRAN.R-project.org/package=mlVAR

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research, 53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Epton, T., Currie, S., & Armitage, C. J. (2017). Unique effects of setting goals on behavior change: Systematic review and meta-analysis. *Journal of Consulting and Clinical Psychology, 85*(12), 1182–1198. https://doi.org/10.1037/ccp0000260

Everaert, J., Bernstein, A., Joormann, J., & Koster, E. H. W. (2020). Mapping dynamic interactions among cognitive biases in depression. *Emotion Review, 12*(2), 93–110. https://doi.org/10.1177/1754073919892069

Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y.-W., Warren, F., Howard, S., Ghinai, R., Fearon, P., & Lowyck, B. (2016). Development and validation of a self-report measure of mentalizing: The Reflective Functioning Questionnaire. *PLoS One, 11*(7), Article e0158678. https://doi.org/10.1371/journal.pone.0158678

Fried, E. I., & Cramer, A. O. J. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science, 12*(6), 999–1020. https://doi.org/10.1177/1745691617705892

Fries, J. F., Lingala, B., Siemons, L., Glas, C. A., Cella, D., Hussain, Y. N., Bruce, B., & Krishnan, E. (2014). Extending the floor and the ceiling for assessment of physical function. *Arthritis & Rheumatology, 66*(5), 1378–1387. https://doi.org/10.1002/art.38342

Frumkin, M. R., Piccirillo, M. L., Beck, E. D., Grossman, J. T., & Rodebaugh, T. L. (2021). Feasibility and utility of idiographic models in the clinic: A pilot study. *Psychotherapy Research, 31*(4), 520–534. https://doi.org/10.1080/10503307.2020.1805133

Gilboa-Schechtman, E. (2024). Case conceptualization in clinical practice and training. *Clinical Psychology in Europe, 6*(Special Issue), Article e12103. https://doi.org/10.32872/cpe.12103

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology, 2*(3), 271–299. https://doi.org/10.1037/1089-2680.2.3.271

Hayes, S. C., Hofmann, S. G., Stanton, C. E., Carpenter, J. K., Sanford, B. T., Curtiss, J. E., & Ciarrochi, J. (2019). The role of the individual in the coming era of process-based therapy. *Behaviour Research and Therapy, 117*, 40–53. https://doi.org/10.1016/j.brat.2018.10.005

Hayes, S. C., Hofmann, S. G., & Wilson, D. S. (2020). Clinical psychology is an applied evolutionary science. *Clinical Psychology Review, 81*, Article 101892. https://doi.org/10.1016/j.cpr.2020.101892

Hevey, D. (2018). Network analysis: A brief overview and tutorial. *Health Psychology and Behavioral Medicine, 6*(1), 301–328. https://doi.org/10.1080/21642850.2018.1521283

Huibers, M. J. H., Lorenzo-Luaces, L., Cuijpers, P., & Kazantzis, N. (2021). On the road to personalized psychotherapy: A research agenda based on cognitive behavior therapy for depression. *Frontiers in Psychiatry, 11*, Article 607508. https://doi.org/10.3389/fpsyt.2020.607508

Hvenegaard, M., Moeller, S. B., Poulsen, S., Gondan, M., Grafton, B., Austin, S. F., Kistrup, M., Rosenberg, N. G. K., Howard, H., & Watkins, E. R. (2020). Group rumination-focused cognitive-behavioural therapy (CBT) v. group CBT for depression: Phase II trial. *Psychological Medicine, 50*(1), 11–19. https://doi.org/10.1017/S0033291718003835

Joseph, S., & Wood, A. (2010). Assessment of positive functioning in clinical psychology: Theoretical and practical issues. *Clinical Psychology Review, 30*(7), 830–838. https://doi.org/10.1016/j.cpr.2010.01.002

Kalokerinos, E. K., Greenaway, K. H., & Casey, J. P. (2017). Context shapes social judgments of positive emotion suppression and expression. *Emotion, 17*(1), 169–186. https://doi.org/10.1037/emo0000222

Kanfer, F. H., & Saslow, G. (1965). Behavioral analysis: An alternative to diagnostic classification. *Archives of General Psychiatry, 12*(6), 529–538. https://doi.org/10.1001/archpsyc.1965.01720360001001

Knowles, M. M., Foden, P., El-Deredy, W., & Wells, A. (2016). A systematic review of efficacy of the attention training technique in clinical and nonclinical samples. *Journal of Clinical Psychology, 72*(10), 999–1025. https://doi.org/10.1002/jclp.22312

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Luborsky, L., Popp, C., Luborsky, E., & Mark, D. (1994). The core conflictual relationship theme. *Psychotherapy Research, 4*(3-4), 172–183. https://doi.org/10.1080/10503309412331334012

McRae, K., & Gross, J. J. (2020). Emotion regulation. *Emotion, 20*(1), 1–9. https://doi.org/10.1037/emo0000703

Mestdagh, M., Pe, M., Pestman, W., Verdonck, S., Kuppens, P., & Tuerlinckx, F. (2018). Sidelining the mean: The relative variability index as a generic mean-corrected variability measure for bounded variables. *Psychological Methods, 23*(4), 690–707. https://doi.org/10.1037/met0000153

Mink, F., Lutz, W., & Hehlmann, M. I. (2025). Ecological momentary assessment in psychotherapy research: A systematic review. *Clinical Psychology Review, 117*, Article 102565. https://doi.org/10.1016/j.cpr.2025.102565

Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time series missing value imputation in R. *The R Journal, 9*(1), 207–218. https://doi.org/10.32614/RJ-2017-009

Nemani, A., Hufschmidt, B., Kohl, V., Sendig, L., Ebert, M., Bonarius, D., Hofmann, S. G., & Stangier, U. (2025). Bipolar vs. unipolar scaling in dynamic network analyses of ecological momentary assessment data. *PLoS One, 20*(3), Article e0314102. https://doi.org/10.1371/journal.pone.0314102

Nestler, E. J. (2015). Role of the brain's reward circuitry in depression: Transcriptional mechanisms. *International Review of Neurobiology, 124*, 151–170. https://doi.org/10.1016/bs.irn.2015.07.003

Phelps, E. A., & Hofmann, S. G. (2019). Memory editing from science fiction to clinical practice. *Nature, 572*(7767), 43–50. https://doi.org/10.1038/s41586-019-1433-7

Posit Team. (2024). *RStudio* (Version 2024.04.2) [Computer software]. Posit Software, PBC. https://posit.co/download/rstudio-desktop/

Rodebaugh, T. L., Scullin, R. B., Langer, J. K., Dixon, D. J., Huppert, J. D., Bernstein, A., Zvielli, A., & Lenze, E. J. (2016). Unreliability as a threat to understanding psychopathology: The cautionary tale of attentional bias. *Journal of Abnormal Psychology, 125*(6), 840–851. https://doi.org/10.1037/abn0000184

Roefs, A., Fried, E. I., Kindt, M., Martijn, C., Elzinga, B., Evers, A. W. M., Wiers, R. W., Borsboom, D., & Jansen, A. (2022). A new science of mental disorders: Using personalised, transdiagnostic, dynamical systems to understand, model, diagnose and treat psychopathology. *Behaviour Research and Therapy, 153*, Article 104096. https://doi.org/10.1016/j.brat.2022.104096

Rubel, J. A., Fisher, A. J., Husen, K., & Lutz, W. (2018). Translating person-specific network models into personalized treatments: Development and demonstration of the dynamic assessment treatment algorithm for individual networks (DATA-IN). *Psychotherapy and Psychosomatics, 87*(4), 249–251. https://doi.org/10.1159/000487769

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Schreiber, F., Heimlich, C., Schweitzer, C., & Stangier, U. (2015). Cognitive therapy for social anxiety disorder: The impact of the "self-focused attention and safety behaviours experiment" on the course of treatment. *Behavioural and Cognitive Psychotherapy, 43*(2), 158–166. https://doi.org/10.1017/S1352465813000672

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Short, N. A., Boffa, J. W., Clancy, K., & Schmidt, N. B. (2018). Effects of emotion regulation strategy use in response to stressors on PTSD symptoms: An ecological momentary assessment study. *Journal of Affective Disorders, 230*, 77–83. https://doi.org/10.1016/j.jad.2017.12.063

Siepe, B. S., Rieble, C. L., Tutunji, R., Rimpler, A., März, J., Proppert, R. K. K., & Fried, E. I. (2025). Understanding ecological-momentary-assessment data: A tutorial on exploring item performance in ecological-momentary-assessment data. *Advances in Methods and Practices in Psychological Science, 8*(1), 1–20. https://doi.org/10.1177/25152459241286877

Šimkovic, M., & Träuble, B. (2019). Robustness of statistical methods when measure is affected by ceiling and/or floor effect. *PLoS One, 14*(8), Article e0220889. https://doi.org/10.1371/journal.pone.0220889

Stefana, A., Fusar-Poli, P., Vieta, E., & Youngstrom, E. A. (2024). Assessing the patient's affective perception of their psychotherapist: Validation of the in-Session Patient Affective Reactions Questionnaire. *Frontiers in Psychiatry, 15*, Article 1346760. https://doi.org/10.3389/fpsyt.2024.1346760

Terwee, C. B., Bot, S. D. M., De Boer, M. R., Van Der Windt, D. A. W. M., Knol, D. L., Dekker, J., Bouter, L. M., & De Vet, H. C. W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology, 60*(1), 34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012

Vacay GmbH. (2020). *Status by Vacay* (Version 3.0.1) [Mobile app]. Google Play Store. https://play.google.com/store/apps/details?id=dev.vacay.sts.android

Van den Bergh, O., & Walentynowicz, M. (2016). Accuracy and bias in retrospective symptom reporting. *Current Opinion in Psychiatry, 29*(5), 302–308. https://doi.org/10.1097/YCO.0000000000000267

van Os, J., Linscott, R. J., Myin-Germeys, I., Delespaul, P., & Krabbendam, L. J. P. M. (2009). A systematic review and meta-analysis of the psychosis continuum: Evidence for a psychosis proneness–persistence–impairment model of psychotic disorder. *Psychological Medicine, 39*(2), 179–195. https://doi.org/10.1017/S0033291708003814

Watkins, E. R., & Moulds, M. L. (2005). Distinct modes of ruminative self-focus: Impact of abstract versus concrete rumination on problem solving in depression. *Emotion, 5*(3), 319–328. https://doi.org/10.1037/1528-3542.5.3.319

Yamada, A. M., Lee, K. K., Dinh, T. Q., Barrio, C., & Brekke, J. S. (2010). Intrinsic motivation as a mediator of relationships between symptoms and functioning among individuals with schizophrenia spectrum disorders in a diverse urban community. *The Journal of Nervous and Mental Disease, 198*(1), 28–34. https://doi.org/10.1097/NMD.0b013e3181c8aa71

Yim, S. J., Lui, L. M. W., Lee, Y., Rosenblat, J. D., Ragguett, R. M., Park, C., Subramaniapillai, M., Cao, B., Zhou, A., Rong, C., Lin, K., Ho, R. C., Coles, A. S., Majeed, A., Wong, E. R., Phan, L., Nasri, F., & McIntyre, R. S. (2020). The utility of smartphone-based, ecological momentary assessment for depressive symptoms. *Journal of Affective Disorders, 274*, 602–609. https://doi.org/10.1016/j.jad.2020.05.116

EACLIPT

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

zpid
Leibniz-Institut für Psychologie

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

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

Preliminary Evaluation of the Psychometric Properties of the Ukrainian Traumatic Grief Inventory-Self Report Plus (TGI-SR+)

Maria Luisa F. Rispa Hoyos [1], Lieke C. J. Nijborg [1], Iryna Norkina [2], Paul A. Boelen [2,3], Lonneke I. M. Lenferink [1]

[1] *Department of Psychology, Health & Technology, Faculty of Behavioural Management and Social Sciences, University of Twente, Enschede, The Netherlands.* **[2]** *Department of Clinical Psychology, Faculty of Social Sciences, Utrecht University, Utrecht, The Netherlands.* **[3]** *ARQ National Psychotrauma Centre, Diemen, The Netherlands.*

Clinical Psychology in Europe, 2026, Vol. 8(2), Article e17367, https://doi.org/10.32872/cpe.17367

Received: 2025-03-17 • **Accepted:** 2025-12-03 • **Published (VoR):** 2026-05-29

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

Corresponding Author: Lonneke I. M. Lenferink, Department of Psychology, Health & Technology, Faculty of Behavioural Management and Social Sciences, University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands. (T: +31534894686). E-mail: l.i.m.lenferink@utwente.nl

Supplementary Materials: Data, Materials [see Index of Supplementary Materials]

Abstract

Background: Prolonged Grief Disorder (PGD) was included in the DSM-5-TR and ICD-11. The Traumatic Grief Inventory Self-Report Plus (TGI-SR+) is a self-report measure to assess PGD symptoms in accordance with both classification systems. It has been translated into various languages and validated across different contexts.

Objective: Evaluating the psychometric properties of the Ukrainian TGI-SR+.

Method: Participants were Ukrainian adults who had lost a loved one at least 6 months ago. One hundred ninety-eight participants completed the TGI-SR+ and measures assessing posttraumatic stress and depression. We examined the factor structure, internal consistency, convergent validity, and known-groups validity of the TGI-SR+. Moreover, rates of probable PGD caseness were calculated, and provisional cut-off scores were determined based on Receiver Operating Characteristic (ROC) analyses.

Results: The one-factor model showed an acceptable fit for DSM-5-TR, but not for ICD-11 PGD symptoms. While some of the factor loadings were low for both criteria sets, the items demonstrated good internal consistency. Convergent validity was supported by strong associations between symptom levels of DSM-5-TR and ICD-11 PGD and posttraumatic stress and depression

severity scores. Known-groups validity was partially supported by DSM-5-TR and ICD-11 PGD severity being related to both cause of death and kinship to the deceased. The provisional cut-off score for detecting both probable DSM-5-TR and ICD-11 PGD caseness, when summing all TGI-SR+ items, was ≥ 75.

Conclusion: The psychometric properties of the Ukrainian TGI-SR+ were mixed. However, pending further research in different and larger samples of Ukrainian bereaved people, this instrument can be used to assess PGD severity in Ukrainians.

Keywords

grief, war, loss, assessment, Russian invasion, bereavement

Highlights

- Prolonged grief symptoms were prevalent among Ukrainians and strongly correlated with posttraumatic stress and depressive symptoms.
- Losses due to violent or unexpected causes were associated with more severe and persistent prolonged grief symptoms than natural losses.
- Provisional cut-off scores were identified to screen for probably prolonged grief disorder among bereaved Ukrainians.

The death of a loved one is a potentially stressful life event that can have a substantial negative effect on the mental health of those left behind (Seiler et al., 2020). Most individuals can adjust to the loss (Nielsen et al., 2019), but about 3-5% experience persistent grief symptoms after losing someone to a natural cause (Rosner et al., 2021; Treml, Linde, et al., 2024), and this rate is four times higher in cases of unexpected losses (e.g., accidents) (Doering et al., 2022).

The definition of pathological grief has generated discussion among experts (Eisma, 2023; Eisma et al., 2022; Haneveld et al., 2022; Lenferink et al., 2021; Prigerson et al., 2024). The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) included Persistent Complex Bereavement Disorder (PCBD) (American Psychiatric Association [APA], 2013) and was replaced by Prolonged Grief Disorder (PGD) in the text revision of the DSM-5 (i.e., DSM-5-TR; APA, 2022). According to the DSM-5-TR, PGD encompasses two separation distress symptoms (i.e., intense yearning or longing for, and preoccupation with thoughts or memories of, the deceased person), and eight accompanying symptoms (e.g., identity disruption, intense emotional pain, and emotional numbness; APA, 2022). DSM-5-TR PGD can be diagnosed one year after the loss for adults and six months for children, when the bereaved person experiences at least one separation distress symptom and three or more accompanying symptoms daily for at least a month, exceeding social, cultural, or religious norms and causing significant distress or impairment in daily activities.

PGD has also been included in the 11th edition of the International Classification of Diseases (ICD-11; World Health Organization [WHO], 2025). Diagnostic criteria in ICD-11 differ slightly from those in DSM-5-TR in terms of the number and content of symptoms. ICD-11 PGD has two separation distress symptoms and ten accompanying symptoms (e.g., sadness, guilt, anger, and denial) and can be diagnosed six months after the loss for both adults and children, when the bereaved person experiences at least one separation distress symptom and one accompanying symptom daily for at least a month, causing functional impairment.

Several instruments have been created to assess PGD levels according to the DSM-5-TR (e.g., Prolonged Grief Disorder-13 – Revised, Prigerson et al., 2021) and ICD-11 diagnostic criteria sets (e.g., International Prolonged Grief Disorder Scale, Killikelly et al., 2020), separately. However, assessing DSM-5-TR and ICD-11 PGD symptoms simultaneously using one measurement tool, such as the Traumatic Grief Inventory-Self Report Plus (TGI-SR+; Lenferink et al., 2022), has several advantages. The TGI-SR+ has been translated into more than 15 different languages and is freely accessible on the Open Science Framework, facilitating cross-national assessment of PGD severity (see https://osf.io/rqn5k/), while mitigating barriers to comparing results across cultures (Comtesse et al., 2024).

The Dutch TGI-SR+ has been validated in distinct samples of people bereaved by various causes and people bereaved by traffic accidents. In both samples, the TGISR+ demonstrated robust psychometric properties, as indicated by good construct validity, internal consistency, temporal stability, convergent and known-groups validity (Lenferink et al., 2022). These findings have been replicated with the French TGI-SR+ with young adults (Kokou-Kpolou et al., 2022), the Swedish TGI-SR+ in bereaved parents (Lenferink, Van Dijk, et al., 2024), the Norwegian TGI-SR+ in people whose child or sibling died unexpectedly or violently (Lenferink, Johnsen, et al., 2024), and the Chinese TGI-SR+ in adults bereaved due to COVID-19 (Tang et al., 2024). However, validation remains largely concentrated in Western Europe. Expanding the psychometric evaluation to non-Western countries is critical, particularly in regions with high mortality rates and armed conflicts, where the risk for PGD is elevated (Kokou-Kpolou et al., 2020).

A reliable and valid screening tool for PGD is crucial for identifying people who may need professional support. The present study evaluated the psychometric properties of the Ukrainian TGI-SR+ to assess DSM-5-TR and ICD-11 based PGD symptoms. The Russian invasion resulted in one of Europe's largest forced displacement crises, affecting more than 44 million Ukrainians (Shevlin et al., 2022), with over 100,000 refugees in the Netherlands (Government of the Netherlands, 2024). Forced refugees have a high risk of developing mental health conditions, such as posttraumatic stress disorder (PTSD), depression, and anxiety (Charlson et al., 2019; Steel et al., 2009). PGD has also been observed in displaced populations (Bryant et al., 2020) and may be commonplace among Ukrainians (Killikelly et al., 2024; Shevlin et al., 2022).

To psychometrically evaluate the Ukrainian TGI-SR+, we evaluated the factor structure of both DSM-5-TR PGD and ICD-11 PGD. Based on prior research (Kokou-Kpolou et al., 2022; Lenferink et al., 2022; Lenferink, Johnsen, et al., 2024), we expected the one-factor model to fit the data best for both DSM-5-TR PGD and ICD-11 PGD. Moreover, we expected to find high internal consistency (Kokou-Kpolou et al., 2022; Lenferink et al., 2022; Lenferink, Van Dijk, et al., 2024; Tang et al., 2024).

To demonstrate convergent validity, we examined to what extent DSM-5-TR PGD and ICD-11 PGD levels are associated with PTSD and depression levels, expecting to find strong positive correlations between these constructs (cf. Boelen, 2021; Eisma et al., 2019; Kokou-Kpolou, 2021; Lenferink et al., 2022; Maccallum & Bryant, 2018; Schaal et al., 2012). Regarding known-groups validity, we hypothesized that people self-identified as women (vs. men) and those who lost a parent, partner, or child (vs. a more distant close person) would report higher DSM-5-TR and ICD-11 PGD levels. Also, we expected that losses due to an unnatural cause (e.g., homicide, accident, suicide) would be associated with higher DSM-5-TR and ICD-11 PGD levels than losses due to a natural cause (e.g., old age, chronic illness) and that time since loss would be negatively associated with these PGD levels (Buur et al., 2024; Doering et al., 2022). Lastly, we determined cut-off scores to distinguish between people who meet the criteria for probable DSM-5-TR PGD and ICD-11 PGD caseness and those who do not.

Material and Method

Participants and Procedures

Data were collected online from Ukrainian adults (≥ 18 years) who had lost a loved one at least six months earlier (for recruitment strategies and procedures, see Supplementary Material A).

In total, 430 individuals started the survey. We excluded responses from people whose loved one had died less than six months before completing the survey (n = 140), who did not complete all three main instruments (TGI-SR+, PCL-5, and PHQ-9; n = 81), who were minors, duplicate entries, or test responses (n = 7), and who provided ambiguous answers regarding the cause of death (e.g., going different ways; n = 4). The final sample included 198 participants.

Measures

Sociodemographic and Loss-Related Characteristics

The following sociodemographic and loss-related characteristics were assessed: gender (1 = man, 2 = woman, 3 = other), age (in years), whether the participant fled their home due to war (1 = no, 2 = yes, I fled my house and live somewhere else in Ukraine, and 3 = yes, now I live in a different country), date of death of their deceased loved one, cause

of death (1 = physical illness, 2 = accident, 3 = suicide, 4 = murder or manslaughter not related to war, 5 = killed in action/combat, 6 = disappearance, 7 = other), and kinship (1 = partner, 2 = child, 3 = parent, 4 = sibling, 5 = grandparent, 6 = grandchild, 7 = friend, and 8 = other).

Traumatic Grief Inventory-Self-Report Plus (TGI-SR+)

The TGI-SR+ is a 22-item extended version of the original TGI-SR (Boelen & Smid, 2017) designed to assess PGD severity as defined in the DSM-5-TR and ICD-11 (Lenferink et al., 2022). Participants rated each item from 1 (never) to 5 (always) to indicate the extent to which they experienced a symptom during the past month (e.g., I avoided places, objects, or thoughts that reminded me that the person I have lost has died). For information about the translation, see Supplementary Material B.

Posttraumatic Stress Disorder Checklist (PCL-5) 6-Item Short Form

PTSD severity was assessed using a six-item ICD-11 subset of the PCL-5 (Items 2, 3, 6, 7, 17, and 18; Heeke et al., 2022; Weathers et al., 2013). Participants rated symptom presence in the preceding month on a 5-point scale with anchors 1 (not at all) and 5 (extremely), with total scores ranging from 6 to 30. The scores were dichotomized such that Scores 1 − 2 indicated symptom absence, while scores ≥ 3 denoted symptom presence (Heeke et al., 2022). Probable caseness required the endorsement of all items.

The 6-item PCL-5 has good psychometric properties (Heeke et al., 2022). The wording of items was changed to refer to the loss instead of the stressful event (e.g., In the past month, how much were you bothered by avoiding external reminders of the death of your loved one (for example, people, places, conversations, activities, objects, or situations)?). We used the items from the Ukrainian version of PCL-5, which has satisfactory psychometric properties (Johnson et al., 2022; Roberts et al., 2019; Ukrainian translation: International Trauma Consortium, n.d.). In this present study, the 6-item PCL-5 showed acceptable internal consistency (ω = .76).

Patient Health Questionnaire – 9 (PHQ-9)

The PHQ-9 consists of nine items used to assess depression severity in the last two weeks (e.g., feeling down, depressed, or hopeless; Kroenke et al., 2001). Each item was rated from 0 (not at all) to 3 (almost every day). A total depression score was calculated by summing all items (range: 0–27). A total score ≥ 10 indicates probable caseness of depression (Gilbody et al., 2007; Kroenke et al., 2001). For this study, the validated Ukrainian translation of the PHQ-9 was used (Hyland et al., 2023; Riad et al., 2022; Ukrainian translation: International Trauma Consortium, n.d.). Internal consistency in the present sample was good (ω = .86).

Statistical Analysis

Confirmatory factor analyses (CFAs) were performed using Mplus 7.4 (Muthén & Muthén, 1998). The remaining analyses were performed using IBM SPSS, version 28.0 (IBM Corporation, 2017).

Factor Structure of DSM-5-TR and ICD-11-Based PGD Symptoms

Kurtosis and skewness values were examined. Kurtosis values < 10 and skewness values < 3 are indicative of a univariate normal distribution of item scores (Kline, 2011). All kurtosis and skewness values were < 2; consequently, the default maximum likelihood estimator was used. There were no missing data on PGD items.

CFAs were performed to evaluate the factor structure of DSM-5-TR PGD and ICD-11 based PGD symptoms. First, a one-factor model was evaluated for DSM-5-TR PGD and ICD-11 PGD, separately. Second, following prior research (Kokou-Kpolou et al., 2022; Lenferink et al., 2022; Lenferink, Johnsen, et al., 2024; Lenferink, Van Dijk, et al., 2024; Tang et al., 2024), a two-factor model was evaluated. In the two-factor model, the two items representing separation distress (i.e., the yearning and preoccupation items) load on one factor, while the items representing accessory symptoms load on a second factor.

Model fit was evaluated using the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI), where values > .95 indicate excellent fit and > .90 acceptable fit. The root-mean-square error of approximation (RMSEA) and standardized root mean square residual (SRMR) were also evaluated, with values < .05 indicating excellent fit and < .10 indicating acceptable fit (Kline, 2011). In addition, the lower the Akaike information criterion (AIC) and Bayesian information criterion (BIC), the better the fit. Factor loadings below 0.6 are considered low.

Internal Consistency

Reliability of the DSM-5-TR PGD and ICD-11 PGD items was evaluated using McDonald's Omega (ω). Values of $\omega > 0.70$ indicate acceptable internal consistency (Hayes & Coutts, 2020; Trizano-Hermosilla & Alvarado, 2016).

Convergent Validity

To examine the convergent validity of the TGI-SR+, correlation analyses were performed. Based on Kolmogorov-Smirnov normality tests, summed scores on the items representing the DSM-5-TR PGD and ICD-11 PGD symptoms were non-normally distributed (i.e., $p < .05$). Spearman's Rho correlations were calculated to examine associations between DSM-5-TR PGD and ICD-11 PGD total scores and PTSD and depression total scores. Correlations ≤ 0.29 were considered weak, between 0.30 and 0.49 moderate, and ≥ 0.50 strong (Cohen, 1988).

Known-Groups Validity

Mann-Whitney U tests were carried out to examine whether DSM-5-TR PGD and ICD-11 PGD total scores differed as a function of gender cause of death (dichotomized into unnatural cause of death (i.e., accident, suicide, killed in combat, disappearance) and natural cause of death [i.e., physical illness, old age, died at birth, etc.]), and kinship (dichotomized into nuclear family member [i.e., child, partner, parent| and other).

Moreover, it was examined whether the time since loss (in months) was related to PGD total scores, using Spearman's Rho correlations.

Rate of Probable PGD Caseness Using Diagnostic Scoring Rules

A participant qualified for probable DSM-5-TR PGD caseness when endorsing at least one of the two Criterion B (i.e., separation distress) symptoms (Items 1 and 3), three or more of the eight Criterion C (i.e., cognitive, emotional, and behavioral) symptoms (Items 6, 9, 10, 11, 19, 21, and the highest score indicated on items 2 and 8[1]), and item 13, representing Criterion D (i.e., functional impairment; APA, 2022).

A participant qualified for probable ICD-11 PGD caseness when endorsing at least one of the two criterion B symptoms (Items 1 and 3), at least one of the 10 criterion C symptoms (Items 2, 5, 8, 9, 10, 16, 19, 20, 21, and 22), and item 13, representing Criterion D (i.e., functional impairment; World Health Organization, 2025). Items were considered "endorsed" when scored at least 3.

Optimal Cut-Off Scores

By performing receiver operating characteristic (ROC) analyses, we determined the optimal cut-off scores to identify probable DSM-5-TR and ICD-11 PGD cases. Specifically, we examined optimal cut-off scores based on (1) the total score of all 22 TGI-SR+ items (range: 22-110), (2) the total score of all DSM-5-TR based PGD items, excluding the functional impairment item (range: 10-50), and (3) the total score of all ICD-11 based PGD items, excluding the functional impairment item (range: 12–60).

A ROC was plotted with the true positive ratio (i.e., sensitivity) as a function of the false positive ratio (i.e., 1 – specificity) for each possible total score. Area Under the Curve (AUC) values ≥ 0.90 indicate that the score has excellent accuracy in distinguishing probable "cases" from "non-cases" (Ferraris, 2019). In addition, the Youden index (i.e., sensitivity index – [1 – specificity index]) was calculated. Values between 0.90 and 1 are excellent and indicate high diagnostic accuracy (Schisterman et al., 2005), between 0.80 and 0.90 are considered good, and between 0.70 and 0.80 are fair. When values < .70, this indicates a poor accuracy in distinguishing probable caseness from probable non-caseness (Ferraris, 2019).

1) One symptom, emotional pain, of criterion C (i.e., C4) is represented by two items (2 and 8).

Results

Sample Characteristics

The characteristics of the sample are shown in Table 1. Nine out of 10 participants self-identified as women. On average, participants were 36 years old (*SD* = 8.79). Approximately three out of 10 participants fled from Ukraine to another country due to the war with Russia. Most participants lost their loved one due to a physical illness, followed by being killed in action/combat, or an accident. A majority of participants reported that they had lost a nuclear family member. On average, the loss took place around four years ago. About one out of 10 people met the criteria for probable PTSD, and six out of 10 people met the criteria for probable depression.

Factor Structure of DSM-5-TR and ICD-11 PGD Items

For DSM-5-TR PGD, both the one- and two-factor models showed acceptable fit, as indicated by the CFI, TLI, RMSEA, and SRMR values (see Table 2). However, the two-factor model did not fit significantly better than the one-factor model ($\Delta\chi^2$ (Δdf) = 2.36 (1), p < .20) and, accordingly, the one-factor and two-factor models had similar AIC and BIC values. The one-factor model was therefore selected as the optimal model (see Table 3 for the standardized factor loadings), despite very low factor loadings for the items "I had intrusive thoughts or images related to the person who died" and "I avoided places, objects, or thoughts that reminded me that the person I lost has died".

For ICD-11 PGD, none of the fit indices indicated an acceptable fit for the one-factor or the two-factor model, except for the SRMR values (see Table 2). When comparing the fit of the two-factor model to the one-factor model, there was no significant improvement in fit ($\Delta\chi^2$ (Δdf) = 2.39 (1), p < .20), and AIC and BIC values of both models were similar. Table 3 shows the factor loadings of the one-factor model. Again, some of the factor loadings were very low, with the lowest loadings for the following two items: "I had intrusive thoughts or images related to the person who died" and "I put an intense blame on others because of his/her death".

Internal Consistency

The DSM-5-TR PGD items and the ICD-11 PGD items displayed acceptable internal consistency (both ω = .86).

Convergent Validity

As expected, significant, positive, and strong associations were found between DSM-5-TR PGD levels and PTSD (ρ = .54, p < .001) and depression levels (ρ = .50, p < .001). Significant, positive, and strong associations were also found between ICD-11 PGD levels and PTSD (ρ = .53, p < .001) and depression levels (ρ = .51, p < .001).

Table 1

Characteristics of the Sample (N = 198)

Variable	*n* (%) or *M* (*SD*) & range
Gender (*n*, %)	
Male	7 (3.5)
Female	184 (92.9)
Other	7 (3.5)
Age (in years; *M*, *SD*, range)	36.23 (8.79); 19 – 59
Fled home due to war (*n*, %)	
No	111 (56.1)
Yes, inside Ukraine	28 (14.1)
Yes, to another country	59 (29.8)
Cause of death (*n*, %)	
Physical illness (e.g., cancer, cardiovascular disease, died at birth)	116 (58.6)
Accident (e.g., traffic accident, drowning, poisoning)	19 (9.6)
Suicide	10 (5.1)
Murder or manslaughter not related to war	3 (1.5)
Killed in action/combat	37 (18.7)
Disappearance	4 (2.0)
Other	9 (4.5)
Kinship, the deceased was the participant's ... (*n*, %)	
Partner (husband, wife, boyfriend, girlfriend)	42 (21.2)
Child	26 (13.1)
Parent	73 (36.9)
Sibling	17 (8.6)
Grandparent	18 (9.1)
Friend(s)	12 (6.1)
Other	10 (5.1)
Time since the loss (in months; *M*, *SD*, range)	45.85 (58.50); 6 – 422
Psychological Outcomes (*M*, *SD*, range)	
DSM-5-TR PGD levels (TGI-SR+)	32.48 (7.85); 13 – 49
ICD-11 PGD levels (TGI-SR+)	40.12 (8.67); 15 – 58
PTSD levels (PCL-5)	14.27 (4.77); 6 – 26
Depression levels (PHQ-9)	11.30 (6.09); 0 – 25

Note. DSM-5-TR = 5th text revised edition of the Diagnostic and Statistical Manual of Mental Disorders; ICD-11 = 11th edition of the International Classification of Diseases; PCL-5 = Posttraumatic stress disorder Checklist for DSM-5; PGD = Prolonged Grief Disorder; PHQ-9 = Patient Health Questionnaire 9; PTSD = posttraumatic stress disorder; TGI-SR + = Traumatic Grief Inventory – Self Report Plus.

Table 2

Fit Indices of the Confirmatory Factor Analyses (N = 198)

Model	χ^2	*df*	CFI	TLI	RMSEA [90% CI]	SRMR	AIC	BIC
DSM-5-TR PGD								
One-factor	92.36	35	.93	.91	.09 [.07, .11]	.05	5538.97	5637.62
Two-factor	90.00	34	.93	.91	.09 [.07, .11]	.05	5538.61	5640.54
ICD – 11 PGD								
One-factor	169.06	54	.87	.84	.10 [.09, .12]	.06	6680.38	6798.76
Two-factor	166.67	53	.87	.84	.10 [.09, .12]	.06	6679.99	6801.65

Note. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; CFI = Comparative Fit Index; CI = Confidence Interval; DSM-5-TR = Diagnostic and Statistical Manual of Mental Disorders 5 Text-Revision; ICD-11 = the 11th edition of the International Classification of Diseases; PGD = Prolonged Grief Disorder; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; TLI = Tucker Lewis Index.

Known-Groups Validity

As expected, participants who lost a nuclear family member reported more severe DSM-5-TR PGD symptoms than those who lost another loved one (U = 3293.5, p = .047). Kinship was unrelated to ICD-11 PGD severity (U = 3383.0, p = .081). Moreover, participants who lost someone due to an unnatural cause reported higher symptoms of DSM-5-TR PGD (U = 3072.5, p <.001) and ICD-11 PGD (U = 3057.5, p <.001) than those who lost someone due to a natural cause.

Time since loss was significantly and negatively associated with the severity of DSM-5-TR PGD (ρ = -.162, p = .02) and trended toward being significantly associated with ICD-11 PGD severity (ρ = -.133, p = .06). Due to the overrepresentation of women, we were unable to examine gender differences.

Rates of Probable PGD Caseness

Seventy-five participants (37.9%) met criteria for probable DSM-5-TR PGD caseness, whereas 80 (40.4%) met criteria for probable ICD-11 PGD caseness.

Optimal Cut-Off Scores

When summing all 22-item TGI-SR+, the optimal cut-off score indicative of probable DSM-5-TR PGD caseness was ≥ 75 (AUC = 0.921, 95% CI: 0.884–0.957). The Youden's index suggested poor diagnostic accuracy (J = 0.65). With this cut-off score, 78% of probable cases were correctly identified as DSM-5-TR PGD cases, and 13% were incorrectly identified. When summing the 10 DSM-5-TR PGD items (range between 10-50), the optimal cut-off score indicative of probable DSM-5-TR PGD caseness was ≥ 34 (AUC =

0.919, 95% CI: 0.882–0.955). With this cut-off score, 85% of probable cases were correctly identified as a DSM-5-TR PGD case, and 17% were incorrectly identified. The Youden's index suggested poor diagnostic accuracy (J = 0.68).

Table 3

Standardized Factor Loadings One-Factor Model for DSM-5-TR and ICD-11 PGD (N = 198)

Variable	Est.	*SE*
1 factor DSM-5-TR PGD		
I found myself longing or yearning for the person who died.	.596	.050
I had intrusive thoughts or images related to the person who died.	.123	.073
It felt as if a part of me has died along with the deceased	.776	.032
It felt unreal that he/she is dead	.558	.053
I avoided places, objects, or thoughts that reminded me that the person I lost has died	.229	.068
I experienced intense emotional pain, sadness, or pangs of grief/I felt bitterness or anger related to his/her death.	.670	.043
I felt that that moving on (e.g., making new friends, pursuing new interests) was difficult for me.	.765	.034
I felt emotionally numb.	.726	.038
I felt that life is unfulfilling or meaningless without him/her.	.880	.022
I felt alone or detached from other individuals.	.725	.038
1 factor ICD-11 PGD		
I found myself longing or yearning for the person who died.	.677	.044
I had intrusive thoughts or images related to the person who died.	.096	.075
I experienced intense emotional pain, sadness, or pangs of grief.	.760	.035
I had negative thoughts about myself in relation to the loss (e.g., thoughts about self-blame).	.472	.060
I felt bitterness or anger related to his/her death.	.533	.055
It felt unreal that he/she is dead	.621	.048
I put an intense blame on others because of his/her death	.305	.069
I had trouble accepting the loss.	.709	.041
It felt as if a part of me has died along with the deceased	.741	.037
I had difficulties experiencing positive feelings	.698	.042
I felt emotionally numb.	.687	.043
I felt that moving on (e.g., making new friends, pursuing new interests) was difficult for me.	.697	.042

Note. DSM-5-TR = Diagnostic and Statistical Manual of Mental Disorders 5 Text-Revision; ICD-11 = the 11th edition of the International Classification of Diseases; PGD = Prolonged Grief Disorder.

For ICD-11 PGD, when summing all 22 items, the optimal cut-off score was also ≥ 75 (AUC = 0.905, 95% CI: 0.865–0.945). The Youden's index was fair (J = 0.71) and 83% of probable cases were correctly identified as ICD-11 PGD cases, whereas 12% were incorrectly identified. When summing the 12 ICD-11 PGD items (range 12-60), the optimal cut-off score was ≥ 44 (AUC = 0.875, 95% CI: 0.826–0.924); 76% of probable cases were correctly identified as an ICD-11 case, and 15% were incorrectly identified. The Youden's index again suggested poor diagnostic accuracy (J = 0.61)

Discussion

This cross-sectional study evaluated the psychometric properties of the Ukrainian TGI-SR+ to assess PGD according to the DSM-5-TR and ICD-11, among 198 bereaved Ukrainians.

We assessed the factor structure according to both the DSM-5-TR and ICD-11 diagnostic criteria set. We found that, as expected, the one-factor model showed adequate fit for the DSM-5-TR PGD items. A poor model fit was found for the ICD-11 PGD items. Notably, for both one-factor DSM-5-TR and ICD-11 PGD models, we found that some items had low factor loadings (e.g., Items 6, 8, 16, and 19, 20), indicating weak inter-item correlations. More specifically, Item 1, "I had intrusive thoughts or images related to the person who died", had the lowest factor loading in both models. This result stands in contrast to prior research, which had shown that this item has strong factor loadings. (Cherblanc et al., 2026; Lenferink, Johnsen, et al., 2024; Treml, Schmidt, et al., 2024). One explanation could be related to the challenge of translating the content of the item into Ukrainian with a clear and unambiguous meaning, which is a common challenge in cross-cultural studies (Cruchinho et al., 2024). These items should be re-translated and reassessed in future research. In addition, the reductionist and stigmatizing culture of how mental health is often viewed among Ukrainians (Frankova et al., 2024) may also play a role in the interpretation of the items by participants and the potential incomprehensibility of the terminology used.

Concerning the poor model fit for the ICD-11 PGD, the items reflecting self-blame and blaming others showed relatively low factor loadings, mirroring previous findings (see Lenferink, Johnsen, et al., 2024; Lenferink, Van Dijk, et al., 2024). This suggests that these items may not align well with the core symptoms of PGD and may contribute to the poor factor structure and overall fit of the ICD-11 model in this study.

Notably, we found acceptable internal consistency for both the DSM-5-TR PGD and ICD-11 PGD items, consistent with previous research (Lenferink et al., 2022; Lenferink, Van Dijk, et al., 2024). Furthermore, the convergent validity was further evidenced by strong positive correlations with measures of PTSD and depression, aligning with the literature (Eisma et al., 2019; Fernández-Alcántara et al., 2025; Kokou-Kpolou, 2021; Kokou-Kpolou et al., 2022; Lenferink, Johnsen, et al., 2024).

Regarding known-groups validity, individuals who lost a close relative or experienced a more recent loss reported higher DSM-5-TR, but not ICD-11, PGD scores. In addition, those who experienced an unnatural loss (e.g., due to an accident) reported higher DSM-5-TR and ICD-11 PGD levels than those who experienced a natural loss. These findings accord with prior evidence that a close relationship to the deceased, a more recent loss, and the unnatural death of a loved one are risk factors for PGD (Buur et al., 2024). There was no significant association between ICD-11 PGD and kinship nor time since loss, which might be related to the poor factor structure of the ICD-11 PGD items.

The association between time since loss and PGD severity was almost similar for DSM-5-TR and ICD-11, with the latter showing a marginally significant effect (Pritschet et al., 2016), suggesting that the lack of significance may be due to limited statistical power. This could explain why we did find significant differences in ICD-11 PGD levels regarding kinship.

The cut-off scores for probable PGD were ≥ 34 for DSM-5-TR PGD (when summing the 10 items) and ≥ 44 for ICD-11 PGD (when summing the 12 items). These cut-off scores are similar to those found in prior research among Dutch, Chinese, German, and Swedish bereaved people, which ranged from 32 to 34 and from 39 to 44, respectively (Lenferink et al., 2022; Lenferink, Van Dijk, et al., 2024; Tang et al., 2024; Treml, Schmidt, et al., 2024). When summing all 22 TGI-SR+ items, the optimal cut-off score for determining probable DSM-5-TR and ICD-11 PGD caseness was ≥ 75. These findings also align with the cut-off scores found in prior research, which varied from 65 to 71 for DSM-5-TR PGD and from 60 to 75 for ICD-11 PGD (Lenferink, Johnsen, et al., 2024; Lenferink, Van Dijk, et al., 2024). These differences in cut-off scores across studies may partly be explained by differences in the accuracy of determining them. In our sample, we obtained poor Youden's indices, indicating low precision in differentiating between bereaved people with and without probable PGD, which is likely due to the combination of our relatively small sample and low factor loadings for some of the items.

When interpreting the results, it is important to consider several limitations. First, the sample was predominantly female; while this is common among bereavement-related research (Eisma & Stroebe, 2021; Kokou-Kpolou et al., 2022; Lenferink et al., 2022), this overrepresentation prevented gender-based comparisons. Notably, in this particular case, the lack of men participating could reflect the ongoing Russian invasion of Ukraine, as data collection coincided with military mobilization. Related to this, some participants recently fled their country or may have been exposed to war-related stressors, which may have affected the results. The underrepresentation of men may also stem from cultural factors, Ukrainian men are generally (even) less open about their emotional state than men from Western countries (Plan International, 2025).

Second, due to the cross-sectional design of the present study, it was impossible to assess the test-retest reliability and predictive validity of the TGI-SR+. Longitudinal research is necessary to evaluate these crucial psychometric properties. Third, while the TGI-SR+ is a valuable tool to assess PGD severity, it is essential to corroborate self-reported symptoms with clinician-administered diagnostic interviews. Future research should employ validated clinical interviews to accurately assess the prevalence of PGD and more reliable cut-off scores. In the present study, ROC analyses were conducted using the self-report data, as no independent diagnostic interviews were available at the time of data collection. While this is common practice in studies evaluating assessment instruments for PGD (Fernández-Alcántara et al., 2025; Kokou-Kpolou et al., 2020; Lenferink, Johnsen, et al., 2024; Lenferink, Van Dijk, et al., 2024; Treml, Schmidt, et al., 2024), the cut-off

scores derived from these analyses should be considered provisional, intended to provide clinicians with a rapid impression of potential PGD caseness, and should never replace a thorough clinical evaluation.

Additionally, while this study examined convergent validity by correlating PGD scores with PTSD and depression, divergent validity was not assessed. Although our CFA suggested a potential two-factor pattern for PGD items, the construct validity of each factor was not formally evaluated. Given the high correlation between the two factors, we did not expect significant differences in their construct validity. Therefore, we did not conduct a separate evaluation of the construct validity for each factor.

To conclude, we found mixed results regarding the psychometric properties of the Ukrainian TGI-SR+. The poor factor loadings of some items need further examination to improve the scale's validity. Despite this, the instrument offers the possibility to assess PGD severity, a priority given the psychological impact of the ongoing war.

Funding: This work was supported by Fonds Slachtofferhulp.

Acknowledgments: The authors would like to express their sincere gratitude to all participants for them valuable contributions to this research. Their willingness to participate was essential to the completion of this study.

Competing Interests: No potential conflict of interest was reported by the author(s).

Ethics Statement: This study was approved by the Behavioral, Management, and Social Sciences (BMS) Ethics Committee of the University of Twente, the Netherlands (ID: 221111).

Preregistration: This study was not preregistered.

Social Media Accounts: *Maria Luisa F. Rispa Hoyos:* LinkedIn, *Lieke C. J. Nijborg:* LinkedIn, *Iryna Norkina:* LinkedIn, *Paul A. Boelen:* LinkedIn, *Lonneke I. M. Lenferink:* LinkedIn

Reporting Guidelines: This article was prepared following the JARS-Quant guidelines.

Data Availability: The data that support the findings of this study are available in the Data Archiving and Networked Services (DANS) repository at https://doi.org/10.17026/SS/OVYGII

Supplementary Materials

The Supplementary Materials contain the following items:

- **Research data and codebook** (Rispa Hoyos & Lenferink, 2026S)
- **Additional information** (Rispa Hoyos et al., 2026S):
  - *Supplementary Material A: Recruitment Strategies and Procedures.* Description of the recruitment strategies and procedures used in the study.

- *Supplementary Material B: TGI-SR+ translation.* Description of the manner in which the TGI-SR+ was translated.

Index of Supplementary Materials

Rispa Hoyos, M. L. F., & Lenferink, L. I. M. (2026S). *Validation of the Ukranian Traumatic Inventory Self-Report (TGISR+)"* [Research data and codebook]. DANS. https://doi.org/10.17026/SS/OVYGII

Rispa Hoyos, M. L. F., Nijborg, L. C. J., Norkina, I., Boelen, P. A., & Lenferink, L. I. M. (2026S). *Supplementary materials to "Preliminary evaluation of the psychometric properties of the Ukrainian Traumatic Grief Inventory-Self Report Plus (TGI-SR+)"* [Additional information]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.21853

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

Boelen, P. A. (2021). Symptoms of prolonged grief disorder as per DSM-5-TR, posttraumatic stress, and depression: Latent classes and correlations with anxious and depressive avoidance. *Psychiatry Research, 302*, Article 114033. https://doi.org/10.1016/j.psychres.2021.114033

Boelen, P. A., & Smid, G. E. (2017). The Traumatic Grief Inventory Self-Report Version (TGI-SR): Introduction and preliminary psychometric evaluation. *Journal of Loss and Trauma, 22*(3), 196–212. https://doi.org/10.1080/15325024.2017.1284488

Bryant, R. A., Edwards, B., Creamer, M., O'Donnell, M., Forbes, D., Felmingham, K. L., Silove, D., Steel, Z., McFarlane, A. C., Van Hooff, M., Nickerson, A., & Hadzi-Pavlovic, D. (2020). A population study of prolonged grief in refugees. *Epidemiology and Psychiatric Sciences, 29*, Article e44. https://doi.org/10.1017/S2045796019000386

Buur, C., Zachariae, R., Komischke-Konnerup, K., Marello, M., Schierff, L., & O'Connor, M. (2024). Risk factors for prolonged grief symptoms: A systematic review and meta-analysis. *Clinical Psychology Review, 107*, Article 102375. https://doi.org/10.1016/j.cpr.2023.102375

Charlson, F., Van Ommeren, M., Flaxman, A., Cornett, J., Whiteford, H., & Saxena, S. (2019). New WHO prevalence estimates of mental disorders in conflict settings: A systematic review and meta-analysis. *Lancet, 394*(10194), 240–248. https://doi.org/10.1016/S0140-6736(19)30934-1

Cherblanc, J., Côté, I., Boever, C., & Zech, E. (2026). Validity of the TGI-SR+ in Francophone populations: Insights from Quebec and Belgium. *Death Studies, 50*(2), 296–306. https://doi.org/10.1080/07481187.2024.2420882

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Comtesse, H., Smid, G. E., Rummel, A.-M., Spreeuwenberg, P., Lundorff, M., & Dückers, M. L. A. (2024). Cross-national analysis of the prevalence of prolonged grief disorder. *Journal of Affective Disorders, 350*, 359–365. https://doi.org/10.1016/j.jad.2024.01.094

Cruchinho, P., López-Franco, M. D., Capelas, M. L., Almeida, S., Bennett, P. M., Da Silva, M. M., Teixeira, G., Nunes, E., Lucas, P., & Gaspar, F. (2024). Translation, cross-cultural adaptation, and validation of measurement instruments: A practical guideline for novice researchers. *Journal of Multidisciplinary Healthcare, 17*, 2701–2728. https://doi.org/10.2147/JMDH.S419714

Doering, B. K., Barke, A., Vogel, A., Comtesse, H., & Rosner, R. (2022). Predictors of prolonged grief disorder in a German representative population sample: Unexpectedness of bereavement contributes to grief severity and prolonged grief disorder. *Frontiers in Psychiatry, 13*, Article 853698. https://doi.org/10.3389/fpsyt.2022.853698

Eisma, M. C. (2023). Prolonged grief disorder in ICD-11 and DSM-5-TR: Challenges and controversies. *The Australian and New Zealand Journal of Psychiatry, 57*(7), 944–951. https://doi.org/10.1177/00048674231154206

Eisma, M. C., Janshen, A., & Lenferink, L. I. M. (2022). Content overlap analyses of ICD-11 and DSM-5 prolonged grief disorder and prior criteria-sets. *European Journal of Psychotraumatology, 13*(1), Article 2011691. https://doi.org/10.1080/20008198.2021.2011691

Eisma, M. C., Lenferink, L. I. M., Chow, A. Y. M., Chan, C. L. W., & Li, J. (2019). Complicated grief and post-traumatic stress symptom profiles in bereaved earthquake survivors: A latent class analysis. *European Journal of Psychotraumatology, 10*(1), Article 1558707. https://doi.org/10.1080/20008198.2018.1558707

Eisma, M. C., & Stroebe, M. S. (2021). Emotion regulatory strategies in complicated grief: A systematic review. *Behavior Therapy, 52*(1), 234–249. https://doi.org/10.1016/j.beth.2020.04.004

Fernández-Alcántara, M., Redondo-Armenteros, A., Pérez-Marfil, M. N., Cabañero-Martínez, M. J., Congost-Maestre, N., & Cruz-Quintana, F. (2025). Psychometric properties of the Spanish version of the Traumatic Grief Inventory Self Report Plus (TGI-SR+). *Psicothema, 37*(1), 81–90. https://doi.org/10.70478/psicothema.2025.37.09

Ferraris, V. A. (2019). Commentary: Should we rely on receiver operating characteristic curves? From submarines to medical tests, the answer is a definite maybe! *The Journal of Thoracic and Cardiovascular Surgery, 157*(6), 2354–2355. https://doi.org/10.1016/j.jtcvs.2018.10.100

Frankova, I., Bahmad, M. L., & Goloktionova, G. (2024). *Mental health and psychosocial support in Ukraine: Coping, help-seeking and health systems strengthening in times of war.* ARQ National Psychotrauma Center & VU Amsterdam. https://doi.org/10.13140/RG.2.2.33092.85123

Gilbody, S., Richards, D., Brealey, S., & Hewitt, C. (2007). Screening for depression in medical settings with the Patient Health Questionnaire (PHQ): A diagnostic meta-analysis. *Journal of General Internal Medicine, 22*(11), 1596–1602. https://doi.org/10.1007/s11606-007-0333-y

Government of the Netherlands. (2024, February). *Two years of war in Ukraine: A look back on the Netherlands' aid efforts.* https://www.government.nl/topics/russia-and-ukraine/dutch-aid-for-ukraine

Haneveld, J., Rosner, R., Vogel, A., Kersting, A., Rief, W., Steil, R., & Comtesse, H. (2022). Same name, same content? Evaluation of DSM-5-TR and ICD-11 prolonged grief criteria. *Journal of Consulting and Clinical Psychology, 90*(4), 303–313. https://doi.org/10.1037/ccp0000720

Hayes, A. F., & Coutts, J. J. (2020). Use Omega rather than Cronbach's Alpha for estimating reliability. But..... *Communication Methods and Measures, 14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Heeke, C., Franzen, M., Hofmann, H., Knaevelsrud, C., & Lenferink, L. I. M. (2022). A latent class analysis on symptoms of prolonged grief, post-traumatic stress, and depression following the loss of a loved one. *Frontiers in Psychiatry, 13*, Article 878773. https://doi.org/10.3389/fpsyt.2022.878773

Hyland, P., Vallières, F., Shevlin, M., Karatzias, T., Ben-Ezra, M., McElroy, E., Vang, M. L., Lorberg, B., & Martsenkovskyi, D. (2023). Psychological consequences of war in Ukraine: Assessing changes in mental health among Ukrainian parents. *Psychological Medicine, 53*(15), 7466–7468. https://doi.org/10.1017/S0033291723000818

IBM Corporation. (2017). *IBM SPSS statistics for windows* (Version 28.0) [Computer software]. https://www.ibm.com

International Trauma Consortium. (n.d.). *Research in Ukraine: Mental health resources for clinicians and researchers*. https://www.traumameasuresglobal.com/ukraine

Johnson, R. J., Antonaccio, O., Botchkovar, E., & Hobfoll, S. E. (2022). War trauma and PTSD in Ukraine's civilian population: Comparing urban-dwelling to internally displaced persons. *Social Psychiatry and Psychiatric Epidemiology, 57*(9), 1807–1816. https://doi.org/10.1007/s00127-021-02176-9

Killikelly, C., Comtesse, H., Lechner-Meichsner, F., Sam, J., & Ogrodniczuk, J. S. (2024). Research breakdowns: A constructive critique of research practice involving grief, trauma and displaced people. *Cambridge Prisms: Global Mental Health, 11*, Article e60. https://doi.org/10.1017/gmh.2024.60

Killikelly, C., Zhou, N., Merzhvynska, M., Stelzer, E., 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

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Kokou-Kpolou, C. K. (2021). Letter to the Editor: Prolonged grief disorder, posttraumatic stress disorder, and depression following traffic accidents among bereaved Balinese family members: Prevalence, latent classes and cultural correlates. *Journal of Affective Disorders, 295*, 1–2. https://doi.org/10.1016/j.jad.2021.08.007

Kokou-Kpolou, C. K., Lenferink, L. I. M., Brunnet, A. E., Park, S., Megalakaki, O., Boelen, P., & Cénat, J. M. (2022). The ICD-11 and DSM-5-TR prolonged grief criteria: Validation of the Traumatic Grief Inventory-Self Report Plus using exploratory factor analysis and item response theory. *Clinical Psychology & Psychotherapy, 29*(6), 1950–1962. https://doi.org/10.1002/cpp.2765

Kokou-Kpolou, C. K., Moukouta, C. S., Masson, J., Bernoussi, A., Cénat, J. M., & Bacqué, M.-F. (2020). Correlates of grief-related disorders and mental health outcomes among adult refugees exposed to trauma and bereavement: A systematic review and future research directions. *Journal of Affective Disorders, 267*, 171–184. https://doi.org/10.1016/j.jad.2020.02.026

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

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

Lenferink, L. I. M., Johnsen, I., Kristensen, P., Lie, N.-E., & Sveen, J. (2024). The Norwegian traumatic grief inventory-self report plus (TGI-SR+): A psychometric evaluation in traumatically bereaved people. *European Journal of Psychotraumatology, 15*(1), Article 2391248. https://doi.org/10.1080/20008066.2024.2391248

Lenferink, L. I. M., Van Dijk, I., Eisma, M. C., Eklund, R., Boelen, P. A., & Sveen, J. (2024). Psychometric evaluation of the Swedish Traumatic Grief Inventory Self-Report Plus (TGI-SR+) in bereaved parents. *Clinical Psychology & Psychotherapy, 31*(1), Article e2922. https://doi.org/10.1002/cpp.2922

Maccallum, F., & Bryant, R. A. (2018). Prolonged grief and attachment security: A latent class analysis. *Psychiatry Research, 268*, 297–302. https://doi.org/10.1016/j.psychres.2018.07.038

Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide*. Muthén & Muthén. https://www.statmodel.com/html_ug.shtml

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

Plan International. (2025). *Invisible wounds: Full report* [Report]. https://plan-international.org/uploads/sites/99/2025/03/Invisible-Wounds-full-report-English-2.pdf

Prigerson, H. G., Boelen, P. A., Xu, J., Smith, K. V., & Maciejewski, P. K. (2021). Validation of the new DSM-5-TR criteria for prolonged grief disorder and the PG-13-Revised (PG-13-R) scale. *World Psychiatry, 20*(1), 96–106. https://doi.org/10.1002/wps.20823

Prigerson, H. G., Singer, J., & Killikelly, C. (2024). Prolonged Grief Disorder: Addressing misconceptions with evidence. *The American Journal of Geriatric Psychiatry, 32*(5), 527–534. https://doi.org/10.1016/j.jagp.2023.10.020

Pritschet, L., Powell, D., & Horne, Z. (2016). Marginally significant effects as evidence for hypotheses. *Psychological Science, 27*(7), 1036–1042. https://doi.org/10.1177/0956797616645672

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

Roberts, B., Makhashvili, N., Javakhishvili, J., Karachevskyy, A., Kharchenko, N., Shpiker, M., & Richardson, E. (2019). Mental health care utilisation among internally displaced persons in Ukraine: Results from a nation-wide survey. *Epidemiology and Psychiatric Sciences, 28*(1), 100–111. https://doi.org/10.1017/S2045796017000385

Rosner, R., Comtesse, H., Vogel, A., & Doering, B. K. (2021). Prevalence of prolonged grief disorder. *Journal of Affective Disorders, 287*, 301–307. https://doi.org/10.1016/j.jad.2021.03.058

Schaal, S., Dusingizemungu, J., Jacob, N., Neuner, F., & Elbert, T. (2012). Associations between prolonged grief disorder, depression, posttraumatic stress disorder, and anxiety in Rwandan genocide survivors. *Death Studies, 36*(2), 97–117. https://doi.org/10.1080/07481187.2011.573177

Schisterman, E. F., Perkins, N. J., Liu, A., & Bondell, H. (2005). Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology, 16*(1), 73–81. https://doi.org/10.1097/01.ede.0000147512.81966.ba

Seiler, A., Von Känel, R., & Slavich, G. M. (2020). The psychobiology of bereavement and health: A conceptual review from the perspective of social signal transduction theory of depression. *Frontiers in Psychiatry, 11*, Article 565239. https://doi.org/10.3389/fpsyt.2020.565239

Shevlin, M., Hyland, P., Karatzias, T., Makhashvili, N., Javakhishvili, J., & Roberts, B. (2022). The Ukraine crisis: Mental health resources for clinicians and researchers. *Journal of Traumatic Stress, 35*(3), 775–777. https://doi.org/10.1002/jts.22837

Steel, Z., Chey, T., Silove, D., Marnane, C., Bryant, R. A., & Van Ommeren, M. (2009). Association of torture and other potentially traumatic events with mental health outcomes among populations exposed to mass conflict and displacement. *Journal of the American Medical Association, 302*(5), 537–549. https://doi.org/10.1001/jama.2009.1132

Tang, S., Chen, Z., Boelen, P. A., Eisma, M. C., & Lenferink, L. I. M. (2024). Psychometric evaluation of the Traumatic Grief Inventory-Self Report Plus (TGI-SR+) in Chinese bereaved people. *Clinical Psychology & Psychotherapy, 31*(5), Article e3066. https://doi.org/10.1002/cpp.3066

Treml, J., Linde, K., Brähler, E., & Kersting, A. (2024). Prolonged grief disorder in ICD-11 and DSM-5-TR: Differences in prevalence and diagnostic criteria. *Frontiers in Psychiatry, 15*, Article 1266132. https://doi.org/10.3389/fpsyt.2024.1266132

Treml, J., Schmidt, V., Braehler, E., Morfeld, M., & Kersting, A. (2024). Psychometric properties of the German version of the Traumatic Grief Inventory-Self Report Plus (TGI-SR+). *European Journal of Psychotraumatology, 15*(1), Article 2421706. https://doi.org/10.1080/20008066.2024.2421706

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best alternatives to Cronbach's Alpha reliability in realistic conditions: Congeneric and asymmetrical measurements. *Frontiers in Psychology, 7*, Article 769. https://doi.org/10.3389/fpsyg.2016.00769

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)*. Boston, MA, USA: National Center for PTSD.

World Health Organization. (2025, January). *ICD-11 for Mortality and Morbidity Statistics: 6B42 Prolonged grief disorder.* https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1183832314

EACLIPT

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

zpid Leibniz-Institut für Psychologie

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

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

Repeated Mirror Exposure in Individuals With Body Dysmorphic Symptoms

Katrin Schoenenberg [1], Alexandra Martin [2]

[1] *Department for Clinical Psychology and Psychotherapy, Kiel University, Kiel, Germany.* **[2]** *Department for Clinical Psychology and Psychotherapy, University of Wuppertal, Wuppertal, Germany.*

Clinical Psychology in Europe, 2026, Vol. 8(2), Article e19097, https://doi.org/10.32872/cpe.19097

Received: 2025-07-30 • **Accepted:** 2026-03-02 • **Published (VoR):** 2026-05-29

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

Corresponding Author: Katrin Schoenenberg, Christian-Albrechts-Platz 4, 24118 Kiel, Germany. Phone: +49 431 880-3617. E-mail: schoenenberg@psychologie.uni-kiel.de

Supplementary Materials: Materials [see Index of Supplementary Materials]

Abstract

Background: Mirror exposure represents a common component in treatment of body dysmorphic disorder (BDD). However, the benefits of repeated mirror exposure have not been investigated as a standalone intervention for BDD. This study aimed to examine the cognitive and affective response to mirror exposure in individuals with high and low levels of body dysmorphic symptoms.

Method: Fifty women participated in two guided full-body mirror exposures (approx. 32 minutes each). Participants were divided into two groups based on the severity of their BDD symptoms. Twenty-three participants reported elevated, not primarily weight related, body dysmorphic symptoms. Body satisfaction and affective responses were assessed before and after the exposure, affective responses were further assessed during the exposure. Post-event processing related to the experience was rated the day afterwards.

Results: Participants with body dysmorphic symptoms reported lower state body satisfaction and higher shame in both sessions, sadness was elevated in the first session only. State body satisfaction dropped from pre to post exposure but improved from the first to the second session. Negative affects did not decrease within but between the two sessions. Post-event processing after the first exposure predicted negative affect at the beginning of the second session.

Conclusion: The results support a positive effect of repeated mirror exposure across sessions, without improvement within the session. They point towards the detrimental role of mental post-processing.

Keywords

body dysmorphic disorder, body image exposure, body satisfaction, shame, post-event processing

Highlights

- Examining the effect of repeated mirror exposure in BDD-symptomatic women.
- State body satisfaction improves between sessions, but is impaired post exposure.
- Women with BDD symptoms experienced more pronounced feelings of shame.
- Negative affects and post-event-processing improved across sessions.

Body Dysmorphic Disorder (BDD) is characterized by excessive preoccupation with one or more subjective flaws in appearance and associated distress or impairment in daily life (American Psychiatric Association, 2013). According to cognitive-behavioral BDD models (Veale & Neziroglu, 2010; Wilhelm et al., 2013) and experimental studies (Barnier & Collison, 2019; Kollei & Martin, 2014; Waldorf et al., 2019), looking into a mirror elicits dysfunctional processing of one's own appearance. As in other body image disorders, mirror checking and mirror avoidance are common safety behaviors in BDD (Naumann et al., 2022). Mirror checking serves as a behavioral response to uncertainty about appearance and is motivated by the hope to be less dissatisfied with the appearance, at the same time it increases selective attention to the disliked body part(s) (Veale & Riley, 2001). Additionally, it is not uncommon for those affected to also avoid, either completely or partially, situations in which they are confronted with their own appearance and thus with the associated aversive feelings. This also strengthens dysfunctional processing of the self and thus selective attention (Veale & Neziroglu, 2010). Selective attention in BDD refers to threatening stimuli associated with rejection or judgment (Johnson et al., 2018), as well as disorder-related stimuli such as body parts.

Kollei and Martin (2014) found that individuals with BDD expressed more body-related and more negative cognitions than participants who were either in good health or suffering from a major depression when confronted with their bodies in the mirror. On the day after the exposure, participants with BDD reported more post-event processing than controls. The results on post-event processing were supported by Schoenenberg and Martin (2022) in people with elevated BDD concern. Neziroglu and colleagues (2010) investigated mirror gazing in BDD and healthy participants at one-minute intervals. The levels of fear and disgust were higher in the BDD group, but decreased significantly over time.

Satisfaction with one's own body or appearance represents the cognitive-evaluative component of body image (Cash, 2012). In a study by Waldorf et al. (2019), men with the BDD subtype muscle dysmorphia had a generally lower level of state body satisfaction and higher negative affect scores than weight-trained or non-weight-trained controls. When confronted with images of their own body, participants in all the groups respon-

ded with a lower level of body satisfaction, but negative affect increased only in the muscle dysmorphia group. No study has investigated state body satisfaction in women with elevated BDD symptoms in the context of mirror exposure so far. For this reason, we aimed to assess state body satisfaction pre and post isolated mirror exposure. We hypothesized the following:

Hypothesis I (BDD group effect): Individuals with distinct BDD symptoms express overall lower state body satisfaction (H1a) and stronger negative affect (H1b) in the context of mirror exposure than individuals with low or no BDD symptoms. Furthermore, individuals with distinct BDD symptoms experience more post-event processing (H1c) about the exposure the following day than a group with low or no BDD symptoms.

In the treatment of BDD, mirror exposure serves to correct attentional and evaluative biases (Veale & Neziroglu, 2010; Wilhelm et al., 2013). Typically, individuals are asked to describe their body in, e.g., a neutral, or non-judgmental way when confronted with their body image in the mirror (Griffen et al., 2018; Schoenenberg & Martin, 2022). Compared to short mirror viewing, it is common to guide the attention to different body parts during mirror exposure to counteract poor global processing and to train finding a less judgmental way of describing the body. It takes longer than short viewing or checking and follows a different approach which is why some manuals prefer the term "mirror retraining" (Wilhelm et al., 2013). Even though mirror exposure or retraining is part of cognitive behavioral treatment manuals, no study has investigated its isolated effect on individuals with BDD (Griffen et al., 2018). Due to the experimental character of the present study the term "mirror exposure" will be used in the following even though we did apply a typical retaining rationale including guidance through the body parts, with instructions to observe upcoming thoughts and emotions and to use a neutral description of the perceived body parts.

Reviews of research on eating disorders summarized the positive effects of repeated video and mirror exposure (Butler & Heimberg, 2020). The few studies on body image exposure in samples with weight-related body dissatisfaction examining within and between sessions effects in detail tend to support between-session effects (Díaz-Ferrer et al., 2017; Moreno-Domínguez et al., 2012). To our knowledge, the effects of mirror exposure within and between sessions with respect to body image evaluations and affective responses in persons with distinct BDD symptoms have not yet been studied. Therefore, we tested the following hypotheses:

Hypothesis II (within-session time effect): At the end of a mirror exposure session, state body satisfaction is higher (H2a) and negative affect is lower (H2b) in comparison to at the beginning of the session.

Hypothesis III (between-session effect): In a second mirror exposure session, body satisfaction is higher (H3a); negative affect is lower (H3b) and post-event processing is lower (H3c) compared to in the first session.

As a third goal, this study aimed to uncover the potentially destructive role of post-event processing in maintaining the disorder. This mechanism is well-known and well-studied for social anxiety (e.g., Fehm et al., 2008; Rachman et al., 2000). It refers to a negatively biased way of reprocessing a previous difficult event and is considered a special form of rumination. Negative distorted post-processing could lead to change in the memorized experience and could interfere with the correction of one's own expectations. If post-event processing contributes to maintenance of appearance-related concerns, stronger post-event processing after a first mirror exposure should predict a poorer body image at the onset of the second session. This could also be of therapeutic relevance, as the barrier to so far avoided mirror exposure could be higher with more pronounced post-event processing and could possibly even risk dropout. Accordingly, we investigated the following hypothesis:

Hypothesis IV (model prediction): Post-event processing about the first mirror exposure predicts lower state body satisfaction (H4a) and higher negative affect (H4b) at the beginning of the second exposure.

To test these hypotheses, we invited individuals with body dysmorphic symptoms and a group without (or with only a low level) of BDD symptoms to participate in the study. All participants took part in two guided mirror exposure sessions separated by a one-week interval. The exposure procedure was standardized and similar to mirror exposure in BDD treatment. No additional intervention components were provided to the study participants.

Method

Design

The experimental study had a 2 (BDD *group*: positive, negative; between factor) x 2 to 4 (*time*: pre, full-body exposure beginning, full-body exposure end, post; within-factor) x 2 (*session*: first, second; within-factor) mixed design.

Participants

Participants were included if they were female and between 18 and 55 years old. Based on a screening, participants who were allocated to the BDD positive group were additionally checked by asking them to name their primary area of concern. If weight was their primary area of concern, they were excluded from participation. Males were excluded from participation because the exposure in underwear to the female examiner could have been a confounding factor. Participants who reported suicidal tendencies were furthermore excluded from participation.

The fifty female participants were on average 22.8 years old (SD = 4.5). As the highest education level, 82% reported a high-school degree and 18% a university degree. Most

participants were living in a relationship (50%) or were single (44%), and some were married (6%). The sample had a body mass index (BMI) in the normal range (M = 21.89, SD = 2.95). Participants with elevated BDD-symptoms were most dissatisfied with their nose (21.7%), belly (17.4%), legs (13.0%), teeth (8.7%) and hands (8.7%). Participants with little BDD-symptoms reported to be most dissatisfied with their belly (22.2%), breast (14.8%), nose (14.8%), body shape (7.4%), legs (7.4%) and teeth (7.4%). All other body parts were named only once in each of the two groups.

Materials

Sample Characteristics and Manipulation Check

The participants indicated ***demographic questions*** regarding their age, gender, relationship status, and highest educational attainment.

The ***Dysmorphic Concern Questionnaire*** (DCQ; Oosthuizen et al., 1998) is a well-established screening instrument for BDD. Its seven items assess the extent of body dysmorphic concerns on 4-point Likert scales (sum range 0 to 21). Different studies have confirmed its good level of reliability and validity (e.g., Schieber et al., 2018; Stangier et al., 2003; α = 0.88 in the current study). In student samples, a cut-off value of nine has proven to be the best balance between sensitivity and specificity (Mancuso et al., 2010). For this reason, participants were assigned to the BDD positive *group* (BDD+; N_{BDD+} = 23) if they scored nine or higher and to the BDD negative *group* (BDD-; N_{BDD-} = 27) if they scored eight or lower.

To assess BDD criteria (DSM-5), a ***self-rated BDD diagnosis*** scale that had previously been used in epidemiologic studies was employed (Schieber et al., 2015). If all criteria were answered with 'yes' and if participants indicated spending at least one hour a day thinking about their appearance or taking actions related to their appearance, and primary weight concerns were excluded, the participants were considered to have a self-rated BDD diagnosis (N = 11). We also recorded the ***body part*** that participants were most dissatisfied with via a self-report question.

Further assessments calculated the appearance ***distortion*** by the difference between the examiners' and the participants' ratings on an 11-point scale from 0 "not at all affected in appearance" to 10 "very strongly affected in appearance" (Stangier et al., 2000). Verbal anchors on the scale supported the examiners' assessment (0: not impaired in appearance at all, 2: slightly impaired in appearance, 5: impaired in appearance; 7: severely impaired in appearance, 10: very severely impaired in appearance).

The ***Patient Health Questionnaire Depression Module*** (PHQ-9; Kroenke et al., 2001) measures depression severity with nine items. The items scale ranges from 0 "not at all" to 3 "nearly every day." Its internal consistency was excellent in the original validation study (α = 0.86 and 0.89; Kroenke et al., 2001) and acceptable in the present study (α = 0.78).

In a short, ***structured interview*** after the exposure, the ability and motivation to follow instructions during the exposure and the stress induced by the exposure were assessed to verify the intended manipulation using 11-point Likert scales from 0 "not at all" to 10 "very much" (Kollei & Martin, 2014; see Supplementary Materials, Table 1).

Dependent Variables

The ***Body Image State Scale*** (BISS; Cash et al., 2002) assesses body image satisfaction. The mean score of the six items (scale from 1 to 9) is typically used. The BISS showed an acceptable to good internal consistency in prior studies with women samples (α = 0.77 to 0.90; Cash et al., 2002; Vocks et al., 2009) as well as in this study (α = 0.79 to 0.89).

The intensity of the four ***negative affects*** tension, shame, sadness, and anxiety were assessed on visual analogue scales ranging from 0 "none" to 100 "strong". The examiner verbally asked the subjects to rate their affect intensity before, during, and after exposure.

The ***Post-Event Questionnaire*** (PEP-Q; Fehm et al., 2008; Rachman et al., 2000) assesses repetitive negative thinking about a past difficult event. We used an adapted version for mirror exposure (Kollei & Martin, 2014) that was administered 24 hours after the exposure. The participants completed the questionnaire online via the LimeSurvey tool. Sixteen items (scale from 0 "none, never, not at all" to 100 "very strong, always") addressed processing about mirror exposure in the past 24 hours. This version showed excellent internal consistency both in prior studies (α = 0.95, Kollei & Martin, 2014; α = 0.89 to 0.93, Schoenenberg & Martin, 2022) and in this study (α = 0.95).

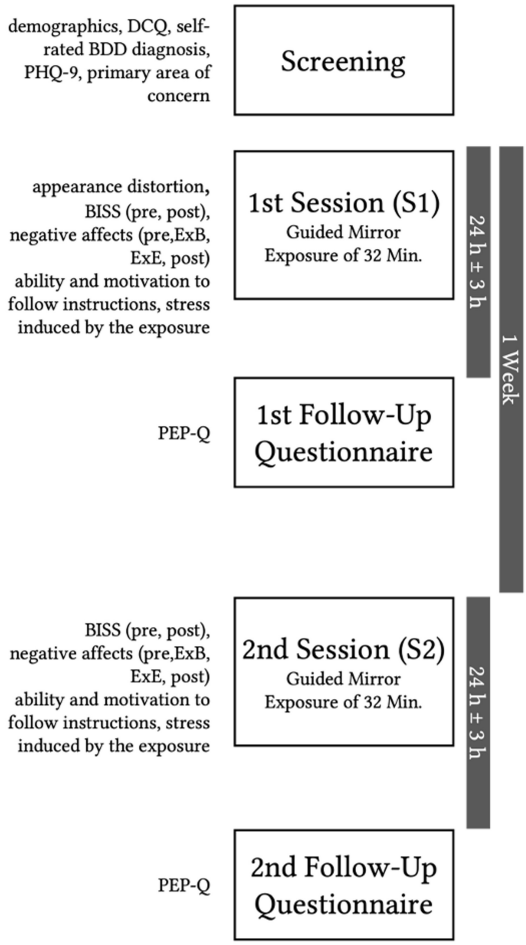
Procedure

The ethics committee of the University of Wuppertal approved the study (MS/BBL 190326 Schoenenberg). Participants were recruited via flyers, online postings, and mailing lists for a study on repeated mirror viewing, with no indications in the advertisement about potentially beneficial effects of the intervention. The participants received detailed information about the study procedure, data protection, and their rights as participants prior to inclusion in the study, and provided informed consent.

Prior to the laboratory appointment, participants responded to an online screening questionnaire (Figure 1) that checked the inclusion and exclusion criteria. Those eligible to participate were invited to attend a first mirror exposure session in the laboratory of the University of Wuppertal. Here, they were again informed about the procedure, and they then responded to a first questionnaire including the first rating of state body satisfaction. The examiner introduced the participant to the task during the mirror exposure, which involved describing all named body parts in detail and in a neutral way. Experience of negative affect and cognitions were permitted and registered but the focus was on describing features of the body or body part in a neutral manner.

Figure 1

Procedure of Repeated Mirror Exposure Study



Note. DCQ = Dysmorphic Concern Questionnaire; PHQ-9 = Patient Health Questionnaire Depression Module; BISS = Body Image State Scale; PEP-Q = Post-Event Questionnaire; Pre = pre-exposure; ExB = full-body view beginning; ExE = full-body view end; Post = post-exposure.

Sample questions were given to participants, e.g., "What is the color/size/texture of this body part?". A sample text with a description of eyes was provided, and participants were asked to describe their thumb to practice the procedure. Afterwards, they undressed to their underwear, which was meant to have a neutral color and design. We chose underwear as clothing to ensure participants could hide almost none of their

potential flaws with their clothes. Psychophysiological measurement equipment for assessing respiration, heart rate, and electrodermal activity was attached to their back with a belt.

Before exposure, the participants were asked to position themselves at a mark 50 cm away from a full-body mirror (height: 175 cm; width: 69 cm; adjustable wings) with their backs to the mirror for two minutes. They were then asked to provide their first affect rating (pre). Afterwards an audio recording with instructions guiding participants through the exposure was played. At first, the participants were asked to look at their entire body for one minute. This was followed by the second affect rating (ExB). The examiner left the room, and the recording directed the subjects' attention to the different body parts of the head, upper and lower torso, and the back view of the body. The female voice prompted them to describe the respective body parts. The participants described their bodies internally to reduce shame and artefacts in psychophysiological measurement. The recording ended with a second full-body view from the front. The recording from the first full body view until this last full body view lasted approximately 32 minutes. The examiner entered the room and assessed affect for the third time (ExE). After completing the exposure task, the participants turned their back to the mirror and did not examine their body in the mirror for two minutes. Affect was then recorded for the fourth time (post). Afterwards, the participants rated their state body satisfaction. The follow-up online questionnaire including the assessment of post-event processing had to be answered 24 hours after the exposure (+/- 3 hours; Figure 1). One week later, the procedure was repeated in the same way and by the same examiner. Students received credit points for participation.

Statistical Analyses

To answer Hypotheses I-III, we conducted mixed analyses of variances (mixed ANOVAs) with two within-subject factors. The first factor *time* had two to four steps depending on the variable: pre-exposure (pre), full-body view at the beginning of exposure (ExB), full-body view at the end of exposure (ExE), and post-exposure (post). The second factor *session* had two steps: first exposure session (S1) and second exposure session (S2). The between-subject factor BDD-*group* was divided into two categories: BDD+ and BDD-. If the assumption of sphericity was not met, we used Greenhouse-Geisser correction. In case of significant effects, post-hoc tests using Bonferroni correction were conducted. To investigate Hypothesis IV, univariate linear regressions were performed. In subsequent sensitivity analyses, the mixed ANOVAs were repeated with a different group factor, categorizing the individuals according to their self-rated BDD diagnosis (DSM-5).

We calculated the sample size with G*Power 3.1 (Faul et al., 2007). For the mixed ANOVAs, we assumed a medium to large effect size for the *group* factor (Kollei & Martin, 2014; Schoenenberg & Martin, 2022; parameters: $f = 0.35$, $r = 0.50$, $1\text{-}\beta = 0.80$, $\alpha = 0.05$) resulting in a sample size of $n = 44$. For the *time* and *session* factors a small to medium

effect was expected (Schoenenberg & Martin, 2022; parameters: $f = 0.17$, $r = 0.70$, 1-β = 0.80, α = 0.05) resulting in a sample size of $n = 44$. For the prediction model a medium effect was assumed (parameters: $f^2 = 0.15$, $r = 0.70$, 1-β = 0.80, α = 0.05) resulting in a sample size of $n = 43$.

Fifty-five people participated in the first session. Five participants (3 BDD+, 2 BDD-), did not participate in the second session and were thus excluded from the sample. One participant (BDD-) reported circulation problems towards the end of the second session and therefore did not provide affect ratings for the measurement of full-body end view (ExE) and post-exposure (Post). Her ratings were constant for the pre-exposure and full-body beginning assessment and were applied for the other two measurements. Hence, a sample of $N = 50$ participants was available for the analysis. The analyses of the PEP-Q after the second session included 45 participants, because five individuals did not respond to this last assessment (3 BDD-, 2 BDD+ cases).

The protocol initially included the assessment of psychophysiological data. Due to severe movement artefacts in the data, we were not able to analyze it.

Results

Sample Characteristics

The group with BDD+ did not differ from the BDD- group with regard to age (Table 1), highest education degree, $\chi^2(1) = 0.01$, or relationship status, $\chi^2(2) = 1.02$. The mean appearance distortion in the BDD+ group was significantly higher than in the BDD- group. The BDD+ group rated their appearance impairment on average 3.43 points, $SD = 2.33$, lower than the examiners in comparison to the BDD- group, $M = 1.33$, $SD = 1.52$. Individuals with higher BDD symptoms reported higher depressiveness, and a lower BMI (Table 1). All 11 cases that fulfilled the self-rated BDD diagnosis were in the in the BDD+ group.

Manipulation Check

Both groups were able to follow the mirror exposure instructions to the same extent and were equally motivated to follow them in the first and second sessions (Table 1). Ratings of stress induced by the exposure were significantly higher in the BDD+ group for both sessions (Table 1). Stress induced by exposure dropped from the first to the second exposure by approximately 1.7 points in the BDD+ group and by approximately only 0.5 in the BDD- group.

Table 1

Sample Characteristics

Variable	BDD+	BDD-		Total Sample
	M (*SD*)	*M* (*SD*)	*t*(*df*)	*M* (*SD*)
Age	22.30 (2.82)	23.19 (5.55)	0.69 (48)	22.78 (4.49)
BMI	20.81 (1.84)	22.81 (3.41)	2.51 (48)*	21.89 (2.95)
DCQ Sum Score	12.35 (2.93)	4.63 (1.84)	-10.92 (35.86)**	8.18 (4.55)
Appearance Distortion	3.43 (2.33)	1.33 (1.52)	-3.83 (48)**	2.30 (2.19)
PHQ-9 Sum Score	11.17 (4.42)	6.52 (3.81)	-4.00 (48)**	8.66 (4.68)
Ability to Follow Instruction				
First exposure	9.00 (1.78)	9.48 (0.94)	1.22 (48)	9.26 (1.40)
Second exposure	9.00 (1.65)	9.26 (1.20)	0.64 (48)	9.14 (1.41)
Motivation to Follow Instructions				
First exposure	8.04 (1.99)	7.89 (1.40)	-0.32 (48)	7.96 (1.68)
Second exposure	8.09 (1.73)	7.85 (1.41)	-0.53 (48)	7.96 (1.55)
Stress Induced by Exposure				
First exposure	5.26 (2.93)	2.63 (2.31)	-3.36 (48)**	3.84 (2.90)
Second exposure	3.57 (2.19)	2.11 (1.89)	-2.52 (48)**	2.78 (2.14)
	n	***n***	χ^2	***n* (%)**
Education			$\chi^2(1) = 0.01$	
A-levels	22	19		41
University Degree	5	4		9
Relationship Status			$\chi^2(2) = 1.02$	
Single	11	11		22
Relationship	15	10		25
Married	1	2		3

Note. $N = 50$; $n_{BDD+} = 23$; $n_{BDD-} = 27$; BMI = Body Mass Index; BDD = Body Dysmorphic Disorder; BDD+ = with BDD symptoms; BDD- = without BDD symptoms.
*$p < .05$. **$p < .01$.

Hypotheses

BDD Group Effect

Differences between the BDD groups were examined in all analysis using the main effects of group, and time-by-group and session-by-group interaction effects (Table 2). People with distinct BDD symptoms (BDD+ group) showed a significantly lower BISS score, as well as higher levels of shame and sadness during the exposure sessions (Table 3). The session-by-group interaction effect was significant for sadness (Table 2). Post-hoc tests indicated significantly higher sadness for the BDD+ than the BDD- group in the first session but not in the second one (Table 3). Reported tension and anxiety did not differ significantly between the groups. In addition, the degree of post-event processing was significantly higher in the BDD+ than in the BDD- group.

Table 2

Results for Repeated Measures Analysis of Variances

Variable/Factor	*F*	*df*	*p*	η_p^2
BISS				
Group	6.43	1, 48	.015	.12
Time	14.10	1, 48	< .001	.23
Session	6.02	1, 48	.018	.11
Time x Group	2.15	1, 48	.149	.04
Session x Group	0.19	1, 48	.668	.01
Time x Session	0.10	1, 48	.754	.01
Session x Time x Group	2.34	1, 48	.132	.05
Tension				
Group	1.08	1, 48	.305	.02
Time	14.94	1.93, 92.74	< .001	.24
Session	34.90	1, 48	< .001	.42
Time x Group	1.20	1.93, 92.74	.304	.02
Session x Group	3.57	1, 48	.065	.07
Time x Session	3.11	2.11, 101. 24	.046	.06
Session x Time x Group	0.02	2.11, 101. 24	.982	.00
Shame				
Group	4.67	1, 48	.036	.09
Time	9.24	2.36, 113.45	< .001	.16
Session	29.52	1, 48	< .001	.38
Time x Group	2.17	2.36, 113.45	.110	.04
Session x Group	2.54	1, 48	.118	.05
Time x Session	3.71	1.82, 87.30	.032	.07
Session x Time x Group	0.31	1.82, 87.30	.716	.01

Variable/Factor	*F*	*df*	*p*	η_p^2
Sadness				
Group	4.52	1, 48	.039	.09
Time	4.93	2.20, 105.69	.007	.09
Session	16.01	1, 48	< .001	.25
Time x Group	1.45	2.20, 105.69	.239	.03
Session x Group	6.07	1, 48	.017	.11
Time x Session	3.07	3, 144	.030	.06
Session x Time x Group	0.71	3, 144	.550	.02
Anxiety				
Group	0.90	1, 48	.348	.02
Time	4.51	2.27, 108.93	.010	.09
Session	7.34	1, 48	.009	.13
Time x Group	1.24	2.27, 108.93	.296	.03
Session x Group	4.02	1, 48	.051	.08
Time x Session	2.06	2.17, 103.96	.129	.04
Session x Time x Group	1.56	2.17, 103.96	.213	.03
PEP-Q				
Group	16.53	1, 43	< .001	.28
Session	13.67	1, 43	.001	.24
Session x Group	0.06	1, 43	.808	.01

Note. $N = 50$; with BDD diagnosis $n = 11$; BDD = Body Dysmorphic Disorder; BISS = Body Image State Scale; PEP-Q = Post-Event Processing Questionnaire.

Within-Session Time Effect

The effect of time was examined in all analysis using the main effects of time (Table 2). The BISS score dropped significantly over time within one session (Figure 2). The post-hoc tests confirmed a significant drop in the first, $M_{Diff} = 0.48$, 95% CI [0.13, 0.84], $p = .008$, and in the second sessions, $M_{Diff} = 0.42$, 95% CI [0.14, 0.70], $p = .004$. Overall, the time effects for tension, shame, sadness, and anxiety were significant (Table 2). The post-hoc tests comparing the "full-body view at the beginning" to the "full body view at the end" of each session did not yield significant effects for any affect variable but a significant drops in tension, shame and anxiety "post" the exposure could be detected (Supplementary Materials, Table 2).

Between Session Effect

The effect between session was examined in all analysis using the main effects of session (Table 2). The BISS score increased significantly between sessions (Figure 2). Tension, shame, sadness, and anxiety decreased significantly between sessions. The PEP-Q score decreased significantly between sessions (Table 2).

Table 3

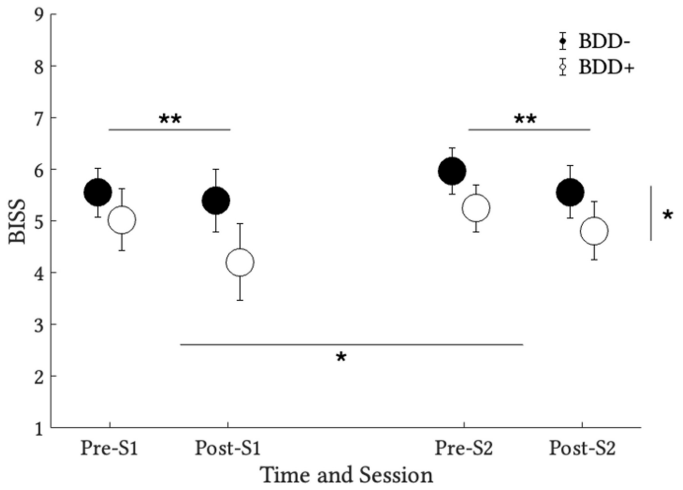
Means and Standard Deviation for Affects and Post-Event Processing

	Session 1		Session 2	
	BDD+	BDD-	BDD+	BDD-
Variable	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**	***M* (*SD*)**
Tension				
Pre	35.22 (4.70)	28.63 (3.45)	16.52 (3.06)	18.41 (3.30)
ExB	40.87 (5.37)	29.81 (3.53)	19.78 (3.42)	18.15 (3.16)
ExE	36.87 (5.80)	25.74 (4.03)	17.61 (3.80)	16.63 (3.36)
Post	28.13 (5.33)	19.78 (3.68)	13.26 (3.45)	13.70 (3.16)
Shame				
Pre	33.30 (6.32)	20.37 (4.00)	13.26 (2.91)	9.33 (2.67)
ExB	38.26 (6.88)	21.11 (3.93)	18.57 (3.51)	10.44 (2.70)
ExE	35.87 (6.61)	19.00 (4.25)	21.26 (4.35)	10.74 (2.98)
Post	28.70 (5.96)	15.00 (3.46)	16.04 (3.55)	8.33 (2.85)
Sadness				
Pre	10.43 (3.58)	3.52 (2.32)	2.22 (1.08)	1.85 (1.20)
ExB	17.04 (4.91)	5.37 (2.40)	5.96 (2.41)	1.85 (1.20)
ExE	19.13 (4.71)	6.48 (2.60)	5.87 (2.39)	3.33 (1.62)
Post	13.70 (3.89)	4.07 (1.73)	5.00 (2.09)	2.59 (1.50)
Anxiety				
Pre	14.22 (4.67)	9.07 (3.79)	1.48 (1.56)	5.56 (2.12)
ExB	15.22 (5.02)	4.41 (1.93)	3.78 (1.96)	5.00 (2.07)
ExE	10.22 (3.67)	5.19 (2.58)	3.30 (1.79)	3.07 (1.49)
Post	9.13 (3.75)	3.33 (2.07)	3.09 (1.74)	2.96 (1.49)
PEP-Q				
Post	34.40 (21.37)	16.52 (14.80)	27.39 (19.35)	8.89 (8.07)

Note. n_{BDD+} = 27; n_{BDD-} = 23; Pre = pre-exposure; ExB = full-body view beginning; ExE = full-body view end; Post = post-exposure; BDD+ = with BDD symptoms; BDD- = without BDD symptoms; BDD = Body Dysmorphic Disorder.

Figure 2

Means and 95% CIs for Body Image Satisfaction



Note. $N = 50$; Pre = pre-exposure; Post = post-exposure; S1 = first exposure session; S2 = second exposure session; BISS = Body Image State Scale; BDD+ = with BDD symptoms; BDD- = without BDD symptoms. $^{*}p < .05$. $^{**}p < .01$.

Model Prediction

The PEP-Q score after the first exposure session was tested as a predictor for the BISS and affect scores at the beginning of the second session. The PEP-Q after the first session positively predicted tension, shame, sadness, and anxiety at the beginning of the "full-body view" in the second session (Table 4). It also significantly predicted shame at the onset of the second session. PEP-Q did not significantly predict BISS prior to the second exposure session.

Sensitivity Analysis

When the groups were categorized according to the self-rated diagnostic criteria, the group main effects were confirmed for the BISS scores, shame, and PEP-Q-scores (Supplementary Materials, Table 3). The main effects for time and session were significant for the BISS score, and all affects; the session effect was significant for the PEP-Q score (Supplementary Materials, Table 3).

Table 4

Univariate Linear Regression Predicting Body Satisfaction and Affects at the Beginning of the Second Session by Post-Event Processing After the First Session

Predicted Variable	Time 2nd Session	R^2	*F*	*B*	*SE*	β	*t*	*p*	*B* [95% CI] *LL*	*UL*
BISS	Pre	0.04	1.90	-0.01	0.01	-0.20	-1.38	.174	-0.03	0.01
Tension										
	Pre	0.06	3.14	0.20	0.11	0.25	1.77	.083	-0.03	0.42
	ExB	0.25	15.58	0.56	0.14	0.50	3.95**	< .001	0.27	0.84
Shame										
	Pre	0.08	4.20	0.20	0.10	0.28	2.05*	.046	0.01	0.40
	ExB	0.22	13.29	0.65	0.18	0.47	3.65**	.001	0.29	1.01
Sadness										
	Pre	0.01	0.29	0.02	0.04	0.08	0.54	.593	-0.06	0.10
	ExB	0.27	17.53	0.49	0.12	0.52	4.19**	< .001	0.26	0.73
Anxiety										
	Pre	0.01	0.36	0.04	0.07	0.09	0.60	.551	-0.10	0.18
	ExB	0.13	7.31	0.34	0.12	0.36	2.70*	.009	0.09	0.59

Note. $N = 49$; $df = (1, 48)$; BISS = Body Image State Scale; Pre = prior to exposure; ExB = full-body view beginning.
*$p < .05$. **$p < .01$.

Discussion

The present study was the first to investigate repeated guided mirror exposure in BDD-symptomatic women as a standalone intervention. We aimed to examine effects within one session and between two sessions of mirror exposure.

As assumed, body image exposure was related to a more negative cognitive evaluation, here, state body satisfaction and dysfunctional rumination about the exposure, which is called post-event processing, in individuals with higher BDD symptoms than in those with fewer or no symptoms. Descriptively, both groups experienced tension and shame with the highest intensity, while the levels of anxiety and sadness were overall lower. A comparison of the two groups showed that shame and sadness were elevated only in individuals with distinct BDD symptoms. Sadness was found to be higher for individuals with BDD in the first exposure session only. Unfortunately, we did not capture the level of disgust in this study. Earlier studies have identified the experience of disgust to be relevant during mirror exposure (Kollei & Martin, 2014; Neziroglu et al., 2010).

The results on affects during exposure are in line with the findings of previous studies which found sadness to be higher in individuals with BDD (Kollei & Martin, 2014; Schoenenberg & Martin, 2022) and reported higher levels of shame in BDD positive participants (Schoenenberg & Martin, 2022). Overall, our results indicate that anxiety is not the main affect in this body confrontation paradigm. Shame has been discussed as a key affect in BDD in previous work (Malcolm et al., 2021). Addressing and changing the experience of shame during mirror exposure may be a crucial mechanism that needs further investigation.

The second research question addressed changes over time within one exposure session. State body satisfaction did not increase, but decreased significantly from pre to post in each session. This implies that state body satisfaction deteriorated with the exposure interventions, even though we applied a relatively long exposure time and a neutral description of body parts. Consistent with this, a previous study found increased or unchanged discomfort with one's body in women who were dissatisfied with their body at the end of a guided mirror exposure (Díaz-Ferrer et al., 2017; Moreno-Domínguez et al., 2012). Affects did not change significantly from the first full-body view at the beginning of the session (ExB) to the full-body view at the end (ExE). It is possible that the participants still felt affected at this point due to the exposure with a difficult stimulus. The intensity of negative affect was significantly lower when exposure was terminated, but negative affects do not seem to decrease during this type of guided exposure. A clear habituation approach without directing attention to the various parts of the body may have led to a greater change in affect from the beginning to the end of the intervention, as it has been found in a study by Díaz-Ferrer et al. (2017) who investigated pure exposure. However, due to the attentional bias associated with body image disorders (Johnson et al., 2018), such an approach would tend to promote a dysfunctional approach to self-image (Veale & Riley, 2001). We did not assess affect while the subjects were viewing their most aversive body part(s). Negative affect might have been increased during this period. In addition, the affect ratings were communicated to the examiner verbally, which may have had an impact on their extend, particularly in the case of shame.

Nevertheless, shame seemed to be more relevant than anxiety during the mirror viewing, as pointed out earlier. The experience of shame is closely related to self-evaluation or evaluation by others (Goss & Allan, 2009). Due to this relationship, changes in shame experience may follow from cognitive changes (e. g. the realization after a mirror exposure that one did not hate the whole body). Consequently, strong affect changes within one session may not be the decisive mechanism for successful mirror exposure and long-term changes in shame experience.

The comparison of the exposure sessions revealed that state body satisfaction was significantly higher and negative affects were significantly lower in the second session, indicating the positive effect of the repeated exposure intervention. An increase of body

satisfaction between sessions was also reported in a study of women with bulimic symptoms (Díaz-Ferrer et al., 2015) and later in a study of body dissatisfied and subclinical eating disordered women (Díaz-Ferrer et al., 2017), both studies used pure and guided exposure. In line with our results, Trentowska et al. (2013) reported a reduction of subjective distress and negative affect between sessions using a similar approach to that used in this study, including guiding attention and a neutral description of the various body parts. A study by Tanck et al. (2021) found positive between-session effects among healthy women in terms of body satisfaction and negative affect for both types of instructions – positive or negative verbalizations – of body parts during exposure. Taken together, all of these findings imply the importance of repeated mirror exposure in clinical practice even though the possible different underlying mechanisms have not yet been clarified.

Another positive effect of repeated mirror exposure was the lower level of post-event processing after the second session. Experiencing the second exposure as less aversive could have reduced engagement in ruminative processes afterwards. Furthermore, higher post-event processing after the first exposure predicted higher tension, shame, sadness, and fear when individuals were initially confronted with their body in the mirror at the second exposure. Addressing dysfunctional post-processing and providing more helpful ways to deal with uncertainty could therefore be very important interventions for clinical practice. Due to the potential impact on memory and associated expectations, addressing post-event processing could possibly amplify the positive effects or reduce negative effects due to dysfunctional processing between exposure sessions.

In sum, repeated mirror exposure may improve important processes, and post-event possessing may be an indicator for successful coping or learning. However, this assumption needs further scientific investigation. Mechanisms of mirror exposure for eating disorders have been discussed by Naumann et al. (2022). They proposed cognitive restructuring and related changes in cognitive biases as plausible explanations. In this study, we did not induce any specific expectations regarding the effects of the exposure and did not examine the expectations of participants. Therefore, we cannot accurately assess the role of expectations in this context. Since current research on anxiety disorders point to the important role of changes in expectations during exposure and the exposure-related learning rate (Pittig et al., 2023), it would be very important to investigate the role of expectation changes and learning rate through mirror exposure in the future. We should incorporate the assessment of expectations prior to the intervention and understand how expectation may influence the results.

Some limitations of the present study need to be considered when interpreting the results. One main limitation is the not fully clinical sample and the lack of confirmation of the clinical diagnosis by means of a clinical interview. Individuals with BDD symptoms who strongly avoid mirrors probably did not take part in the study. Differential effects may become apparent in comparison to a fully clinical sample. We investigated

women only, even though there is a substantial group of males with BDD. The internal description of appearance during exposure may have reduced the intensity of affect intensity. Although participants said they were able to follow the instructions, we cannot know, i.e., whether they shifted their visual attention to different body parts. In addition, possibly ongoing safety behavior could have been better recognized by a therapist guiding the participant. We focused on assessing in-session and short-term between-session effects. Hence, further investigation of the stability of the improvements is required. It is of interest to see whether further sessions, including repetition as homework, contribute to an even greater change in individuals with BDD.

Conclusions

In sum, in spite of the limitations, this research provides further insights into mirror exposure in individuals with BDD symptoms. It presents evidence for a positive effect of repeated guided mirror exposure in terms of state body image and affective response in women. The role of cognitive post-processing of mirror confrontation is highlighted, and new research questions are derived to address the key change mechanisms of mirror exposure.

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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

Competing Interests: We have no known conflict of interest to disclose.

Ethics Statement: The study received approval by the ethics committee of the University of Wuppertal. The study was conducted according to common research standards, including having obtained informed consent of all participants prior to participation.

Preregistration: The study was not submitted to a preregistration platform.

Reporting Guidelines: The study adheres to the JARS-Quant Guidelines of the American Psychological Association.

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

Supplementary Materials

The Supplementary Materials (Schoenenberg & Martin, 2026S) include a table listing the interview questions about the exposure experience (Supplementary Table 1), post-hoc tests on affect for the factor time (Supplementary Table 2), and results for the Repeated Measures Analysis of Variances comparing the group with BDD-Diagnosis to the group without diagnosis (Supplementary Table 3).

Index of Supplementary Materials

Schoenenberg, K., & Martin, A. (2026S). *Supplementary materials to "Repeated mirror exposure in individuals with body dysmorphic symptoms"* [Supplementary tables]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.21860

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Barnier, E. M., & Collison, J. (2019). Experimental induction of self-focused attention via mirror gazing: Effects on body image, appraisals, body-focused shame, and self-esteem. *Body Image, 30*, 150–158. https://doi.org/10.1016/j.bodyim.2019.07.003

Butler, R. M., & Heimberg, R. G. (2020). Exposure therapy for eating disorders: A systematic review. *Clinical Psychology Review, 78*, Article 101851. https://doi.org/10.1016/j.cpr.2020.101851

Cash, T. F. (2012). Cognitive-behavioral perspectives on body image. In T. F. Cash (Ed.), *Encyclopedia of body image and human appearance* (pp. 334–342). Elsevier.

Cash, T. F., Fleming, E. C., Alindogan, J., Steadman, L., & Whitehead, A. (2002). Beyond body image as a trait: The development and validation of the Body Image States Scale. *Eating Disorders, 10*(2), 103–113. https://doi.org/10.1080/10640260290081678

Díaz-Ferrer, S., Rodríguez-Ruiz, S., Ortega-Roldán, B., Mata-Martín, J. L., & Carmen Fernández-Santaella, M. (2017). Psychophysiological changes during pure vs guided mirror exposure therapies in women with high body dissatisfaction: What are they learning about their bodies? *European Eating Disorders Review, 25*(6), 562–569. https://doi.org/10.1002/erv.2546

Díaz-Ferrer, S., Rodríguez-Ruiz, S., Ortega-Roldán, B., Moreno-Domínguez, S., & Fernández-Santaella, M. C. (2015). Testing the efficacy of pure versus guided mirror exposure in women with bulimia nervosa: A combination of neuroendocrine and psychological indices. *Journal of Behavior Therapy and Experimental Psychiatry, 48*, 1–8. https://doi.org/10.1016/j.jbtep.2015.01.003

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

Fehm, L., Hoyer, J., Schneider, G., Lindemann, C., & Klusmann, U. (2008). Assessing post event processing after social situations: A measure based on the cognitive model for social phobia. *Anxiety, Stress, and Coping, 21*(2), 129–142. https://doi.org/10.1080/10615800701424672

Goss, K., & Allan, S. (2009). Shame, pride and eating disorders. *Clinical Psychology & Psychotherapy, 16*(4), 303–316. https://doi.org/10.1002/cpp.627

Griffen, T. C., Naumann, E., & Hildebrandt, T. (2018). Mirror exposure therapy for body image disturbances and eating disorders: A review. *Clinical Psychology Review, 65*, 163–174. https://doi.org/10.1016/j.cpr.2018.08.006

Johnson, S., Williamson, P., & Wade, T. D. (2018). A systematic review and meta-analysis of cognitive processing deficits associated with body dysmorphic disorder. *Behaviour Research and Therapy, 107*, 83–94. https://doi.org/10.1016/j.brat.2018.05.013

Kollei, I., & Martin, A. (2014). Body-related cognitions, affect and post-event processing in body dysmorphic disorder. *Journal of Behavior Therapy and Experimental Psychiatry, 45*(1), 144–151. https://doi.org/10.1016/j.jbtep.2013.09.005

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

LimeSurvey GmbH. (n.d.). *LimeSurvey: An open source survey tool* (Version 3) [Computer software]. LimeSurvey GmbH. https://www.limesurvey.org

Malcolm, A., Pikoos, T., Castle, D. J., Labuschagne, I., & Rossell, S. L. (2021). Identity and shame in body dysmorphic disorder as compared to obsessive-compulsive disorder. *Journal of Obsessive-Compulsive and Related Disorders, 31*, Article 100686. https://doi.org/10.1016/j.jocrd.2021.100686

Mancuso, S. G., Knoesen, N. P., & Castle, D. J. (2010). The dysmorphic concern questionnaire: A screening measure for body dysmorphic disorder. *The Australian and New Zealand Journal of Psychiatry, 44*(6), 535–542. https://doi.org/10.3109/00048671003596055

Moreno-Domínguez, S., Rodríguez-Ruiz, S., Fernández-Santaella, M. C., Jansen, A., & Tuschen-Caffier, B. (2012). Pure versus guided mirror exposure to reduce body dissatisfaction: A preliminary study with university women. *Body Image, 9*(2), 285–288. https://doi.org/10.1016/j.bodyim.2011.12.001

Naumann, E., Werthmann, J., Vocks, S., Svaldi, J., & Hartmann, A. S. (2022). Die Spiegelkonfrontationstherapie zur Behandlung von Körperbildstörungen bei Essstörungen – Evidenz, Wirkmechanismen und Vorgehensweise [Mirror confrontation therapy for treating body image disorders in eating disorders – Evidence, mechanisms of action, and procedure]. *Psychologische Rundschau, 73*(4), 243–259. https://doi.org/10.1026/0033-3042/a000558

Neziroglu, F., Hickey, M., & McKay, D. (2010). Psychophysiological and self-report components of disgust in body dysmorphic disorder: The effects of repeated exposure. *International Journal of Cognitive Therapy, 3*(1), 40–51. https://doi.org/10.1521/ijct.2010.3.1.40

Oosthuizen, P., Lambert, T., & Castle, D. J. (1998). Dysmorphic concern: Prevalence and associations with clinical variables. *The Australian and New Zealand Journal of Psychiatry, 32*(1), 129–132. https://doi.org/10.3109/00048679809062719

Pittig, A., Heinig, I., Goerigk, S., Richter, J., Hollandt, M., Lueken, U., Pauli, P., Deckert, J., Kircher, T., Straube, B., Neudeck, P., Koelkebeck, K., Dannlowski, U., Arolt, V., Fydrich, T., Fehm, L., Ströhle, A., Totzeck, C., Margraf, J., . . . Wittchen, H. U. (2023). Change of threat expectancy as

mechanism of exposure-based psychotherapy for anxiety disorders: Evidence from 8,484 exposure exercises of 605 patients. *Clinical Psychological Science, 11*(2), 199–217. https://doi.org/10.1177/21677026221101379

Rachman, S., Grüter-Andrew, J., & Shafran, R. (2000). Post-event processing in social anxiety. *Behaviour Research and Therapy, 38*(6), 611–617. https://doi.org/10.1016/S0005-7967(99)00089-3

Schieber, K., Kollei, I., de Zwaan, M., & Martin, A. (2015). Classification of body dysmorphic disorder – What is the advantage of the new DSM-5 criteria? *Journal of Psychosomatic Research, 78*(3), 223–227. https://doi.org/10.1016/j.jpsychores.2015.01.002

Schieber, K., Kollei, I., de Zwaan, M., & Martin, A. (2018). The dysmorphic concern questionnaire in the German general population: Psychometric properties and normative data. *Aesthetic Plastic Surgery, 42*(5), 1412–1420. https://doi.org/10.1007/s00266-018-1183-1

Schoenenberg, K., & Martin, A. (2022). Cognitive-affective reactions to a non-judgmental and judgmental mirror gazing task in individuals with body dysmorphic concerns. *Journal of Behavior Therapy and Experimental Psychiatry, 77*, Article 101779. https://doi.org/10.1016/j.jbtep.2022.101779

Stangier, U., Hungerbühler, R., Meyer, A., & Wolter, M. (2000). Diagnostische Erfassung der Körperdysmorphen Störung: Eine Pilotstudie [Assessment of body dysmorphic disorder: A pilot study]. *Der Nervenarzt, 71*, 876–884. https://doi.org/10.1007/s001150050678

Stangier, U., Janich, C., Adam-Schwebe, S., Berger, P., & Wolter, M. (2003). Screening for body dysmorphic disorder in dermatological outpatients. *Dermatology and Psychosomatics / Dermatologie und Psychosomatik, 4*(2), 66–71. https://doi.org/10.1159/000072194

Tanck, J. A., Hartmann, A. S., Svaldi, J., & Vocks, S. (2021). Effects of full-body mirror exposure on eating pathology, body image and emotional states: Comparison between positive and negative verbalization. *PLoS One, 16*(9), Article e0257303. https://doi.org/10.1371/journal.pone.0257303

Trentowska, M., Bender, C., & Tuschen-Caffier, B. (2013). Mirror exposure in women with bulimic symptoms: How do thoughts and emotions change in body image treatment? *Behaviour Research and Therapy, 51*(1), 1–6. https://doi.org/10.1016/j.brat.2012.03.012

Veale, D., & Neziroglu, F. (2010). *Body dysmorphic disorder: A treatment manual*. Wiley-Blackwell. https://doi.org/10.1002/9780470684610

Veale, D., & Riley, S. (2001). Mirror, mirror on the wall, who is the ugliest of them all? The psychopathology of mirror gazing in body dysmorphic disorder. *Behaviour Research and Therapy, 39*(12), 1381–1393. https://doi.org/10.1016/S0005-7967(00)00102-9

Vocks, S., Hechler, T., Rohrig, S., & Legenbauer, T. (2009). Effects of a physical exercise session on state body image: The influence of pre-experimental body dissatisfaction and concerns about weight and shape. *Psychology & Health, 24*(6), 713–728. https://doi.org/10.1080/08870440801998988

Waldorf, M., Vocks, S., Düsing, R., Bauer, A., & Cordes, M. (2019). Body-oriented gaze behaviors in men with muscle dysmorphia diagnoses. *Journal of Abnormal Psychology, 128*(2), 140–150. https://doi.org/10.1037/abn0000403

Wilhelm, S., Phillips, K. A., & Steketee, G. (2013). *Cognitive-behavioral therapy for body dysmorphic disorder: A treatment manual.* Guilford Press.

EACLIPT

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

zpid
Leibniz-Institut für Psychologie

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

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

Are You Really Paranoid, Just Because They Are After You? Exploring the Underlying Sensitization Processes of Intersectional Discrimination on Everyday-Life Paranoia

Björn Schlier[1], Felix Strakeljahn[2], Christoph Kahl[2], Katharina Winkler[2]

[1] *Department of Clinical Child and Adolescent Psychology and Psychotherapy, University of Wuppertal, Wuppertal, Germany.* **[2]** *Department of Clinical Psychology and Psychotherapy, University of Hamburg, Hamburg, Germany.*

Clinical Psychology in Europe, 2026, Vol. 8(2), Article e19263, https://doi.org/10.32872/cpe.19263

Received: 2025-08-11 • **Accepted:** 2025-12-16 • **Published (VoR):** 2026-05-29

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

Corresponding Author: Björn Schlier, Friedrichstraße 40, 42105 Wuppertal, Germany. Tel.: +49 202-439 5502. E-mail: schlier@uni-wuppertal.de

Supplementary Materials: Materials [see Index of Supplementary Materials]

Abstract

Background: Elevated paranoia levels have been found in discriminated, minoritized groups. Social-cognitive models of paranoia posit that experiences of discrimination strengthen negative core beliefs, which in turn foster clinical paranoia. In contrast, the healthy cultural mistrust hypothesis proposes that elevated paranoia in minoritized groups reflects an adaptive response to increased exploitation/discrimination. To explore whether elevated paranoia in discriminated groups can be fully understood within clinical models, we tested whether a history of discrimination amplifies the association between everyday-life stressors and subsequent state-paranoia and whether this moderation effect remains when controlling for established cognitive risk factors for paranoia (i.e., dysfunctional beliefs).

Method: A general population sample (n = 108) answered a baseline self-report questionnaire of lifetime discrimination experiences (LDE) and core beliefs about oneself and others. Next, they reported state-paranoia, negative affect, and exclusion experiences in 2.5h intervals for seven days of ambulatory assessment. LDE and core-beliefs were tested as concurrent moderators for the associations between state-paranoia and putative triggers (negative-affect, exclusion) using multilevel regression.

Results: More LDE amplified the association between paranoia, negative affect, and exclusion experiences. When controlling for the moderation effects of negative beliefs, effect sizes for the LDE moderation were lower but largely remained significant.

Conclusions: Increased state-paranoia in the daily lives of people with discrimination experiences can be partially, but not fully, explained by clinical, cognitive risk factors. Consequently, healthy cultural mistrust can be considered an additional contributing factor to higher mistrust/paranoia in minoritized groups. Future in-depth research needs to disentangle mechanisms of emerging paranoia and adaptive mistrust in minoritized groups.

Keywords

discrimination, minority, intersectionality, paranoia, psychosis, ambulatory assessment

Highlights

- Lifetime discrimination experiences (LDE) correlate strongest with lifetime paranoia among psychotic symptoms.
- LDE amplifies cross-sectional associations between state-paranoia, –negative affect, and –exclusion.
- LDE amplifies time-lagged associations from exclusion to paranoia to negative affect.
- LDE effects remain when controlling for clinical risk factors, implying healthy cultural mistrust.

Various etiological models highlight the crucial role of social adversity in the emergence of clinical levels of psychosis and specific symptoms such as paranoia (e.g., Jaya et al., 2016; van Os et al., 2010). Cognitive theories (Beck et al., 2011; Freeman et al., 2002; Garety et al., 2001; Kesting & Lincoln, 2013; Morrison, 2001) suggest that paranoia results from such experiences via cognitive biases rooted in dysfunctional, i.e., more negative and/or less positive, beliefs about oneself and others. The *Social Defeat Hypothesis* (Selten et al., 2013) proposes that the underlying common factor of established risk factors for psychotic symptoms, including paranoia, such as migration, urban upbringing, or even illicit drug use, is experiencing social defeat or exclusion. Specifically, chronic social defeat/exclusion sensitizes the mesolimbic dopamine system, increasing the risk that subsequent stressors elicit excessive stress responses and thus the emergence of psychotic symptoms. Such frameworks allow us to explain how stressors found to be linked to psychotic symptoms may lead to the emergence of symptom-states in some individuals but not in others. According to the *affective pathway model to psychosis* (Myin-Germeys & Van Os, 2007), fluctuations in negative affect that follow a stressful event constitute a common stressor, with multiple ambulatory assessment studies linking negative affect levels to subsequent emergence and/or increase in psychotic symptoms in daily life (Krkovic et al., 2020; Ludwig et al., 2020). In recent years, some experimental (e.g., Lincoln et al., 2018) and ambulatory assessment studies (e.g., Monsonet et al., 2022; Schlier et al., 2018) have shown that momentary experiences of social defeat or social exclusion can also serve as a situational trigger of positive symptoms such as hallucinations or paranoia. This tentatively suggests that not only a history of social

defeat/exclusion (that affects vulnerability), but also current momentary experiences of exclusion (as a specific trigger) constitute relevant risk factors for psychosis. Which symptoms are directly triggered by social defeat/exclusion, however, is still an open question (Shovestul et al., 2023).

Some researchers have tried to focus on individual symptoms rather than broad diagnostic or syndrome categories to reveal risk factors and/or stressors that are specific to the emergence of paranoia, hallucinations, or other psychotic symptoms (Bentall et al., 2014). Such a symptom-specific approach may not just expand our understanding of the aetiology of heterogeneous expressions of psychosis but also guide case formulation via the identification of an individual's most effective treatment targets for psychotherapeutic interventions. For paranoia, one such putatively specific risk factor is discrimination. A review on the topic of the association between discrimination experience and psychosis showed that a substantial number of studies found a specific association between discrimination experience and paranoia, specifically in at-risk and population samples (Pearce et al., 2019). Discussing these finding, Pearce and colleagues (2019) drew a parallel to findings of more pronounced negative core beliefs about the self and others in minority samples as well as "intrinsic social inequalities that underpin discrimination" that might lead to experiences of social defeat, proposing discrimination to be a vulnerability factor for paranoia that could be explained within the framework of existing clinical models.

At the same time, disparities in the U.S.-American health-care system, particularly increased rates of misdiagnosis of psychosis in ethnic minorities and their under-utilization of the mental health care system, have been viewed under the assumption of *healthy cultural mistrust* (Whaley, 2001b). Cultural mistrust describes a response style that results from experiences of discrimination/oppression, which informs a person's attitudes about the majority group in society, leading to a functional apprehensiveness in subsequent interactions with members of that group (Whaley, 2001a). Thus, the healthy cultural mistrust hypothesis suggests that what is phenomenologically perceived as paranoia in minoritized groups could reflect a healthy, adaptive response to discrimination (Whaley, 2001b). From this perspective, paranoid responses rooted in a history of discriminatory experiences could be etiologically distinct from paranoid responses rooted in the aforementioned clinically relevant social-cognitive mechanisms. At worst, they could constitute a phenomenological expression of a distinct form of distress, e.g., depression clad in contextually-evoked suspiciousness, that is diagnostically misclassified as paranoia.

Pearce and colleagues (2019) found that evidence on the connection between discrimination and paranoia primarily stems from cross-sectional studies, limiting our understanding of the underlying mechanisms connecting the two. Utilizing a large multi-cultural dataset, Kingston et al. (2023) highlighted potential differences in paranoia as reported by minoritized groups vs. the majority. They showed that paranoia levels were higher in people from minoritized groups and increased with more intersection-

ality: With each additional minoritized group people identified with, paranoia levels were higher. Intriguingly, various associations between established social-cognitive risk factors found to correlate with paranoia in majority groups (i.e., positive core beliefs about the self, about others, and low social rank beliefs as an indicator of habitual social defeat) did not show the same association in (intersectionally) minoritized groups. This tentatively suggests that there is a distinct underlying mechanism in minoritized groups, such as healthy cultural mistrust.

To further investigate these putative etiological differences, the current study re-analyses an existing ambulatory assessment dataset to investigate the influence of lifetime experiences of discrimination on current experiences of paranoia and their association with established momentary triggers of paranoid states. In this (due to the authors' prior knowledge of the data) not pre-registered secondary analysis, we aim to explore to what extent and by what mechanism the lifetime experiences of discrimination sensitize people to the experience of paranoia. In a first step, we explore to what degree the lifetime discrimination experience (LDE) attributable to belonging to one or more minoritized groups, i.e., increasing levels of intersectional LDE, correlate with lifetime experiences of paranoia. Additionally, we compare this correlation with associations of LDE and the lifetime experience of other psychotic symptoms. Further, we explore to what extent LDE covaries with changes in core beliefs about oneself and other people. In the second step, we test whether LDE affects the daily life experience of aversive social situations (i.e., social exclusion), negative affect, and paranoid states. To this end, we first test whether more LDE covaries with higher person-average levels of social exclusion, negative affect, and paranoid states in daily life. Next, we test whether LDE modulates the effects of state-social exclusion, -negative affect, and -paranoia exerted on each other. Finally, we investigate whether any such modulating effects can be explained by differences in core beliefs associated with discrimination.

Method

Design and Procedure

The data for this study was derived from a larger study on social antecedents of psychosis symptoms. It consisted of a baseline assessment followed by a seven-day ambulatory assessment period. Participants provided informed consent and then completed a 20-to-30-minute questionnaire battery comprising self-report measures on lifetime psychotic experience, core beliefs about the self, minority status, discrimination, and other interpersonal and aversive experiences. The questionnaire was presented on the survey platform Questback EFS.

Next, an ambulatory assessment was conducted using the app movisensXS. The ambulatory assessment started immediately after the baseline assessment and contin-

ued for seven days. Participants either received an Android smartphone with the app movisensXS preinstalled (n = 12) or installed the app on their smartphones (n = 96). Participants received the first assessment prompt during the introductory session to ensure there were no technical problems with the app and to answer any questions on handling the app. On the day of the baseline assessment, participants received one to five prompts, depending on the time of the first alarm. On the following six days, participants received five alarms per day, presented between 10 am and 9 pm in 2.5-hour intervals (+/-15min). Thus, the total number of questionnaires presented during the ambulatory assessment phase ranged from 31 to 35. The completion of an assessment questionnaire took approximately five minutes. If participants were unable to answer the questionnaire immediately after the prompt, it could be postponed by an additional ten minutes.

After seven days, participants returned to our lab for debriefing, to return the smartphone they received for the study, and to provide brief feedback on any problems or unusual events during the assessment.

Materials

Baseline Assessment

At baseline, lifetime experience of discrimination (LDE) was assessed with a modified version of the discrimination assessment from the NEMESIS study (Janssen et al., 2003), which has been used in this form multiple times in large-scale surveys (Jaya et al., 2016; Kingston et al., 2023). Using a yes/no answer format, participants first indicated whether they belonged to any of five minority group categories, i.e., ethnicity, religion, sexual orientation/identity, disability, or physical differences (such as obesity or visible scars). Following this, they were presented with two to seven additional yes/no questions asking whether they experienced discrimination due to being a member of any of the previously mentioned minority groups (individual questions only presented when participants indicated they were a member of the corresponding minority) or because of their gender or age (questions presented to every participant). A total score of all positively answered discrimination questions, ranging from 0 to 7, was calculated as an indicator of the extent of intersectional LDE. Higher values indicate LDE due to more coinciding forms of discrimination.

Lifetime prevalence of psychotic experiences was assessed with the validated German version of the Community Assessment of Psychic Experience (CAPE). This 42-item self-report questionnaire measures psychotic experiences in nine symptom categories that can further be summarized into three dimensions (Schlier et al., 2015): The dimension positive symptoms (20 items, e.g., “Do you ever feel as if people seem to drop hints about you or say things with a double meaning?”), consisted of the symptoms paranoid beliefs (five items), bizarre experiences (seven items), hallucinations (four items), grandiosity (two items), and magical thinking (two items); Negative symptoms (14 items, e.g., “Do you ever feel that you are not a very animated person?”), consisted of the symptoms

amotivation (seven items), anhedonia (three items), and social withdrawal (four items). Finally, depressive symptoms (eight items, e.g., "Do you ever feel sad?") constituted a third dimension with no subfactors. All items are answered on four-point Likert-scales (1 = "never", 2 = "sometimes", 3 = "often", 4 = "nearly always"). The German version of the CAPE has shown good reliability and validity across population and patient samples (Schlier et al., 2015). Mean scores for symptom dimensions and individual symptom factors were calculated.

Core beliefs about oneself and others were assessed with the Brief Core Schema Scale (BCSS; Fowler et al., 2006), a 24-item questionnaire assessing cognitive-schemas particularly relevant in psychosis. Specifically, the four dimensions negative-self-beliefs (e.g., "I am worthless"), positive-self-beliefs (e.g., "I am good"), negative-other-beliefs (e.g., "other people are nasty"), and positive-other-beliefs (e.g., "other people are supportive") are assessed with six items each. Participants rated the items on five-point Likert scales (1 = "no, do not believe it"; 5 = "yes, fully believe it"). The BCSS has shown good validity and high internal consistency across all four dimensions (Fowler et al., 2006), Cronbach's α ranged from .78 to .88, and its German translation has been used in multiple population studies (e.g., Kingston et al., 2023). Mean scores for each of the four core-schema dimensions were calculated for this study.

Additionally, we assessed habitually experienced social exclusion as an indicator of baseline social defeat, using the three-item inclusion-subscale from the *Social Comparison Scale* (SCS; Allan & Gilbert, 1995). The SCS measures different aspects of social comparisons with others (i.e., social rank, social attractiveness, and inclusion/exclusion). All items are ten-point semantic differentials between adjective pairs. The three-item inclusion-subscale focuses on the subjective feeling of being accepted vs. excluded by others (e.g., "In relationship to others I feel: left out – accepted"). At baseline, participants were asked to indicate how they "usually" feel in relationships with others. Social defeat assessments vary and ideally include multiple indicators – often encompassing feelings of exclusion, defeatist beliefs and self-reported discrimination or minority status (e.g., Shovestul et al., 2023). Nevertheless, prolonged/repeated feelings of exclusion are often described as core features of chronic social defeat (e.g., Selten et al., 2016). We calculated a reversed mean-score as an indicator of exclusion, with higher values indicating more intense feelings of social defeat.

Ambulatory Assessment

The ambulatory assessment included short self-report measures on current negative affect, psychotic experiences, and social interactions. All 32 items (2-6 per variable, see Supplementary Materials, Table S1) referred to the time period since the last prompt, except for negative affect, which referred to the current moment. For the analyses in this study, negative affect, paranoia, and social exclusion assessments are used.

To assess current negative affect, participants were asked to rate their current feelings of anxiety, anger, sadness, and shame based on a list of four adjectives (e.g., anxiety: "anxious, fearful, afraid, worried"). Participants rated each item on an 11-point Likert scale (0 = "does not apply to me at all", 10 = "strongly applies to me"). These items have been used in multiple ambulatory assessment studies (e.g., Krkovic et al., 2018, 2020). A negative affect mean score for each assessment was calculated.

State paranoid ideation was assessed using the five-item Brief State Paranoia Checklist (Schlier et al., 2016), which was developed as an ambulatory assessment ready, change-sensitive version of the Paranoia Checklist (Freeman et al., 2005). It includes items on milder social-evaluative concerns (e.g., "People laughed at me.") to more prototypically clinical paranoia (e.g., "My actions/thoughts might have been controlled by others"). Items were answered on 11-point Likert scales (0 = "does not apply to me at all", 10 = "strongly applies to me"). The five-item version of the Paranoia Checklist has shown sufficient reliability with a Cronbach's α of .83, within-person variability, and validity in its validation samples (Schlier et al., 2016).

State social exclusion was again assessed with the three-item subscale "group fit" from the *SCS* (Allan & Gilbert, 1995). The German translation of the scale was adopted from a previous ambulatory assessment study (Schlier et al., 2018). Participants were asked how they perceived themselves in relation to others since the last prompt. A reversed mean-score was calculated as an indicator of state social exclusion (i.e., higher values indicate more social exclusion).

Participants

Participants were recruited at the University of Hamburg via advertisement on Campus and in the digital participant recruitment database of the department of psychology. Demographic characteristics of the 108 participants are summarized in Table 1. The participants were young (mean age: 24.24 years), consisted mostly of university students (95.37%) with up to 1,000€ of monthly income (76.85%), and predominantly identified as women (77.78%). About half of the sample reported they have experienced one or more types of discrimination (52.78%), with the most prevalent forms being discrimination due to gender (33.33%), age (25.93%), visible physical differences (e.g., obesity or scars: 8.33%), and ethnicity/religion (6.48%, respectively).

The study was approved by the ethics committee of the University of Hamburg. All participants provided written informed consent prior to participation in accordance with the Declaration of Helsinki.

Table 1

Demographic Information of the Study Sample

Demographic Variable	*M* or %	*SD*	*Range*
Age	24.24	5.68	18-43
Gender			
Man	22.22%	–	–
Woman	77.78%	–	–
Other	–	–	–
No Answer	–	–	–
Current occupation			
Student	95.37%	–	–
Employed	3.70%	–	–
Self-employed	0.92%	–	–
Monthly income			
up to 500€	40.74%	–	–
500€ to 1,000€	36.11%	–	–
1,000€ to 1,500€	13.89%	–	–
1,500€ to 2,000€	5.56%	–	–
more than 2,000€	3.70%	–	–
Minority status due to…			
…ethnic minority	12.96%	–	–
…sexual orientation/identity	6.48%	–	–
…religious beliefs	7.41%	–	–
…visual physical difference	12.00%	–	–
Discrimination Experience			
Reporting any type of lifetime discrimination	52.78%	–	–
Experience of discrimination due to…			
…Age	25.93%	–	–
…Gender	33.33%	–	–
…Ethnicity	6.48%	–	–
…Sexual orientation/identity	3.70%	–	–
…Religious beliefs	6.48%	–	–
…Visual physical difference	8.33%	–	–
Number of lifetime discrimination types experienced			
All participants	0.84	0.98	0-4
…Participants reporting any discrimination (n = 57)	1.60	0.78	1-4

Note. *M* = Mean; *SD* = standard deviation.

Data Analysis

To test for and compare the association between LDE and different types of lifetime psychotic experiences, we calculated Spearman correlation tests between LDE and CAPE symptom dimensions/factors. Additionally, Spearman correlation tests were calculated between LDE and BCSS scores to identify cognitive changes that potentially underlie sensitization processes.

Next, we tested for the effect of LDE on state negative affect, paranoia, and social exclusion by calculating random-intercept, fixed-slope multilevel models with LDE as level-two predictors and either state negative affect, state paranoia, or state social exclusion as outcome. Further, to test for the sensitization effects of LDE, we calculated multilevel models with one of the state variables, negative affect, paranoia, or social exclusion as the dependent variable and another of the three state variables, as well as the intersectional LDE index, as independent variables. Both cross-sectional and time-lagged model variations were calculated for each of these combinations. Time-lagged models were built with a one-assessment-interval lag, i.e., 2.5 hours, and also controlled for the dependent variable scores at the prior time-point. Interaction effects between the respective independent variable x LDE were tested.

Finally, we tested for potential sensitization effects due to baseline core schema levels that putatively vary as a function of LDE within a clinical vulnerability perspective on paranoia. We entered interactions with any risk factor significantly associated with LDE in the initial correlation tests as parallel moderators to the multilevel models. In these models, we tested for a combination of significant risk factor x independent variable interaction and no-longer-significant LDE x independent variable interactions. The significance level for each individual test was $p < .05$.

Results

All participants fully answered all questionnaires at baseline. In the ambulatory assessment, total compliance was 86.14%, with 57 participants answering more than 90% of their questionnaires, 39 answering 70%-90%, nine answering 50%-70%, and three answering between 40%-50% of their questionnaires.

Association of LDE With Lifetime Psychotic Experiences and Risk Factors

The results of all correlation tests between LDE and lifetime psychotic symptoms are summarized in Table 2. There were significant associations with positive symptom scores, paranoia scores, bizarre experiences scores, amotivation scores, and depression scores. As expected, the highest correlation was between LDE and lifetime paranoia, $r = .48$. Lifetime positive symptoms also showed a significant, yet only medium association,

r = .39. The majority of the remaining correlations amounted to non-significant, small to medium effect sizes, 0.11 ≤ r ≤ 26, except for grandiosity, r = .07, and social withdrawal, r = -.02. Post-hoc z-tests for differences between correlations controlling for the association between the symptom variable and paranoia (added as suggested during revision) showed paranoia to correlate significantly stronger with LDE than all other symptoms and symptom-dimensions, except for the positive symptom sum-score, z = 1.60, p = .055.

Table 2

Correlation Test Results Between Lifetime Discrimination Experience (LDE) and Lifetime Psychotic Symptoms, Social Defeat, and Self and Other Beliefs at Baseline

	Descriptive Values			Correlation with LDE			Test for difference with correlation LDE-Paranoia	
Symptom/Risk Factor	***M***	***SD***	***Range***	***r***	***t*(106)**	***p***	***z***	***p***
Positive symptoms	1.40	0.20	1.05-2.10	**0.39**	**4.37**	**< .001**	1.60	.055
Paranoia	1.79	0.37	1.00-2.80	**0.48**	**5.67**	**< .001**	–	–
Bizarre experiences	1.19	0.23	1.00-2.00	**0.23**	**2.48**	**.015**	**2.62**	**.004**
Hallucinations	1.05	0.17	1.00-2.50	0.13	1.34	.183	**3.19**	**.001**
Grandiosity	1.87	0.61	1.00-3.50	0.07	0.70	.488	**3.63**	**< .001**
Magical Thinking	1.41	0.45	1.00-3.00	0.15	1.56	.122	**2.98**	**.001**
Negative symptoms	1.91	0.34	1.21-2.79	0.18	1.91	.059	**2.79**	**.003**
Amotivation	1.96	0.42	1.00-3.14	**0.25**	**2.62**	**.010**	**2.22**	**.013**
Anhedonia	1.65	0.57	1.00-3.33	0.11	1.10	.273	**3.08**	**.001**
Social withdrawal	2.03	0.44	1.25-3.25	-0.02	-0.23	.819	**4.18**	**< .001**
Depression	1.98	0.37	1.13-3.63	**0.26**	**2.72**	**.008**	**2.25**	**.012**
Baseline social defeat	14.49	5.41	3.00-27.00	**0.24**	**2.55**	**.012**	–	–
Negative self-beliefs	3.22	2.95	0.00-18.00	**0.22**	**2.32**	**.022**	–	–
Positive self-beliefs	14.61	5.10	2.00-24.00	-0.03	-0.34	.735	–	–
Negative other-beliefs	3.77	3.18	0.00-12.00	**0.29**	**3.14**	**.002**	–	–
Positive other-beliefs	13.02	4.38	3.00-24.00	-0.06	-0.58	.565	–	–

Note. Symptom scores range: 1 to 4; self/other belief scores range: 0 to 24; social defeat score range: 3 to 33. Significant correlations (p < .05) are printed in bold.

Regarding the association between LDE and core beliefs, there were significant associations with habitual social defeat, r = .24, negative-self-beliefs, r = .22, and negative-other-beliefs, r = .29, but not with positive-self-beliefs, r = -.03, or positive-other-beliefs, r = -.06 (see Table 2).

The Effects of LDE on Everyday Life Experiences

Multilevel models using LDE as a level two predictor showed that people with more LDE reported higher levels of negative affect, β = 0.27, *b* = 0.16, *SE* = 0.06, *t* = 2.74, *p* = .007, and paranoia, β = 0.46, *b* = 0.22, *SE* = 0.04, *t* = 5.10, *p* < .001, but not social exclusion in everyday life, β = 0.11, *b* = 0.16, *SE* = 0.14, *t* = 1.13, *p* = .263.

Table 3 provides an overview of all the moderation effects of LDE. As can be seen, the effect of social exclusion on paranoia is amplified by LDE in both the cross-sectional, β = 0.24, and time-lagged regression, β = 0.13. Further, there are significant moderation effects in the cross-sectional models for negative affect on paranoia, β = 0.35, and social exclusion, β = 0.10, plus paranoia on social exclusion, β = 0.10, with no corresponding effects in time-lagged analyses. Finally, the effect of paranoia on negative affect was significantly moderated in time-lagged analysis only, β = 0.08.

Comparison of Moderation Effects of LDE vs. Negative Core and Social Beliefs

The results of the moderation models testing the simultaneous effect of LDE and clinical cognitive risk factors from baseline (negative-other-beliefs, negative-self-beliefs, and social defeat) are summarized in Table 3 (right columns). All previously significant LDE moderation effects remained significant, except for the paranoia x LDE effect on subsequent negative affect.

Overall, standardized moderation effect sizes for LDE decreased when controlling for the clinical risk factors. However, in the models with a significant LDE-moderation effect, effect sizes of the concurrently tested moderations (see rightmost column of Table 3 for an overview and Table 4 for a detailed list of other moderation effects) were mostly within the same range. Notably, the controlled cross-sectional effect of social-exclusion x LDE, β = 0.15, was only marginally smaller than the highest other interaction effect for social exclusion x negative-other-beliefs, β = 0.17. For the time-lagged association between social exclusion and subsequent paranoia, the LDE moderation effect was highest, β = 0.13, albeit only marginally higher than the second-highest moderation effect, negative-self-beliefs: β = 0.11.

Clinical Psychology in Europe
2026, Vol. 8(2), Article e19263
https://doi.org/10.32872/cpe.19263

Table 3

Discrimination Experiences as Moderator for the Associations Between Everyday-Life Social Exclusion, Negative Affect, and Paranoia

		Models Including Predictor and Discrimination										Models Controlling for Core-Schemas and Social Defeat					
		Main Effect Predictor					Discrimination x Predictor					Discrimination x Predictor					Other Mod.
Outcome	**Predictor**	**β**	***b***	***SE***	***t***	***p*** [a]	**β**	***b***	***SE***	***t***	***p*** [a]	**β**	***b***	***SE***	***t***	***p*** [a]	**Range β**
Cross-Sectional Models																	
Paranoia	Social excl.	**0.22**	**0.07**	**0.01**	**6.41**	**< .001**	**0.24**	**0.04**	**0.01**	**6.31**	**< .001**	**0.15**	**0.03**	**0.01**	**3.63**	**< .001**	-0.14; 0.17
Social excl.	Paranoia	**0.25**	**0.45**	**0.06**	**7.19**	**< .001**	**0.10**	**0.09**	**0.03**	**2.76**	**.006**	**0.09**	**0.08**	**0.04**	**2.20**	**.028**	-0.05; 0.11
Paranoia	Neg. affect	**0.27**	**0.16**	**0.02**	**10.38**	**< .001**	**0.35**	**0.12**	**0.01**	**11.62**	**< .001**	**0.26**	**0.09**	**0.01**	**8.08**	**< .001**	-0.13; 0.29
Neg. affect	Paranoia	**0.48**	**0.57**	**0.04**	**14.45**	**< .001**	0.03	0.02	0.02	0.88	.378	0.00	0.00	0.02	< 0.01	0.999	-0.03; 0.10
Social excl.	Neg. affect	**0.34**	**0.42**	**0.04**	**12.32**	**< .001**	**0.10**	**0.07**	**0.02**	**3.08**	**.002**	0.06	0.05	0.03	1.80	0.072	-0.01; 0.22
Neg. affect	Social excl.	**0.46**	**0.22**	**0.02**	**13.86**	**< .001**	0.00	0.00	0.01	0.07	.943	**-0.08**	**-0.02**	**0.01**	**-2.04**	**.041**	-0.08; 0.32
Lagged Models																	
Paranoia	Social excl.	-0.02	-0.01	0.01	-0.63	.529	**0.13**	**0.02**	**0.01**	**3.16**	**.002**	**0.12**	**0.02**	**0.01**	**2.65**	**.008**	-0.09; 0.11
Social excl.	Paranoia	0.01	0.02	0.08	0.28	.780	0.05	0.04	0.04	1.27	.206	0.02	0.02	0.04	0.51	.612	-0.03; 0.05
Paranoia	Neg. affect	0.03	0.02	0.02	0.96	.340	-0.02	-0.01	0.01	-0.67	.500	-0.06	-0.02	0.01	-1.61	.107	0.03; 0.06
Neg. affect	Paranoia	-0.03	-0.04	0.05	-0.83	.405	**0.08**	**0.05**	**0.02**	**2.08**	**.037**	0.04	0.03	0.03	1.01	.315	-0.08; 0.13
Social excl.	Neg. affect	-0.01	-0.01	0.04	-0.26	.489	0.05	0.03	0.03	1.33	.182	0.01	0.01	0.03	0.35	.729	-0.02; 0.07
Neg. affect	Social excl.	0.02	0.01	0.02	0.45	.654	0.07	0.02	0.01	1.80	.072	0.06	0.01	0.01	1.31	.190	-0.01; 0.13

Note. Neg. affect = negative affect; Social excl. = social exclusion. All models are random-intercept, fixed slope multilevel-regressions with predictor, discrimination index, and predictor x discrimination index as independent variables; Controlled models (third column-cluster) also include grand-mean-centered BCSS negative other-beliefs scores, BCSS negative self-beliefs scores, and SCS social rank beliefs and corresponding interaction effects with the predictor. Range β provides an overview of the standardized effects of the moderations entered as control variables. Time-lagged models are additionally controlled for the effect of the dependent variable scores from the previous assessment.

[a]*p*-values calculated with Satterthwaite approximation, with significant effects ($p < .05$) printed in bold.

Clinical Psychology in Europe
2026, Vol. 8(2), Article e19263
https://doi.org/10.32872/cpe.19263

Table 4

Discrimination Experiences as Moderator for the Associations Between Social Exclusion, Negative Affect, and Paranoia: Overview of all Tested Moderators

		Single moderator	Model with all moderation effects included											
		Predictor x Discrimination	Predictor x Discrimination			Predictor x Negative Other-Beliefs			Predictor x Negative Self-Beliefs			Predictor x Baseline Social Defeat		
Outcome	**Predictor**	**β**	**β**	***t***	***p*** [a]	**β**	***t***	***p*** [a]	**β**	***t***	***p*** [a]	**β**	***t***	***p*** [a]
Cross-Sectional Models														
Paranoia	Social excl.	**0.24**	**0.15**	**3.63**	**< .001**	**0.17**	**3.78**	**< .001**	**-0.14**	**-3.13**	**.002**	**0.12**	**2.64**	**.008**
Social excl.	Paranoia	**0.10**	**0.09**	**2.20**	**.028**	-0.05	-1.34	.179	**-0.08**	**-2.12**	**.034**	**0.11**	**3.19**	**.001**
Paranoia	Neg. affect	**0.35**	**0.26**	**8.08**	**< .001**	**0.29**	**10.04**	**<.001**	0.00	-0.01	.994	**-0.13**	**-3.97**	**< .001**
Neg. affect	Paranoia	0.03	0.00	< 0.01	.999	**0.10**	**3.01**	**.003**	0.06	1.60	.109	-0.03	-0.99	.322
Social excl.	Neg. affect	**0.10**	0.06	1.80	.072	-0.01	-0.46	.648	**-0.14**	**-3.96**	**< .001**	**0.22**	**6.27**	**< .001**
Neg. affect	Social excl.	0.00	**-0.08**	**-2.04**	**.041**	0.07	1.70	.089	**-0.13**	**-3.02**	**.003**	**0.32**	**7.12**	**< .001**
Time-Lagged Models														
Paranoia	Social excl.	**0.13**	**0.12**	**2.65**	**.008**	0.03	0.61	.539	**0.11**	**2.07**	**.039**	-0.09	-1.91	.057
Social excl.	Paranoia	0.05	0.02	0.51	.612	0.05	1.20	.229	-0.03	-0.74	.459	0.05	1.30	.192
Paranoia	Neg. affect	-0.02	-0.06	-1.61	.107	0.04	0.98	.327	0.06	1.47	.141	0.03	0.71	.476
Neg. affect	Paranoia	**0.08**	0.04	1.01	.315	0.05	1.32	.185	**-0.08**	**-2.06**	**.040**	**0.13**	**3.23**	**.001**
Social excl.	Neg. affect	0.05	0.01	0.35	.729	0.03	0.83	.406	-0.02	-0.41	.684	0.07	1.82	.068
Neg. affect	Social excl.	0.07	0.06	1.31	.190	0.00	-0.10	.920	**0.13**	**2.58**	**.010**	-0.01	-0.24	.809

Note. Neg. affect = negative affect; Social excl. = social exclusion. Models are random-intercept, fixed slope multilevel-regressions with predictor and discrimination index as well as, grand-mean centered negative other-beliefs scores, negative self-beliefs scores, social defeat and the listed interaction effects as independent variables. Time-lagged associations are additionally controlled for the effect of the dependent variable scores from the previous assessment.

[a]*p*-values calculated with Satterthwaite approximation, with significant effects ($p < .05$) printed in bold.

Discussion

In this study, we aimed to explore whether the extent of intersectional lifetime discrimination experience (LDE) continuously heightens susceptibility to feelings of paranoia/mistrust in minoritized groups. Additionally, we investigated whether this heightened sensitivity can be fully explained by key factors identified in socio-cognitive etiological models of clinical paranoia. In line with previous empirical and theoretical models (Bentall et al., 2014), we found that among all types of psychotic (positive, negative, and depressive) symptoms, paranoia shows the highest association with discrimination. Additionally, LDE is correlated with habitual feelings of social exclusion (i.e., social defeat), one's negative self-beliefs, and negative beliefs about others. This replicates the established link between LDE and both paranoia and socio-cognitive factors at the core of clinical paranoia models (Freeman et al., 2002; Selten et al., 2013).

When investigating the role of LDE on paranoia in daily life, we consistently found that an increasing complexity of intersectional LDE amplifies the dynamic interrelations of negative affect, state-exclusion, and state-paranoia. Among these, the largest amplifying effects were found for negative affect and state-exclusion on state-paranoia, with the effects for state-exclusion on state-paranoia being found in cross-sectional and time-lagged analyses. These directional effects mirror well-established triggers that have been shown to elicit paranoia both in lab studies (Lamster et al., 2017; Lincoln et al., 2018; Stewart et al., 2017) and numerous ambulatory assessment studies (e.g., Bell et al., 2023). With LDE intensifying this trigger-response-association at the heart of various vulnerability-stress models of paranoia and psychosis (negative emotions, e.g., Freeman et al., 2002; Myin-Germeys & Van Os, 2007; adversity, e.g., Selten et al., 2013), these findings at first glance indicate that discrimination indeed purely constitutes a clinical risk factor.

However, within the healthy cultural mistrust framework, increased mistrust in response to situational stressors could also be found in minoritized groups as an adaptive response to increased exploitation and discrimination (Whaley, 2001b): A person with more LDE might respond with mistrust to social stressors, but this response may not necessarily fall within the spectrum of clinically relevant paranoia and thus would not be governed by the same set of social-cognitive risk factors for clinical paranoia. Consequently, finding moderation effects of LDE that still exist when negative beliefs about oneself, others, and habitual social defeat are controlled for can be interpreted as a sign of non-clinical sensitization processes. Especially the stable moderation effect LDE found on state social exclusion and concurrent/subsequent state-paranoia shows that the differential reaction to current social adversity in people with LDE cannot be fully explained by changes in beliefs and/or habitual feelings of social defeat. This specificity to the association with a current social stressor tentatively suggests that the effects we found could be partially caused by healthy cultural mistrust.

Nevertheless, the reduced effect size of moderating effects of LDE when clinical risk factors were controlled for indicate that at least some of these influences can be explained within a clinical framework. In sum, this everyday-life perspective offers a potential explanation for previous findings from cross-sectional studies that showed a lowered associations between paranoia levels and some, but not all socio-cognitive risk factor levels in intersectionally minoritized groups (Kingston et al., 2023): Based on our results, we could assume that some seemingly-paranoid responses of healthy mistrust emerge in response to situational factors of the eliciting social-exclusion experience, i.e., an increased tendency towards facing new situations of discrimination; and these responses are not rooted in a cognitively biased perception of the situation due to dysfunctional core-schemas. In consequence, retrospective reports of core schemas about oneself and others and paranoid beliefs would be less strong/ubiquitous in minoritized groups due to confounding clinical paranoia and healthy mistrust. To further verify this hypothesis, future research needs to provide a more detailed picture of the nature of the everyday-life situations that elicit feelings of social exclusion in minoritized people and elucidate whether discriminatory features of these situational triggers contribute to an increase in paranoid responses when compared to non-minoritized groups.

This is, to our knowledge, the first attempt at disentangling the clinical spectrum of paranoia and healthy mistrust using ambulatory assessment and a competing moderation approach. Given that multiple clinical models exist besides the social and cognitive ones included here (for an overview, see Denecke et al., 2024), we can only preliminarily draw the conclusion that the moderating effects of LDE are outside the working mechanisms of clinical models of paranoia. It might well be that other mechanisms described in clinical models, for example, reasoning biases as recently framed within Bayesian inference paradigms (Coltheart et al., 2010; Fletcher & Frith, 2009), fully account for the LDE moderation effect we found. Even if that were to be the case, however, it would lead to theoretically and practically important follow-up questions of where to draw the line and declare a response of mistrust as clinical paranoia, and which associated mechanisms are crucial to this distinction.

Finally, some general limitations need to be considered: First, our sample is a population sample. While previous work has demonstrated multiple times that the spectrum of paranoia extends to the general population (Elahi et al., 2017), a comparison of clinical groups with and without minority backgrounds would provide more substantial evidence that the differences we found persist when treatment-relevant levels of distress are reached. Also, our study constituted a secondary analysis of existing data, so a pre-registered verification of our findings (preferably in population and clinical samples) is needed. Second, the vast majority of the healthy cultural mistrust research revolved around ethnicity. Some recent studies (Ellett et al., 2025; Kingston et al., 2023) provided evidence that this type of research can be generalized to other minority groups and other forms of discrimination, but with our comparatively small subsamples for individ-

ual minorities, we are unable to adequately analyse and compare different types of discrimination individually. Follow-up studies may benefit from closing this gap between our research and the bulk of healthy mistrust studies by stratifying participants along one or multiple types of minority group association. Finally, lifetime discrimination has been based on a broadly defined self-report assessment in this study. Consequently, certain degrees of nuance in intersectional minority status or discrimination could not be adequately captured (e.g., differentiation between gender-based discrimination and transphobic discrimination or discrimination because of sexual orientation vs. because of sexual orientation and sexual identity). This degree of imprecision could have led us to systematically underestimate associations involving complex intersectional discrimination. Further, relying on self-report leads to a risk of confounding an accurate assessment of discrimination with a bias due to paranoid thinking, possibly leading to an inflation of reported discrimination. While the complexity of our study, specifically the separation of state-paranoia and state-triggers in ambulatory assessment and discrimination as part of the baseline questionnaire as well as the pattern of our results showing associations between discrimination and subsequent situational dynamics of state-paranoia can be seen as arguments against this risk playing a prominent role, a more extensive assessment of discrimination experience (e.g., a interview-based assessment of factual descriptions of the situations perceived as discriminatory or a third person assessment via participants relatives or close friends) could provide a more valid measure of this variable in future research. Similarly, the baseline assessment of chronic social exclusion/social defeat was rather brief. While some aspects of chronic social defeat commonly used as additional indicators are derived from self-reported discrimination experiences and cannot be implemented as a social defeat measure in a study that aims to disentangle minoritization effects from social defeat, future studies could expand the assessment of chronic defeat by including additional constructs (e.g., self-reported lifetime feelings of ostracism; Shovestul et al., 2023) to increase the quality of chronic social defeat assessment.

In conclusion, our research shows that a history of (intersecting) discrimination experiences increases state-mistrust in response to (social) stress and that psychosocial clinical models of paranoia cannot fully explain this effect. Thus, healthy cultural mistrust could partially explain such reactions that are mistakenly labelled as paranoid ideation. By incorporating this perspective, we can further refine our understanding of paranoia and, through future research, optimize evidence-based guidelines for diversity-sensitive treatment.

Funding: The study received no external funding.

Acknowledgments: The authors would like to thank the Fatima Chebil for her help in conducting the study and the participants for their participation.

Competing Interests: The authors declare that they have no competing interests.

Ethics Statement: The study was approved by the Ethics Committee of the University of Hamburg (AZ 2018_191 Schlier KA). Informed consent was obtained from all participants prior to their participation in this study.

Preregistration: This study is not pre-registered.

Reporting Guidelines: The JARS-Quant guidelines for non-experimental quantitative research were followed in this study.

Data Availability: Data supporting the conclusions of this study are available upon reasonable request to the corresponding author. Data are not publicly available due to lack of explicit consent to open data publication in the informed consent.

Supplementary Materials

The Supplementary Materials contain an overview of the full ambulatory assessment questionnaire (for access, see Schlier et al., 2026S).

Index of Supplementary Materials

Schlier, B., Strakeljahn, F., Kahl, C., & Winkler, K. (2026S). *Supplementary materials to "Are you really paranoid, just because they are after you? Exploring the underlying sensitization processes of intersectional discrimination on everyday-life paranoia"* [Overview of the full ambulatory assessment questionnaire]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.21893

References

Allan, S., & Gilbert, P. (1995). A social comparison scale: Psychometric properties and relationship to psychopathology. *Personality and Individual Differences, 19*(3), 293–299. https://doi.org/10.1016/0191-8869(95)00086-L

Beck, A. T., Rector, N. A., Stolar, N., & Grant, P. (2011). *Schizophrenia: Cognitive theory, research, and therapy*. Guilford Press.

Bell, V., Velthorst, E., Almansa, J., Myin-Germeys, I., Shergill, S., & Fett, A.-K. (2023). Do loneliness and social exclusion breed paranoia? An experience sampling investigation across the

psychosis continuum. *Schizophrenia Research: Cognition, 33*, Article 100282. https://doi.org/10.1016/j.scog.2023.100282

Bentall, R. P., De Sousa, P., Varese, F., Wickham, S., Sitko, K., Haarmans, M., & Read, J. (2014). From adversity to psychosis: Pathways and mechanisms from specific adversities to specific symptoms. *Social Psychiatry and Psychiatric Epidemiology, 49*(7), 1011–1022. https://doi.org/10.1007/s00127-014-0914-0

Coltheart, M., Menzies, P., & Sutton, J. (2010). Abductive inference and delusional belief. *Cognitive Neuropsychiatry, 15*(1–3), 261–287. https://doi.org/10.1080/13546800903439120

Denecke, S., Schönig, S. N., Bott, A., Faße, J. L., & Lincoln, T. M. (2024). Bridging perspectives—A review and synthesis of 53 theoretical models of delusions. *Clinical Psychology Review, 114*, Article 102510. https://doi.org/10.1016/j.cpr.2024.102510

Elahi, A., Perez Algorta, G., Varese, F., McIntyre, J. C., & Bentall, R. P. (2017). Do paranoid delusions exist on a continuum with subclinical paranoia? A multi-method taxometric study. *Schizophrenia Research, 190*, 77–81. https://doi.org/10.1016/j.schres.2017.03.022

Ellett, L., Krkovic, K., Gaudiano, B., Thompson, E., & Kingston, J. (2025). Discrimination, minority group endorsement and paranoia in adolescents: The moderating role of loneliness. *International Journal of Adolescence and Youth, 30*(1), Article 2437029. https://doi.org/10.1080/02673843.2024.2437029

Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience, 10*(1), 48–58. https://doi.org/10.1038/nrn2536

Fowler, D., Freeman, D., Smith, B., Kuipers, E., Bebbington, P., Bashforth, H., Coker, S., Hodgekins, J., Gracie, A., Dunn, G., & Garety, P. (2006). The Brief Core Schema Scales (BCSS): Psychometric properties and associations with paranoia and grandiosity in non-clinical and psychosis samples. *Psychological Medicine, 36*(6), 749–759. https://doi.org/10.1017/S0033291706007355

Freeman, D., Garety, P. A., Bebbington, P. E., Smith, B., Rollinson, R., Fowler, D., Kuipers, E., Ray, K., & Dunn, G. (2005). Psychological investigation of the structure of paranoia in a non-clinical population. *British Journal of Psychiatry, 186*(5), 427–435. https://doi.org/10.1192/bjp.186.5.427

Freeman, D., Garety, P. A., Kuipers, E., Fowler, D., & Bebbington, P. E. (2002). A cognitive model of persecutory delusions. *British Journal of Clinical Psychology, 41*(4), 331–347. https://doi.org/10.1348/014466502760387461

Garety, P. A., Kuipers, E., Fowler, D., Freeman, D., & Bebbington, P. E. (2001). A cognitive model of the positive symptoms of psychosis. *Psychological Medicine, 31*(2), 189–195. https://doi.org/10.1017/S0033291701003312

Janssen, I., Hanssen, M., Bak, M., Bijl, R. V., De Graaf, R., Vollebergh, W., McKenzie, K., & Van Os, J. (2003). Discrimination and delusional ideation. *British Journal of Psychiatry, 182*(1), 71–76. https://doi.org/10.1192/bjp.182.1.71

Jaya, E. S., Ascone, L., & Lincoln, T. M. (2016). Social adversity and psychosis: The mediating role of cognitive vulnerability. *Schizophrenia Bulletin, 43*(3), 557–565. https://doi.org/10.1093/schbul/sbw104

Kesting, M.-L., & Lincoln, T. M. (2013). The relevance of self-esteem and self-schemas to persecutory delusions: A systematic review. *Comprehensive Psychiatry, 54*(7), 766–789. https://doi.org/10.1016/j.comppsych.2013.03.002

Kingston, J. L., Schlier, B., Lincoln, T., So, S. H., Gaudiano, B. A., Morris, E. M. J., Phiri, P., & Ellett, L. (2023). Paranoid thinking as a function of minority group status and intersectionality: An international examination of the role of negative beliefs. *Schizophrenia Bulletin, 49*(4), 1078–1087. https://doi.org/10.1093/schbul/sbad027

Krkovic, K., Clamor, A., Schlier, B., & Lincoln, T. M. (2020). Emotions and persecutory ideation in daily life: On the trail of the "chicken and egg" problem. *Journal of Abnormal Psychology, 129*(2), 215–223. https://doi.org/10.1037/abn0000495

Krkovic, K., Schlier, B., & Lincoln, T. (2018). An experience sampling study on the nature of the interaction between traumatic experiences, negative affect in everyday life, and threat beliefs. *Schizophrenia Research, 201*, 381–387. https://doi.org/10.1016/j.schres.2018.05.030

Lamster, F., Nittel, C., Rief, W., Mehl, S., & Lincoln, T. (2017). The impact of loneliness on paranoia: An experimental approach. *Journal of Behavior Therapy and Experimental Psychiatry, 54*, 51–57. https://doi.org/10.1016/j.jbtep.2016.06.005

Lincoln, T. M., Sundag, J., Schlier, B., & Karow, A. (2018). The relevance of emotion regulation in explaining why social exclusion triggers paranoia in individuals at clinical high risk of psychosis. *Schizophrenia Bulletin, 44*(4), 757–767. https://doi.org/10.1093/schbul/sbx135

Ludwig, L., Mehl, S., Schlier, B., Krkovic, K., & Lincoln, T. M. (2020). Awareness and rumination moderate the affective pathway to paranoia in daily life. *Schizophrenia Research, 216*, 161–167. https://doi.org/10.1016/j.schres.2019.12.007

Monsonet, M., Kwapil, T. R., & Barrantes-Vidal, N. (2022). A time-lagged study of predictors of paranoia and psychotic-like experiences in daily life across the schizotypy continuum. *Journal of Psychopathology and Clinical Science, 131*(1), 98–108. https://doi.org/10.1037/abn0000726

Morrison, A. P. (2001). The interpretation of intrusions in psychosis: An integrative cognitive approach to hallucinations and delusions. *Behavioural and Cognitive Psychotherapy, 29*(3), 257–276. https://doi.org/10.1017/S1352465801003010

Myin-Germeys, I., & Van Os, J. (2007). Stress-reactivity in psychosis: Evidence for an affective pathway to psychosis. *Clinical Psychology Review, 27*(4), 409–424. https://doi.org/10.1016/j.cpr.2006.09.005

Pearce, J., Rafiq, S., Simpson, J., & Varese, F. (2019). Perceived discrimination and psychosis: A systematic review of the literature. *Social Psychiatry and Psychiatric Epidemiology, 54*(9), 1023–1044. https://doi.org/10.1007/s00127-019-01729-3

Schlier, B., Jaya, E. S., Moritz, S., & Lincoln, T. M. (2015). The Community Assessment of Psychic Experiences measures nine clusters of psychosis-like experiences: A validation of the German version of the CAPE. *Schizophrenia Research, 169*(1–3), 274–279. https://doi.org/10.1016/j.schres.2015.10.034

Schlier, B., Moritz, S., & Lincoln, T. M. (2016). Measuring fluctuations in paranoia: Validity and psychometric properties of brief state versions of the Paranoia Checklist. *Psychiatry Research, 241*, 323–332. https://doi.org/10.1016/j.psychres.2016.05.002

Schlier, B., Winkler, K., Jaya, E. S., & Lincoln, T. M. (2018). Fluctuations in hallucination spectrum experiences co-vary with social defeat but not with social deafferentation: A 3-week daily assessment study. *Cognitive Therapy and Research, 42*(1), 92–102. https://doi.org/10.1007/s10608-017-9871-8

Selten, J.-P., Van Der Ven, E., Rutten, B. P. F., & Cantor-Graae, E. (2013). The Social Defeat Hypothesis of schizophrenia: An update. *Schizophrenia Bulletin, 39*(6), 1180–1186. https://doi.org/10.1093/schbul/sbt134

Selten, J.-P., Van Os, J., & Cantor-Graae, E. (2016). The social defeat hypothesis of schizophrenia: Issues of measurement and reverse causality. *World Psychiatry, 15*(3), 294–295. https://doi.org/10.1002/wps.20369

Shovestul, B., Scharf, M., Liu, G., & Dodell-Feder, D. (2023). Social defeat and psychosis-related outcomes: Associative and experimental tests related to the nature of defeat, specificity of outcomes, and psychosis-proneness. *Psychiatry Research Communications, 3*(4), Article 100149. https://doi.org/10.1016/j.psycom.2023.100149

Stewart, C., Rogers, F., Pilch, M., Stewart, I., Barnes-Holmes, Y., & Westermann, S. (2017). The effect of social exclusion on state paranoia and explicit and implicit self-esteem in a non-clinical sample. *Journal of Behavior Therapy and Experimental Psychiatry, 57*, 62–69. https://doi.org/10.1016/j.jbtep.2017.04.001

van Os, J., Kenis, G., & Rutten, B. P. F. (2010). The environment and schizophrenia. *Nature, 468*(7321), 203–212. https://doi.org/10.1038/nature09563

Whaley, A. L. (2001a). Cultural mistrust and mental health services for African Americans: A review and meta-analysis. *The Counseling Psychologist, 29*(4), 513–531. https://doi.org/10.1177/0011000001294003

Whaley, A. L. (2001b). Cultural mistrust and the clinical diagnosis of paranoid schizophrenia in African American patients. *Journal of Psychopathology and Behavioral Assessment, 23*(2), 93–100. https://doi.org/10.1023/A:1010911608102

EACLIPT

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

zpid
Leibniz-Institut für Psychologie

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

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

Promoting Flexibility in Expectations: A Randomized-Controlled Online-Intervention Study for Mild Psychopathological Symptoms

Anne-Catherine I. Ewen [1], Marcel Wilhelm [1]

[1] *Department of Clinical Psychology and Psychotherapy, Philipps-University of Marburg, Marburg, Germany.*

Clinical Psychology in Europe, 2026, Vol. 8(2), Article e15079, https://doi.org/10.32872/cpe.15079

Received: 2024-07-10 • **Accepted:** 2025-11-28 • **Published (VoR):** 2026-05-29

Handling Editor: Julia A. Glombiewski, RPTU University Kaiserslautern-Landau, Landau, Germany

Corresponding Author: Anne-Catherine I. Ewen, Centre Hospitalier Neuro-Psychiatrique CHNP, 17 Avenue des Alliés, L-9012 Ettelbruck, Luxembourg. Tel.: +352 621 728 470. E-mail: ac.ewen@outlook.com

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

Abstract

Background: Research demonstrates that mental disorders are associated with specific dysfunctional expectations and a reduced ability to adjust them, even after expectation-violating experiences. The *ViolEx* (violated expectation) model offers a framework to explain why expectations persist or change, introducing the concept of *cognitive immunization* as a potential explanation for differences in information processes. Expectation-focused psychological interventions (EFPI) aim to promote expectation adaptation.

Method: This study examines the effectiveness of an online EFPI platform for individuals with mild depressive and/or anxiety symptoms. A total of 128 participants, screened with the PHQ-9 (scores 5-9) and/or BAI (scores 8-25), were randomly assigned to one of three groups. The EFPI group and the active control group (ACG) received a psychoeducational video about expectations and their influence on behavior. Over four weeks, the EFPI group completed behavioral experiments to test their personal expectations, while the ACG used only cognitive strategies to challenge personal expectations. A third group (control group; CG) received no intervention. Surveys were administered at baseline, four weeks, and eight weeks after the initial assessment.

Results: A significant reduction in cognitive immunization was observed over measurement timepoints, with a significant difference to the CG at the follow-up. Anxiety symptoms appear to moderate this effect, whereas EFPI did not influence depressive symptoms, nor did depressive symptoms moderate changes in cognitive immunization.

Conclusion: This study is the first to evaluate online EFPI for mild depression and/or anxiety symptoms, suggesting that EFPI may reduce cognitive immunization. Future studies should investigate therapist-delivered EFPI in clinical populations with more severe symptoms.

Keywords

expectation violation, immunization, ViolEx-model, expectation focused psychological intervention EFPI, online interventions, subclinical to mild symptoms

Highlights

- Dysfunctional expectations and cognitive immunization maintain mental disorders by blocking belief/expectation updating.
- Online expectation-focused psychological interventions (EFPI) can reduce cognitive immunization over time.
- EFPI effects appear stronger in individuals with elevated anxiety symptoms.
- EFPI is promising, but more therapist-guided and clinical research is needed.

Research has increasingly focused on non-specific factors in psychotherapy, known as *common factors* (Wampold & Imel, 2015). These include concepts such as the therapeutic relationship, activation of resources, actualization of the patient's problems, motivational clarification, and problem solving, as well as alliance, empathy, expectations, cultural adaptation, and therapist differences (Wampold, 2015; Wampold & Imel, 2015). One promising factor receiving growing attention is the role of expectations in the therapeutic process (Constantino et al., 2011, 2012; Greenberg et al., 2006). It is well-established that expectations significantly influence therapy outcomes. Placebo research highlights the power of treatment expectations (Bingel, 2020; Evers et al., 2018; Kirsch, 2018; Kirsch et al., 2016; Wampold et al., 2005), and specific expectations have been shown to impact therapeutic relationships (Al-Darmaki & Kivlighan, 1993; Finsrud et al., 2022; Wright & Davis, 1994), proper self-efficacy (Bandura, 1977; Lightsey, 1999), or the effectiveness of different intervention techniques (Craske et al., 1988; Rief & Glombiewski, 2017).

In clinical psychology research, studies suggest that individuals with mental disorders not only hold more dysfunctional expectations but also struggle to adjust their expectations after disconfirming experiences (Kirchner et al., 2022; Kube, Kirchner, et al., 2019; Kube et al., 2020; Rief & Joormann, 2019). In this context, dysfunctional expectations are defined as dysfunctional thoughts that are unhelpful, distorted, idiosyncratic, and negatively biased (Lam & Cheng, 2001). Not only should the content of a certain expectation be considered, but also information processing factors should be explicitly addressed that are responsible for expectation origination, maintenance, and modification.

The concept of expectation is also prominent in action planning and decision-making models (Atkinson & Feather, 1966; Kahneman & Tversky, 2013). A similar generalized model called the ViolEx (violated expectation) model was developed, offering a frame-

work for understanding rigid expectations as core features of mental disorders (Rief et al., 2015). According to this model, generalized expectations are shaped by past experiences, social influences, and individual differences (Gollwitzer et al., 2018; Rief et al., 2015). These generalized expectations lead to situation-specific predictions, which may be confirmed or violated by experience. Healthy individuals generally adjust their expectations following disconfirming experiences. However, research suggests that in mental disorders such as depression, expectation updating may occur less frequently or require stronger disconfirmatory evidence to take place. A key mechanism for maintaining dysfunctional expectations, as proposed by the ViolEx model, is *cognitive immunization* (Gollwitzer et al., 2018; Kube, Rief, et al., 2019; Pinquart et al., 2021). This involves a reappraisal of an expectation-violating experience to preserve the original expectation, often by dismissing disconfirming events as exceptions ("*it was only an exception").* A well-known process in cognitive-behavioral theories that contributes to the persistence of dysfunctional expectations or beliefs is cognitive or behavioral avoidance (Aldao et al., 2010; Hofmann & Hay, 2018; Servatius, 2016).

In the treatment of anxiety disorders, dysfunctional expectations such as *"Something bad will happen"* or "*They will laugh at me"* are addressed explicitly through exposure therapy, which aims to reduce avoidance behavior and promote expectation adjustment (Clark, 1999; Craske et al., 1988; Craske et al., 2014). Research indicated that individuals with mental disorders exhibit not only increased dysfunctional expectations but also difficulty in accommodating these dysfunctional expectations after new expectation-disconfirming experiences (Kube, D'Astolfo, et al., 2017; Rief & Joormann, 2019). In depression, for instance, patients often struggle to change negative performance expectations even after experiencing positive outcomes (Kube, Rief, et al., 2019; Kube, Rief, & Glombiewski, 2017).

These findings suggest that persistent dysfunctional expectations and impaired adaptation processes may be central to various mental disorders. Addressing these mechanisms directly in therapy could foster therapeutic change (Craske et al., 2014). Expectation focused psychological interventions (EFPI) in psychotherapy has been proposed as a strategy to address dysfunctional expectations by fostering conscious awareness, encouraging behavioral testing of expectations, and promoting expectation-disconfirming experiences (Kube, Glombiewski, & Rief, 2019; Rief & Glombiewski, 2016).

This study investigates the effectiveness of online EFPI in reducing cognitive immunization processes in individuals with mild depressive and/ or anxious symptoms. The intervention aims to encourage participants to consciously explore how expectations influence their feelings and behaviors, as well as actively test these expectations to foster adaptive changes.

In this online randomized controlled trial, mildly depressed or anxious participants were assigned to one of three groups. The experimental group (EFPI group) received a psychoeducational video and conducted behavioral experiments to test personally

burdensome expectations over four weeks. By 'burdensome expectations,' we refer to expectations that cause psychological distress or hinder goal achievement.

The active control group (ACG) received the same psychoeducational video but was only asked to observe personal expectations over four weeks. The control group (CG) received no interventions. It was assumed that the EFPI group would experience a significant reduction in cognitive immunization over time, with the greatest reduction compared to the ACG and CG. Additionally, it was expected that symptom severity, e.g., depressive and anxiety symptom severity, would moderate the effect of the online interventions on cognitive immunization. This assumption was based on evidence suggesting that individuals with higher symptom burden may exhibit stronger cognitive immunization mechanisms, making them less responsive to interventions aimed at promoting expectation adaptation (Berg et al., 2022; Ewen et al., 2022; Kube, Kirchner, et al., 2019). Since cognitive immunization functions as a defense mechanism to maintain dysfunctional expectations, individuals with more severe symptoms may require stronger interventions or repeated exposure to expectation-violating experiences to enhance meaningful change.

Materials and Method

Participants

Recruitment

Participants were recruited through flyers displayed in public locations such as university buildings, supermarkets, general practitioner practices, pharmacies, and hospitals, as well as through social media groups, the SONA system (Research Participation System of the Philipps-University of Marburg; https://www.sona-systems.com/), and mailing lists at both Philipps-University of Marburg (students and employees) and other universities in German-speaking countries (i.e. Luxemburg and Austria). In addition, support groups for individuals with depression and anxiety disorders were contacted. Recruitment occurred from April 2021 to February 2022. For each participant, one Euro was donated to an organization working to reduce the stigma surrounding mental illness. Psychology students could receive credit points for their bachelor's degree.

Inclusion Criteria

The target sample included individuals with mild depressive and/ or anxiety symptoms. A score between 5-9 on the German version of the Patient Health Questionnaire-9 (PHQ-9; Gräfe et al., 2004) was considered indicative of mild depression. Alternatively, a score between 8-25 on the Beck Anxiety Inventory (BAI; Margraf & Ehlers, 2007), indicating mild to moderate anxiety not clinically relevant, was accepted. These cutoffs were selected for ethical reasons, as the planned behavioral experiments by participants

were not controlled by a licensed therapist. Participants were required to be at least 18 years old, have sufficient proficiency in the German language, and have access to a personal email account.

Sample Size

The literature suggests that online interventions tend to have small to medium effect sizes (Lintvedt et al., 2013; Sander et al., 2016; Zhou et al., 2016). Prior power analyses were calculated for a MANOVA with repeated measures on within- and between-subject factors, using an effect size of $f = 0.2$ and a power of .80, indicated that 152 participants would be needed. This power analysis was conducted as part of the study's preregistration (Ewen et al., 2021S) to ensure that the sample size was adequate for detecting meaningful effects given the expected variance in cognitive immunization processes. The selected parameters were based on previous research on online psychological interventions with comparable effect sizes.

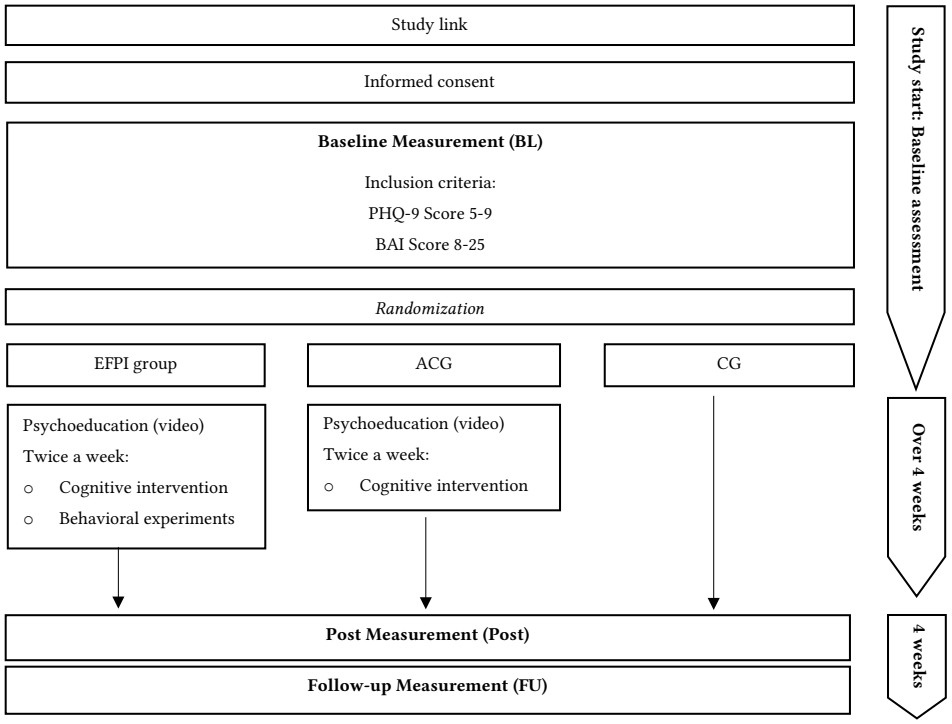
Study Design

The study was conducted entirely online using the survey platform formr (Arslan et al., 2020). At each measurement point, the questionnaires were automatically sent via email to participants through formr. Email addresses were collected at the beginning of the study and stored pseudonymously by the program. Upon clicking on the study link, participants were directed to the study information and informed consent page, which they had to agree to before proceeding. Participants were then automatically randomized by the program (Arslan et al., 2020) into one of three groups: the EFPI group, the ACG, or the CG (see below). Additionally, all the participants were asked to complete baseline questionnaires (see variables for details). After the baseline assessment, the procedures varied among the groups (see Figure 1). For the EFPI group, a psychoeducation video was presented immediately after the completion of the baseline questionnaires. This video covered the role of expectations in shaping behavior and emotions, as well as the impact of dysfunctional expectations, those that contribute to dysfunctional behaviors and long-term negative consequences. Three days later, participants were for the first time introduced to behavioral experiments, designed to test their burdensome situation-specific expectations (see Instructions, Table A1, Appendix). 'Situation-specific expectations' refer to beliefs about particular situations or events that may be unrealistic or rigid, contributing to maladaptive responses to those situations. The chosen expectation had to be very specific and testable over the following three days. Participants documented the planned experiment, their specific expectation regarding the situation, and their level of belief in this expectation (rated from 0-100%). They also recorded the associated emotion. Furthermore, they were asked to reflect on their future behavior in the given situation by considering how their actions could either confirm or disconfirm the expectation (*How can I confirm/disconfirm the expectation through my behavior?*). After three days,

participants evaluated the experiment and planned a new one. This process was repeated twice a week for four weeks.

Figure 1

Procedure of the Experiment



Note. PHQ-9 = Patient Health Questionnaire; BAI = Beck Anxiety Inventory; BL = Baseline measurement; Post = Post measurement; FU = Follow-up measurement; EFPI = expectation focused psychological interventions; ACG = active control group; CG = control group.

For the ACG, the same psychoeducation video was shown. Participants received a standardized psychoeducational intervention aimed at promoting cognitive restructuring, but without the active expectation testing component. The intervention focused on providing participants with knowledge about cognitive biases and common distortions, aiming to reduce maladaptive thinking patterns. This group was chosen to account for the effects of general psychoeducation, and we expected that any changes in expectations in this group would be less pronounced compared to the EFPI group, given the lack of behavioral experimentation. They were asked every three days, over four weeks, to document their burdensome expectations in the days between the questionnaires and

to evaluate them as helpful or unhelpful, as well as noting the associated emotions and behaviors.

The CG did not receive any psychoeducation or interim questionnaires.

After four weeks, all groups completed the post-questionnaires. A follow-up measurement was conducted four weeks later.

Materials

Psychoeducation Video

The animated video using the program Powtoon (Puspitarini et al., 2019) guided participants through the five key questions (Anne-Catherine Ewen, 2021; https://www.youtube.com/watch?v=q_L56tmckos): *What are expectations? Why do humans have expectations? How do expectations arise? Why do certain expectations remain stable? When are expectations causing problems?* The questions were addressed using the basic principles of the ViolEx model and psychoeducational content based on Rief and Glombiewski (2016). The video explained that expectations were future-directed thoughts that could be neutral, positive, or negative and interact with one's environment. They were arising automatically and were not always consciously recognized. Expectations were shaped by experience, social influences, and individual differences. The evolutionary advantage of planning behavior to prevent harm was highlighted. While the stability of functional expectations would be adaptive, the persistence of dysfunctional expectations could lead to distress by reinforcing avoidance behaviors. Avoidance and immunization processes contribute to the rigidity of expectations, making them resistant to change. Avoidance reduces exposure to corrective experiences, leading to a distorted perception of reality. Even when an expectation is violated, the brain may dismiss the event as insignificant, preventing the correction of distorted perceptions of reality. The video emphasized the importance of consciously observing negative, burdensome, dysfunctional expectations and periodically testing them.

Cognitive Intervention

The cognitive intervention was implemented in both the EFPI and ACG groups. It was based on the SORC model, commonly used in cognitive-behavioral therapy (Borg-Laufs, 2020), which facilitates a structured analysis of situations by identifying thoughts, physical sensations, emotions, behaviors, and consequences. Participants selected a burdensome expectation they had in the past three days. They were asked to identify the emotions and the physical sensations triggered by this specific expectation, describe their behavioral response, and reflect on its consequences. The cognitive intervention aimed to help participants become aware of their expectations, understand their influence on emotions and behavior, and encourage a more flexible adaptation of expectations over time. In the EFPI group, this intervention was combined with behavioral experiments to actively test and modify expectations, while in the ACG group, participants engaged only

in cognitive reflection without direct behavioral testing. The intervention was designed to reduce cognitive immunization and promote more adaptive expectation processes.

Behavioral Experiments

Behavioral experiments were conducted exclusively by the EFPI group. The instructions for the behavioral experiments were adapted from Rief and Glombiewski (2016). Participants selected a specific negative or burdensome expectation and formulated it in a precise, situation-specific manner to ensure testability over the next three days. They rated their belief in the expectation (0-100%) and identified the associated emotion. Additionally, they reflected on their potential behavior in the given situation, considering how they could act in ways that either confirmed or disconfirmed their expectation (*How can I behave to confirm my expectation? How can I behave to disconfirm my expectation?*). A key aim of the behavioral experiments was to counter cognitive immunization by actively engaging participants in testing their expectations. By requiring them to plan and execute experiments that challenge their expectations, the intervention sought to make it more difficult for participants to dismiss disconfirming experiences as mere exceptions. This structured approach was intended to promote a more flexible adaptation of expectations. After three days, they reassessed their belief in the expectation, documented any perceived evidence supporting or contradicting it, and considered alternative interpretations of the situation.

Measures

Demographics

Age, gender, nationality, education, profession, current and past diagnosed mental disorder, and current and past psychotherapeutic treatment were assessed.

Cognitive Immunization

The Immunization Scale (IMS; Ewen et al., 2022) was chosen as the primary outcome measure because it assesses key mechanisms involved in expectation persistence, making it a relevant indicator of cognitive immunization. The IMS consists of three subscales: (1) Cognitive Immunization, which measures reappraisal strategies used to maintain expectations despite disconfirming experiences, (2) Negative Expectations, capturing the persistence of negative beliefs that often accompany cognitive immunization, and (3) Assimilation, referring to the integration of expectation-violating information without fundamentally altering the expectation. While the Cognitive Immunization subscale directly measures immunization processes, the other two subscales assess related mechanisms that sustain rigid expectations, aligning with the ViolEx model. The IMS was validated through factor analyses and demonstrated excellent internal consistency (Cronbach's $\alpha = .94$). Significant correlations with depression, anxiety, and psychological flexibility

further support its construct validity. Cognitive immunization was selected as the primary outcome instead of direct expectation change because expectations are highly specific and individual, making it difficult to systematically assess whether a particular expectation was tested and updated. Instead, we focused on cognitive immunization as the assumed underlying mechanism, as it represents a generalized process by which individuals maintain their expectations despite disconfirming evidence. Depressive and anxiety symptoms were considered moderators rather than primary outcomes because the study aimed to investigate mechanisms of expectation maintenance and change rather than direct symptom reduction.

Psychopathology

Depressive symptoms were assessed using the Patient Health Questionnaire – German Version, section PHQ-9 (Gräfe et al., 2004). This 9-item scale primarily evaluates depression criteria as defined by the Diagnostic and Statistical Manual of Mental Disorders DSM-IV. Severity levels are classified as follows: a sum score of 1 to 4 indicates minimal depressive symptoms, 5 to 9 mild, 10 to 14 moderate, and 15 to 27 severe depressive symptoms. Anxiety levels were measured using the Beck Anxiety Inventory BAI (Margraf & Ehlers, 2007), a 21-item scale. A sum score of 0 to 7 indicates minimal anxiety, 8 to 15 mild, 16 to 25 moderate, and 26 to 63 severe anxiety symptoms. A sum score of 26 or higher is considered clinically relevant anxiety.

Ethics

The study was approved by the local ethics committee of the Department of Psychology at Philipps-University of Marburg (reference number 2020-84k). The study was preregistered on the Open Science Framework (https://osf.io/kvuj7).

Statistical Analyses

All analyses were conducted using RStudio version 1.2.5042 (Posit PBC, 2009-2020). To account for missing values at the second and third assessment time points, incomplete sum scores of the dependent variables were estimated using multiple imputation (Donders et al., 2006) with the MICE package (van Buuren & Groothuis-Oudshoorn, 2011).

Before conducting the analyses, the Mahalanobis distance was calculated and compared against a χ^2-cut-off of $\alpha = .001$, revealing no outliers. Mixed-effects analyses were performed using the *lme4* package (Bates et al., 2018) and *lmerTest* (Kuznetsova et al., 2015). The interaction term *time*group* was included as a predictor, while participant intercepts and slopes were included as random effects, allowing trends over time to vary for each participant. Analyses of variance ANOVA were calculated for an overview of fixed effects. The package emmeans (Lenth et al., 2018) was used to compute contrasts.

The results of the mixed linear models (LLM) were provided by the output of the *tab_model* function of the *sjPlot* package (Lüdecke, 2024). For moderator analyses, moderator variables were included as a triple interaction into the model. The models with and without the moderating term were compared using the Chi-square difference test. Homoscedasticity and normality were checked through residual plots, which exhibited the expected pattern. Due to a high number of missing values, the number of imputations was set to 50 (Bodner, 2008; White et al., 2011). The missing values were assumed to be missing at random (MAR), meaning that their occurrence was independent of the values themselves. The baseline assessment had no missing values (n = 1,218), while the second and third assessment time points had 54% and 66% missing values, respectively.

Results

Sample Characteristics

A total of 543 individuals agreed to the informed consent. However, 230 participants were excluded due to incomplete baseline assessments, and 185 participants did not meet the inclusion criteria, resulting in a final sample of 128 participants. Among these, 41 participants were allocated to the EFPI group, 38 to the ACG, and 49 to the CG (see Figure 2).

The mean age of the sample was 27 (SD = 11.46). Most of the participants were female (82.81%), while 16.41% identified as male and 0.78% as other. Most participants (91.41%) held German nationality. Higher education, defined as holding a university degree, was reported by 24.22% of the sample, while 78.13% were current university students. A history of diagnosed mental disorders was reported by 21.88% of participants, with 8.59% indicating a current diagnosis. Furthermore, 33.59% reported having received psychotherapy at some point in their lives, and 12.50% were engaged in psychotherapeutic treatment during the study period. The means and standard deviations of the assessed questionnaires, comparing imputed and non-imputed data for the post and follow-up measurements, are presented in Table 1.

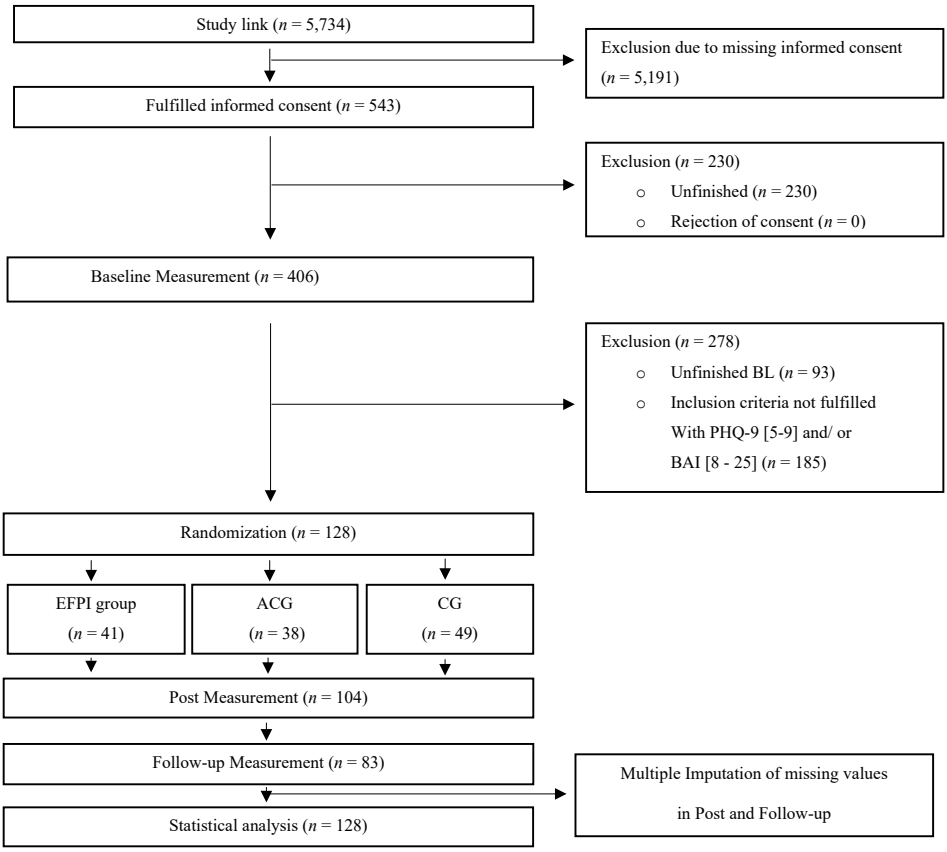
Cognitive Immunization

An ANOVA of the mixed-effects model revealed a significant main effect of timepoint (F(2, 122) = 19.96, p < .001), whereas the main group (F(2,122) = 2.21, p = .114) was not significant. The interaction of timepoint and group (F(4,122) = 2.12, p = .082) was not significant. The interaction plot can be seen in Figure 3.

The results of the linear mixed-effects model further revealed the main significant effect of timepoint, indicating a decrease in the IMS score at the post and follow-up measurement compared to the baseline measurement (post: β = -6.05, 95% CI [-9.90, -2.19], p = .002; follow-up: β = - 11.64, 95% CI [-16.00, -7.28], p < .001). No significant

Figure 2

Flow Chart



Note. *n* = number of participants; PHQ-9 = Patient Health Questionnaire; BAI = Beck Anxiety Inventory; BL = Baseline measurement; Post = Post measurement; FU = Follow-up measurement; EFPI = expectation focused psychological interventions; ACG = active control group; CG = control group.

effect was found for the group variable (ACG: β = -0.18, 95% CI [-6.26, 5.89], p = .953; CG: β = 1.97, 95% CI [-3.74, 7.67], p = .498), nor were any significant interactions observed between the timepoint and group, while a significant interaction emerged for the follow-up timepoint and the CG (β = 7.43, 95% CI [1.52, 13.35], p = .014; see Table 2) indicating that the change in the IMS score over the timepoints differs significantly between the control group and the other groups.

Table 1

Demographics: Mean and Standard Deviations of Different Variables for the Assessment Points Post and Follow-Up

Variable	Total Sample $N = 128$	EFPI $n = 41$	ACG $n = 38$	CG $n = 49$
Age (*M*, *SD*)	28.88 (13.00)	28.07 (12.83)	26.89 (9.89)	31.10 (15.02)
Gender (F, M, N)[a]	106 / 21 / 1	32 / 8 / 1	34 / 4 / 0	40 / 9 / 0
PHQ-9$_{Baseline}$ (*M*, *SD*)	6.27 (1.53)	6.54 (1.34)	6.03 (1.53)	6.24 (1.65)
PHQ-9$_{Post}$ (*M*, *SD*)	6.81 (3.03)	6.71 (2.60)	6.18 (3.20)	6.35 (2.95)
PHQ-9$_{Follow-up}$ (*M*, *SD*)	6.04 (3.04)	5.61 (2.74)	5.89 (3.17)	6.51 (3.16)
BAI$_{Baseline}$ (*M*, *SD*)	10.23 (5.52)	9.37 (5.71)	10.82 (5.28)	10.51 (5.49)
BAI$_{Post}$ (*M*, *SD*)	10.85 (7.89)	11.00 (8.39)	9.29 (6.53)	9.55 (8.40)
BAI$_{Follow-up}$ (*M*, *SD*)	7.89 (7.06)	6.49 (6.49)	7.53 (6.39)	8.98 (7.96)
IMS$_{Baseline}$ (*M*, *SD*)	56.44 (13.45)	55.56 (14.13)	55.53 (12.52)	57.88 (13.73)
IMS$_{Post}$ (*M*, *SD*)	51.31 (13.15)	49.51 (11.84)	50.84 (12.31)	53.18 (14.75)
IMS$_{Follow-up}$ (*M*, *SD*)	48.57 (14.05)	43.88 (10.27)	47.08 (13.40)	53.65 (15.78)

Note. *N* = sample size; *M* = mean; *SD* = standard deviation; PHQ-9 = Patient Health Questionnaire-9; BAI = Beck Anxiety Inventory; IMS = Immunization Scale; EFPI = Expectation Focused Psychological Intervention; ACG = active control group; CG = control group.

[a]F = female, M = male, N = neutral numbers of participants of each group.

Random effects analysis showed that the variance in the intercepts across sessions (τ00) was 132.35, and the variance in the slopes (τ11) for the post timepoint and the follow-up timepoint was 51.86 and 94.79, respectively. The intraclass correlation coefficient (ICC) was 0.71, indicating that 71% of the total variability in the outcome was attributable to differences between groups and not within individuals. The model's conditional R^2 was 0.735, suggesting that the model explains a substantial proportion of the variability in the outcome and indicating the importance of the grouping structure, while the marginal R^2 was 0.084, indicating that the fixed effects alone explain a smaller proportion of the variability.

Psychopathology

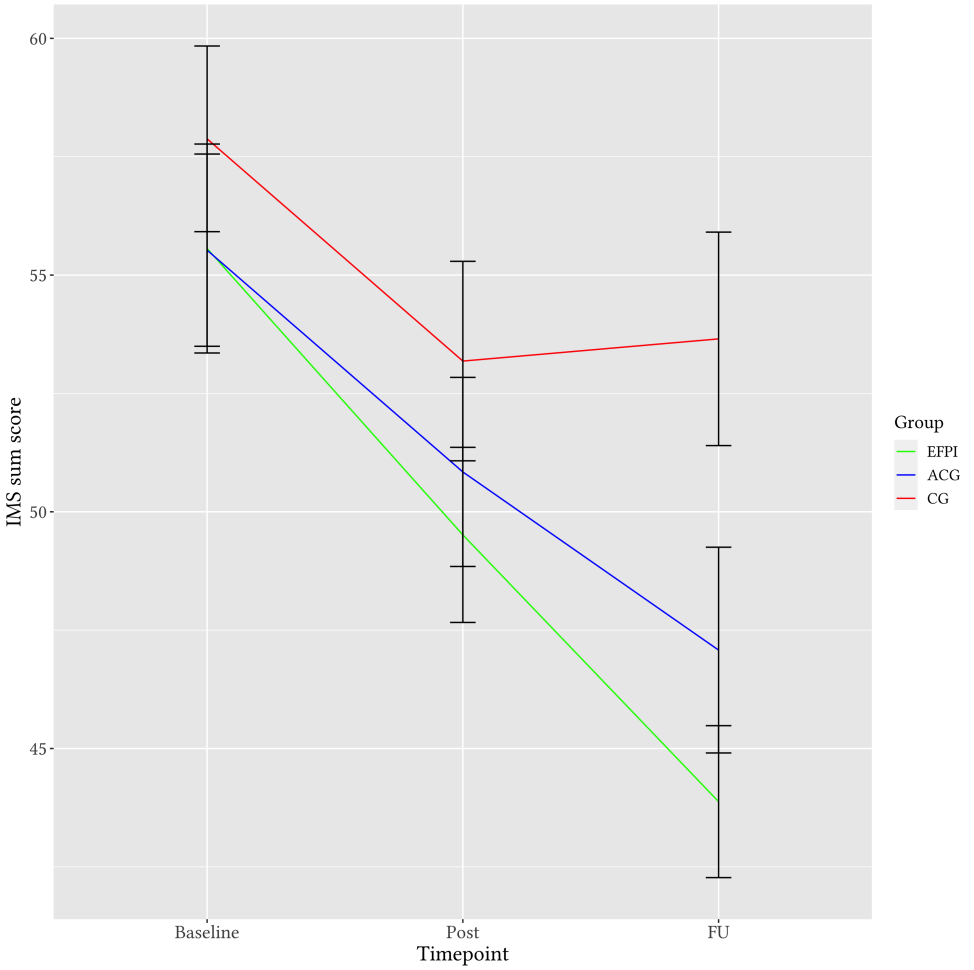
Depressiveness PHQ-9

An ANOVA of the mixed-effects model revealed no significant main (timepoint: $F(2,121) = 0.18$, $p = .311$; group: $F(2,122) = 0.20$, $p = .820$) or interaction ($F(4,121) = 1.23$, $p = .300$) effects (see Figure 4).

The results of the linear mixed-effects model revealed no significant effects of timepoint or group, as well as no significant interactions could be found (see Table 3). The PHQ-9 score did not significantly vary between the groups or measurement timepoints.

Figure 3

Interaction Plot Mapped for the Immunization Scale



Note. Figure 3 shows the Immunization Level measured by the IMS over the three measurement time points. The results of the mixed models indicate only one significant interaction at the follow-up timepoint between the groups. IMS = Immunization Scale; Baseline = Baseline measurement; Post = Post measurement; FU = Follow up measurement; EFPI = expectation focused psychological interventions; ACG = active control group; CG = control group.

Random effects analysis showed that the intercept variance (τ00) was 0.77. The variance in the slopes (τ11) for the post timepoint was 4.41, and for the follow-up timepoint, 5.54. The ICC was 0.77, suggesting that most variability in the PHQ-9 scores

Table 2

Output of Mixed Model Including Immunization Scale, Timepoint and Group

Predictors	Estimates	95% CI	*p*
Intercept	55.59	51.38 – 59.80	**< .001**
Post	-6.05	-9.90 – -2.19	**.002**
Follow-up	-11.64	-16.00 – -7.28	**< .001**
ACG	-0.18	-6.26 – 5.89	.953
CG	1.97	-3.74 – 7.67	.498
Post * ACG	1.61	-3.95 – 7.17	.569
Follow-up * ACG	3.65	-2.65 – 9.94	.255
Post * CG	1.38	-3.84 – 6.61	.603
Follow-up * CG	7.43	1.52 – 13.35	**.014**
Random Effects			
σ^2	52.01		
τ_{00}	132.35		
τ_{11Post}	51.86		
$\tau_{11Follow\text{-}up}$	94.79		
ρ01	-0.39		
	-0.43		
ICC	0.71		
N	125		
Observations	384		
Marginal R^2/Conditional R^2	0.084/0.735		

Note. CI = confidence interval; ACG = active control group; CG = control group; σ^2 = residual variance; τ_{00} = random intercept variance; τ_{11} = random slope variance; ρ01 = correlation between random intercept and slope; ICC = interclass correlation coefficient; *N* = sample size; significant *p*-values ($p \leq .05$) are shown in bold.

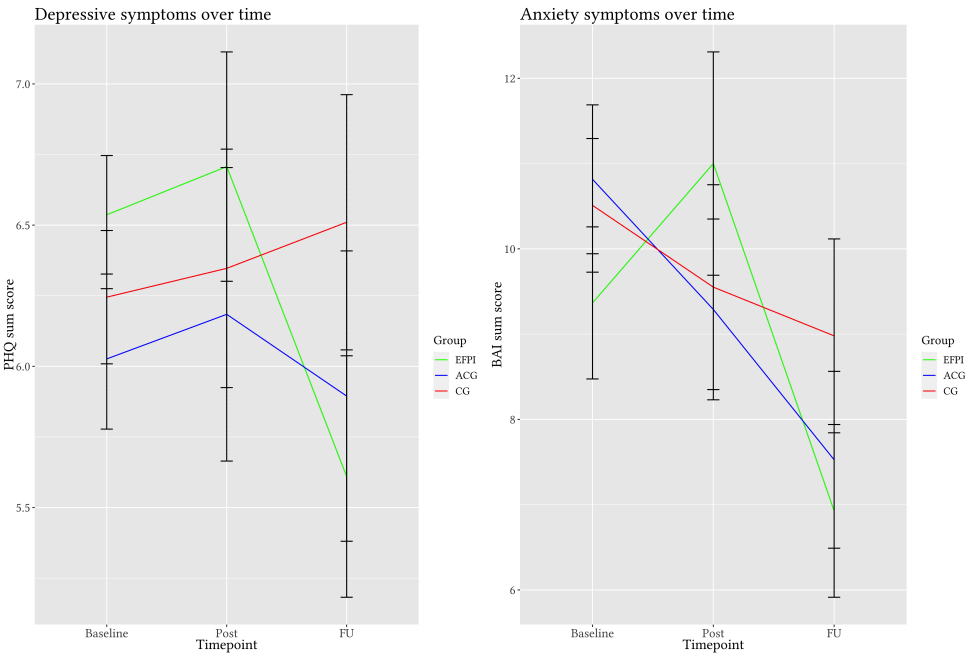
is due to between-group differences. The model's conditional R^2 was 0.769, suggesting that the model explains a substantial proportion of the variability in the outcome, while the marginal R^2 was 0.014, indicating that the fixed effects alone explain only a small proportion of the variability.

Anxiety BAI

For the calculated ANOVA, a significant main effect for timepoint could be found ($F(2,125) = 14.46$, $p < .001$). The main effect group ($F(2,122) = 0.20$, $p = .819$) was non-significant, the interaction effect ($F(4,125) = 2.89$, $p = .025$) showed a significant result. Contrast analyses looking at the change scores per group, only the EFPI group showed a significant difference between post and follow-up measurement ($t(119) = 5.03$, $p < .001$; see interaction plot Figure 4).

Figure 4

Interaction Plot Mapped for the Depressive and Anxiety Symptoms



Note. No significant results could be found for the depressive symptoms over the three measurement timepoints (left). A significant interaction could be found for the anxiety symptoms over the three measurement timepoints (right). PHQ-9 = Patient Health Questionnaire; BAI = Beck Anxiety Inventory; Baseline = Baseline measurement; Post = Post measurement; FU = Follow up measurement; EFPI = expectation focused psychological interventions; ACG = active control group; CG = control group.

The results of the linear mixed-effects model further revealed the main significant effect of the follow-up timepoint (β = - 2.55, 95% CI [-4.75, -0.36], p = .023). No significant effect was found for the group variable (see Table 4), nor were any significant interactions observed between the timepoint and group.

The random intercept variance (τ00) was 18.28, the random slope variance (τ11) at the post timepoint and the follow-up timepoint was 30.49 and 24.84, respectively. The total variability was attributable to 74% of the differences between the groups. The model's conditional R^2 was 0.744, the marginal R^2 was 0.034.

Table 3

Output of Mixed Model Including Patient Health Questionnaire, Timepoint and Group

Predictors	Estimates	95% CI	*p*
Intercept	6.53	6.06 – 7.00	**< .001**
Post	0.14	-0.71 – 0.99	.744
Follow-up	-0.91	-1.82 – 0.00	.051
ACG	-0.49	-1.17 – 0.19	.159
CG	-0.29	-0.93 – 0.35	.378
Post * ACG	0.08	-1.15 – 1.31	.901
Follow-up * ACG	0.81	-0.51 – 2.12	.231
Post * CG	-0.07	-1.23 – 1.08	.902
Follow-up * CG	1.17	-0.07 – 2.41	.065
Random effects			
σ^2	1.59		
τ_{00}	0.77		
τ_{11Post}	4.41		
$\tau_{11Follow\text{-}up}$	5.54		
ρ01	0.51		
	0.35		
ICC	0.77		
N	125		
Observations	384		
Marginal R^2/Conditional R^2	0.014/0.769		

Note. CI = confidence interval; ACG = active control group; CG = control group; σ^2 = residual variance; τ_{00} = random intercept variance; τ_{11} = random slope variance; ρ01 = correlation between random intercept and slope; ICC = interclass correlation coefficient; *N* = sample size; significant *p*-values ($p \leq .05$) are shown in bold.

Moderations

IMS and PHQ-9

The model integrating the depressive symptoms as a moderator (AIC 3013.2) did not explain significantly more variance than the model without a moderator (AIC 3016.5; $\chi^2(9) = 16.28$, $p = .061$).

In the triple interaction model, ANOVA calculations showed a significant main effect of timepoint ($F(2,123) = 4.68$, $p = .011$) and a significant interaction of timepoint and PHQ-9 ($F(2,123) = 3.48$, $p = .034$). The triple interaction was not significant ($F(4,123) = 3.48$, $p = .417$). The results of the linear mixed-effects model revealed no significant main effects or interactions (see Table 5).

Table 4

Output of Mixed Model Including the Beck Anxiety Inventory Timepoint and Group

Predictors	Estimates	95% CI	*p*
Intercept	9.41	7.68 – 11.13	**< .001**
Post	1.51	-0.80 – 3.83	.199
Follow-up	-2.55	-4.75 – -0.36	**.023**
ACG	1.45	-1.03 – 3.94	.251
CG	1.20	-1.14 – 3.53	.314
Post * ACG	-2.98	-6.32 – 0.36	.080
Follow-up * ACG	-0.69	-3.85 – 2.47	.668
Post * CG	-2.40	-5.53 – 0.73	.133
Follow-up * CG	1.07	-1.90 – 4.04	.479
Random effects			
σ^2	12.67		
τ_{00}	18.28		
τ_{11Post}	30.49		
$\tau_{11Follow-up}$	24.84		
ρ01	0.03		
	-0.14		
ICC	0.74		
N	125		
Observations	384		
Marginal R^2/Conditional R^2	0.034/0.744		

Note. CI = confidence interval; ACG = active control group; CG = control group; σ^2 = residual variance; τ_{00} = random intercept variance; τ_{11} = random slope variance; ρ01 = correlation between random intercept and slope; ICC = interclass correlation coefficient; *N* = sample size; significant *p*-values ($p \leq .05$) are shown in bold.

Random effects analysis showed that the variance in the intercepts across sessions (τ00) was 131.26, and the variance in the slopes (τ11) for the post and the follow-up timepoint was 54.11 and 102.20, respectively. The ICC was 0.71. The model's conditional R^2 was 0.749, and the marginal R^2 was 0.123.

IMS and BAI

Integrating anxiety as a moderator in the mixed model (interaction model: AIC 2981.6, triple interaction model: AIC 2974.1) explained the data significantly better ($\chi^2(9) = 25.45$, $p = .003$). The ANOVA revealed a significant main effect in group ($F(2,119) = 6.23$, $p = .002$). The interaction between timepoint x BAI ($F(2,122) = 6.43$, $p = .002$) and group x BAI ($F(2,119) = 4.27$, $p = .016$) were significant. The triple interaction was non-significant ($F(4,122) = 0.40$, $p = .806$).

Table 5

Output of Mixed Model Including Immunization Scale, Timepoint and Group and Patient Health Questionnaire as Moderator

Predictors	Estimates	95% CI	*p*
Intercept	56.30	35.49 – 77.11	**< .001**
Post	8.61	-10.60 – 27.82	.379
Follow-up	-16.79	-38.78 – 5.21	.134
ACG	-12.80	-40.10 – 14.50	.357
CG	-11.50	-37.13 – 14.13	.378
PHQ-9	-0.11	-3.23 – 3.01	.945
Post * ACG	-3.01	-28.22 – 22.19	.814
Follow-up * ACG	9.44	-19.42 – 38.29	.520
Post * CG	-14.23	-37.89 – 9.43	.238
Follow-up * CG	9.70	-17.39 – 36.79	.482
Post * PHQ-9	-2.24	-5.12 – 0.63	.127
Followup * PHQ-9	0.79	-2.51 – 4.09	.639
ACG * PHQ-9	2.08	-2.14 – 6.31	.332
CG * PHQ-9	2.15	-1.74 – 6.04	.277
Post * ACG * PHQ-9	0.58	-3.32 – 4.48	.770
Follow-up * ACG * PHQ-9	-0.89	-5.35 – 3.57	.695
Post * CG * PHQ-9	2.40	-1.19 – 5.98	.190
Follow-up * CG * PHQ-9	-0.33	-4.44 – 3.78	.876
Random effects			
σ^2	50.36		
τ_{00}	131.26		
τ_{11Post}	54.11		
$\tau_{11Follow-up}$	102.20		
ρ01	-0.40		
	-0.46		
ICC	0.71		
N	125		
Observations	384		
Marginal R^2/Conditional R^2	0.123/0.749		

Note. CI = confidence interval; ACG = active control group; CG = control group; PHQ-9 = Patient Health Questionnaire-9; σ^2 = residual variance; τ_{00} = random intercept variance; τ_{11} = random slope variance; ρ01 = correlation between random intercept and slope; ICC = interclass correlation coefficient; *N* = sample size; significant *p*-values ($p \leq .05$) are shown in bold.

The results of the linear mixed effects model further revealed the main significant effect of the CG, indicating an increase in the IMS score in the third group compared to the EFPI group (CG: β = 14.52, 95% CI [3.47, 35.57], p = .010). Higher BAI scores were significantly associated with higher IMS scores (β = 1.05, 95% CI [0.35, 1.75], p = .003). A significant interaction between the follow-up timepoint and the BAI score was found (β = -0.87, 95% CI [-1.61, -0.13], p = .021), indicating that the relationship between the BAI score and the IMS score differs significantly at the follow-up timepoint compared to the other timepoints. Another significant interaction was found between the third group, CG, and the BAI score (β = -1.30, 95% CI [-2.26, -0.33], p = .008; see Table 6), indicating that the relationship of the BAI score and the IMS score differs significantly in the CG compared to the other groups.

The random intercept variance (τ00) was 113.50, and the random slope variance (τ11) was for the post timepoint and the follow-up timepoint 54.34 and 85.34, respectively. The ICC was 0.71, indicating that 71% of the total variability in the outcome could be explained through the grouping factors and not individual variance. The model's conditional R^2 was 0.752, suggesting that the model explains a substantial proportion of the variability in the outcome, while the marginal R^2 was 0.148.

Discussion

The primary aim of this study was to evaluate the effectiveness of expectation-focused online interventions in reducing cognitive immunization. The first hypothesis, which proposed a greater reduction in immunization in the EFPI group among the participants with mild symptom severity, can be interpreted as only partially supported. The findings indicate a significant effect of the timepoint on IMS scores with reductions observed at both post-measurement and follow-up. However, no significant main effect of group was found. The results of the mixed models only showed one significant interaction at the follow-up measurement for the CG, suggesting that changes in the IMS scores at the follow-up timepoint differed between the CG and the other groups.

The lack of statistical significance of the interaction between group x timepoint after the ANOVA calculations, as well as the interactions between the ACG and the timepoints in the mixed model results, may be attributed to the smaller sample size than suggested by the a priori power analyses, and the nature of the online intervention, which limits control over participants' adherence to the interventions. Further studies could address this limitation by incorporating phone consultations with study therapists to enhance adherence. However, the findings are in line with other studies (Kube, Glombiewski, Gall, et al., 2019; Kube, Glombiewski, & Rief, 2019; Rief & Joormann, 2019).

Table 6

Output of Mixed Model Including Score of the Immunization Scale, Timepoint and Group, and Beck Anxiety Inventory as Moderator

Predictors	Estimates	95% CI	*p*
Intercept	45.67	37.99 – 53.35	**< .001**
Post	-0.92	-8.35 – 6.51	.808
Follow-up	-3.43	-11.59 – 4.72	.408
ACG	-3.56	-15.62 – 8.50	.562
CG	14.52	3.47 – 25.57	**.010**
BAI	1.05	0.35 – 1.75	**.003**
Post * ACG	2.99	-8.68 – 14.65	.615
Follow-up * ACG	6.22	-6.58 – 19.03	.340
Post * CG	-3.14	-13.81 – 7.53	.563
Follow-up * CG	4.65	-7.07 – 16.36	.436
Post * BAI	-0.55	-1.22 – 0.13	.114
Follow-up * BAI	-0.87	-1.61 – -0.13	**.021**
ACG * BAI	0.117	-0.86 – 1.21	.742
CG * BAI	-1.30	-2.26 – -0.33	**.008**
Post * ACG * BAI	-0.06	-1.06 – 0.95	.913
Follow-up * ACG * BAI	-0.12	-1.23 – 0.98	.826
Post * CG * BAI	0.49	-0.45 – 1.42	.304
Follow-up * CG * BAI	0.36	-0.67 – 1.38	.492
Random effects			
σ^2	49.57		
τ_{00}	113.50		
τ_{11Post}	54.34		
$\tau_{11Follow\text{-}up}$	85.34		
ρ01	-0.31		
	-0.35		
ICC	0.71		
N	125		
Observations	384		
Marginal R^2/Conditional R^2	0.148/0.752		

Note. CI = confidence interval; ACG = active control group; CG = control group; PHQ = Patient Health Questionnaire-9; BAI = Beck Anxiety Inventory; σ^2 = residual variance; τ_{00} = random intercept variance; τ_{11} = random slope variance; ρ01 = correlation between random intercept and slope; ICC = interclass correlation coefficient; *N* = sample size; significant *p*-values ($p \leq .05$) are shown in bold.

While no significant results were found in the model including depressive symptoms (PHQ-9), the EFPI appeared to have an influence on anxiety symptoms. Looking at the model incorporating anxiety symptoms (BAI), ANOVA calculations revealed a significant main effect of timepoint, and a significant interaction between timepoint and group. Further contrast analyses indicated a significant difference between post-measurement and follow-up within the EFPI group. The results of the mixed models showed as well a significant reduction of the BAI score at the follow-up timepoint. This suggests that EFPI may contribute to a significant reduction in anxiety symptoms. However, no significant interactions between timepoint and group were found. This indicates that after accounting for individual differences (random effects), the interaction effect does not remain strong enough. Further studies with bigger sample sizes should clarify the effects of EFPI on anxiety symptoms.

Moderator analyses further supported the findings that EFPI influences anxiety symptoms. The model, including the BAI, could significantly better explain the data than the model without the BAI as a moderator. Looking at the results of the mixed models, significant moderations could be found at the follow-up timepoint and the third group (CG). Notably, at the follow-up timepoint the relationship between BAI and IMS scores was significantly different, indicating that this relationship varied over time. Additionally, the relationship between BAI and IMS scores was significantly smaller in the CG group compared to other groups. This is in line with previous research on the effectiveness of behavioral experiments in anxiety disorders and the role of exposure in dysfunctional expectations (Craske et al., 2014; McMillan & Lee, 2010).

Overall, differences between the EFPI group and the CG could be found, as well as differences between baseline measurements and follow-up. Moreover, anxiety levels seem to have an influence on the IMS levels and the effect of EFPI on IMS scores. The reduction in immunization at the four-week follow-up suggests that cognitive (ACG) and experience-based (EFPI) interventions may promote enduring changes in information-processing mechanisms. Through a continuous updating process, the discrepancy between preexisting assumptions and reality may decrease - a process that is often impaired in individuals with mental health conditions (Berg et al., 2022). In other words, these interventions may enhance cognitive flexibility by fostering adaptive updating mechanisms, allowing individuals to consider alternative situational interpretations (Fröber et al., 2018; Liknaitzky et al., 2017, 2018; Meiran et al., 2011).

Strengths and Limitations

This study proposes and tests expectation focused psychological interventions EFPI in a longitudinal online format, offering an economical approach. Until now, cognitive immunization has primarily been examined through complex experimental designs. This study provides initial evidence of the effectiveness of EFPI in reducing immunization over time using a simple self-report questionnaire in a naturalistic setting (Ewen et al.,

2022). Additionally, the study contributes to the validation of the IMS, demonstrating that immunization can be influenced following specific interventions. The results suggest that cognitive immunization remains rather stable over time in the CG, highlighting the potential of the interventions. Another advantage of the study design is that it provided implicit support during an uncertain period during the COVID-19 pandemic. Moreover, the findings suggest that EFPI may have a beneficial effect on subclinical or mild anxiety symptoms.

However, the results should be interpreted with caution due to several limitations. A substantial number of participants were lost to post and follow-up assessments, likely due to the lack of therapeutic contact in this online program, which may have facilitated early dropout. Additionally, as participants were contacted via email, the study had limited control over questionnaire completion and participant accessibility. Consequently, data interpretation should be approached with caution, particularly given the need for missing data imputation, even though this method is already well established (Enders, 2017). Another limitation is the restrictive inclusion criteria, which resulted in a relatively small sample size. Additionally, an unequal distribution of female and male participants occurred, with a higher proportion of female participants. This imbalance in gender distribution is common in psychological research, particularly in mental health and online intervention studies, where female individuals tend to be overrepresented (Richmond et al., 2015). While our analyses do not suggest gender-related effects, future studies should aim for a more balanced sample to ensure broader generalizability.

Future Research and Practical Implications

Future research should include participants with diagnosed mental disorders to evaluate the use of EFPI in psychotherapy over the course of structured Cognitive Behavioral Therapy (Ewen et al., 2023; Wilhelm et al., 2022). Furthermore, future research should explore the potential moderating effects in more depth (i.e., anxiety symptoms). To better understand the underlying mechanisms of cognitive immunization, it would be valuable to investigate cognitive flexibility, as examined in previous studies (Fröber et al., 2018; Liknaitzky et al., 2018). Additionally, future research should explore the connection between expectation management and predictive coding approaches, which is an emerging field of study (Kube et al., 2022). Another important question is whether different mental disorders exhibit similar patterns of increased cognitive immunization and whether EFPI could be effective in these populations as well.

Regarding practical implications, this study could support the integration of EFPI into psychotherapy (Kube, Glombiewski, & Rief, 2019; Rief & Glombiewski, 2016), whereby the influence of EFPI on more severe psychopathology could be analyzed. The interventions appear to have a long-term impact. By actively addressing persistent and dysfunctional expectations in psychotherapy, clinicians may help patients modify rigid cognitive

patterns in general. EFPI provides a structured approach to seemingly reduce cognitive immunization, thereby fostering greater flexibility in expectation adaptation.

Conclusion

This study is the first to evaluate the effectiveness of expectation-focused psychological interventions over a two-month period. Delivered online without offering therapeutic support, these interventions showed an ability to reduce cognitive immunization processes. Since cognitive immunization contributes to the persistence of dysfunctional expectations, reducing these processes may lead to greater adaptability to situational experiences, particularly in individuals with mild depressive and/or anxiety symptoms. Furthermore, the findings suggest a certain relationship between cognitive immunization and anxiety symptoms, with a possible link between the reduction of cognitive immunization and improvement in anxiety symptoms that should be tested in further studies.

Funding: This project was financially supported by the Philipps-University Marburg by Quality assurance of teaching (QSL – Qualitätssicherung der Lehre).

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: *Anne-Catherine I. Ewen:* Conceptualization, Methodology, Software, Formal Analysis, Writing – Original draft, Project Administration. *Marcel Wilhelm:* Conceptualization, Writing – Review & Editing, Supervision.

Ethics Statement: The local ethics committee of the Department of Psychology, Philipps-University Marburg approved the study (Reference Number: 2020-84k).

Preregistration: This study has been preregistered on the Open Science Framework OSF (https://osf.io/kvuj7).

Related Versions: This article is derived from the doctoral thesis of the first author. Some overlap in content is expected, as the thesis served as the basis for the present manuscript. The thesis is publicly accessible (Ewen, 2022). The current article has been revised and adapted for journal publication.

Reporting Guidelines: This study applied the CONSORT-Statement Guidelines.

Data Availability: Materials and the statistical code supporting this study are available from the corresponding author upon reasonable request. Data sharing may be subject to ethical, legal, or confidentiality restrictions, in line with applicable regulations and participant consent.

Supplementary Materials

The Supplementary Materials contain the following items:

- *Preregistration* (Ewen et al., 2021S)
- *Appendix:* The appendix includes the instructions of the expectation testing experiments and the table showing the differences between imputed and non-imputed data (Ewen & Wilhelm, 2026S).

Index of Supplementary Materials

Ewen, A.-C. I., Rief, W., & Wilhelm, M. (2021S). *Promote flexibility in expectations (FLEX)* [Preregistration]. OSF Registries. https://doi.org/10.17605/OSF.IO/KVUJ7

Ewen, A.-C. I., & Wilhelm, M. (2026S). *Supplementary materials to "Promoting flexibility in expectations: A randomized-controlled online-intervention study for mild psychopathological symptoms"* [Appendix]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.21808

References

Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review, 30*(2), 217–237. https://doi.org/10.1016/j.cpr.2009.11.004

Al-Darmaki, F., & Kivlighan, D. M. (1993). Congruence in client-counselor expectations for relationship and the working alliance. *Journal of Counseling Psychology, 40*(4), 379–384. https://doi.org/10.1037/0022-0167.40.4.379

Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods, 52*(1), 376–387. https://doi.org/10.3758/s13428-019-01236-y

Atkinson, J. W., & Feather, N. T. (1966). *A theory of achievement motivation* (Vol. 66). John Wiley & Sons.

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

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Dai, B., Scheipl, F., Grothendiek, G., Green, P., Fox, J., Bauer, A., Krivitsky, P., Tamaka, E., Jagan, M., Boylan, R., & Ly, A. (2018). *Package 'lme4'* (Version 1(17), 437). [Computer software].

Berg, M., Feldmann, M., & Kirchner, L., & Kube, T. (2022). Oversampled and undersolved: Depressive rumination from an active inference perspective. *Neuroscience and Biobehavioral Reviews, 142*, Article 104873. https://doi.org/10.1016/j.neubiorev.2022.104873

Bingel, U. (2020). Placebo 2.0: The impact of expectations on analgesic treatment outcome. *Pain, 161*, S48–S56. https://doi.org/10.1097/j.pain.0000000000001981

Bodner, T. E. (2008). What improves with increased missing data imputations? *Structural Equation Modeling, 15*(4), 651–675. https://doi.org/10.1080/10705510802339072

Borg-Laufs, M. (2020). *Die funktionale Verhaltensanalyse: Ein praktischer Leitfaden für Psychotherapie, Sozialarbeit und Beratung* [Functional behavior analysis: A practical guide for psychotherapy, social work, and counseling]. Springer. https://doi.org/10.1007/978-3-658-30812-4

Clark, D. M. (1999). Anxiety disorders: Why they persist and how to treat them. *Behaviour Research and Therapy, 37*(Suppl. 1), S5–S27. https://doi.org/10.1016/S0005-7967(99)00048-0

Constantino, M. J., Ametrano, R. M., & Greenberg, R. P. (2012). Clinician interventions and participant characteristics that foster adaptive patient expectations for psychotherapy and psychotherapeutic change. *Psychotherapy, 49*(4), 557–569. https://doi.org/10.1037/a0029440

Constantino, M. J., Arnkoff, D. B., Glass, C. R., Ametrano, R. M., & Smith, J. Z. (2011). Expectations. *Journal of Clinical Psychology, 67*(2), 184–192. https://doi.org/10.1002/jclp.20754

Craske, M. G., Rapee, R. M., & Barlow, D. H. (1988). The significance of panic-expectancy for individual patterns of avoidance. *Behavior Therapy, 19*(4), 577–592. https://doi.org/10.1016/S0005-7894(88)80025-X

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Donders, A. R. T., van der Heijden, G. J. M. G., Stijnen, T., & Moons, K. G. M. (2006). Review: A gentle introduction to imputation of missing values. *Journal of Clinical Epidemiology, 59*(10), 1087–1091. https://doi.org/10.1016/j.jclinepi.2006.01.014

Enders, C. K. (2017). Multiple imputation as a flexible tool for missing data handling in clinical research. *Behaviour Research and Therapy, 98*, 4–18. https://doi.org/10.1016/j.brat.2016.11.008

Evers, A. W., Colloca, L., Blease, C., Annoni, M., Atlas, L. Y., Benedetti, F., Bingel, U., Büchel, C., Carvalho, C., Colagiuri, B., Crum, A., Enck, P., Gaab, J., Geers, A. L., Howick, J., Jensen, K., Krisch, I., Meissner, K., Napadow, V., . . . Kelly, J. M. (2018). Implications of placebo and nocebo effects for clinical practice: Expert consensus. *Psychotherapy and Psychosomatics, 87*(4), 204–210. https://doi.org/10.1159/000490354

Ewen, A.-C. I. (2021, March 12). *ViolEx-Modell* [Video]. YouTube. https://www.youtube.com/watch?v=q_L56tmckos

Ewen, A.-C. I. (2022). *The role of dysfunctional expectation persistence in psychopathology* [Doctoral dissertation, Philipps-University of Marburg]. open_UMR. https://doi.org/10.17192/z2023.0054

Ewen, A.-C. I., Bleichhardt, G., Rief, W., Von Blanckenburg, P., Wambach, K., & Wilhelm, M. (2023). Protocol: Expectation focused and frequency enhanced cognitive behavioural therapy for patients with major depression (EFFECT): A study protocol of a randomised active-control trial. *BMJ Open, 13*(3), Article e065946. https://doi.org/10.1136/bmjopen-2022-065946

Ewen, A.-C. I., Rief, W., & Wilhelm, M. (2022). Exploring the path of persisting dysfunctional expectations—Development of the immunization scale IMS. *Frontiers in Psychology, 13*, Article 1033078. https://doi.org/10.3389/fpsyg.2022.1033078

Finsrud, I., Nissen-Lie, H. A., Vrabel, K., Høstmælingen, A., Wampold, B. E., & Ulvenes, P. G. (2022). It's the therapist and the treatment: The structure of common therapeutic relationship factors. *Psychotherapy Research, 32*(2), 139–150. https://doi.org/10.1080/10503307.2021.1916640

Fröber, K., Raith, L., & Dreisbach, G. (2018). The dynamic balance between cognitive flexibility and stability: The influence of local changes in reward expectation and global task context on voluntary switch rate. *Psychological Research, 82*(1), 65–77. https://doi.org/10.1007/s00426-017-0922-2

Gollwitzer, M., Thorwart, A., & Meissner, K. (2018). Editorial: Psychological responses to violations of expectations. *Frontiers in Psychology, 8*, Article 2357. https://doi.org/10.3389/fpsyg.2017.02357

Gräfe, K., Zipfel, S., Herzog, W., & Löwe, B. (2004). Screening psychischer Störungen mit dem "Gesundheitsfragebogen für Patienten (PHQ-D)" [Screening of mental disorders using the 'Patient Health Questionnaire (PHQ-D)']. *Diagnostica, 50*(4), 171–181. https://doi.org/10.1026/0012-1924.50.4.171

Greenberg, R. P., Constantino, M. J., & Bruce, N. (2006). Are patient expectations still relevant for psychotherapy process and outcome? *Clinical Psychology Review, 26*(6), 657–678. https://doi.org/10.1016/j.cpr.2005.03.002

Hofmann, S. G., & Hay, A. C. (2018). Rethinking avoidance: Toward a balanced approach to avoidance in treating anxiety disorders. *Journal of Anxiety Disorders, 55*, 14–21. https://doi.org/10.1016/j.janxdis.2018.03.004

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In L. C. MacLean & W. T. Ziemba (Eds.), *Handbook of the fundamentals of financial decision making: Part I* (pp. 99–127). World Scientific. https://doi.org/10.1142/9789814417358_0006

Kirchner, L., Schummer, S. E., Krug, H., Kube, T., & Rief, W. (2022). How social rejection expectations and depressive symptoms bi-directionally predict each other–A cross-lagged panel analysis. *Psychology and Psychotherapy: Theory. Research and Practice, 95*(2), 477–492. https://doi.org/10.1111/papt.12383

Kirsch, I. (2018). Response expectancy and the placebo effect. *International Review of Neurobiology, 138*, 81–93. https://doi.org/10.1016/bs.irn.2018.01.003

Kirsch, I., Wampold, B., & Kelley, J. M. (2016). Controlling for the placebo effect in psychotherapy: Noble quest or tilting at windmills? *Psychology of Consciousness: Theory, Research, and Practice, 3*(2), 121–131. https://doi.org/10.1037/cns0000065

Kube, T., D'Astolfo, L., Glombiewski, J. A., Doering, B. K., & Rief, W. (2017). Focusing on situation-specific expectations in major depression as basis for behavioural experiments – Development of the Depressive Expectations Scale. *Psychology and Psychotherapy: Theory, Research and Practice, 90*(3), 336–352. https://doi.org/10.1111/papt.12114

Kube, T., Glombiewski, J. A., Gall, J., Touissant, L., Gärtner, T., & Rief, W. (2019). How to modify persisting negative expectations in major depression? An experimental study comparing three strategies to inhibit cognitive immunization against novel positive experiences. *Journal of Affective Disorders, 250*, 231–240. https://doi.org/10.1016/j.jad.2019.03.027

Kube, T., Glombiewski, J. A., & Rief, W. (2019). Erwartungsfokussierte psychotherapeutische Interventionen bei Personen mit depressiver Symptomatik [Expectation-focused psychotherapeutic interventions for people with depressive symptoms]. *Verhaltenstherapie, 29*(4), 281–291. https://doi.org/10.1159/000496944

Kube, T., Kirchner, L., Lemmer, G., & Glombiewski, J. A. (2022). How the discrepancy between prior expectations and new information influences expectation updating in depression—The greater, the better? *Clinical Psychological Science, 10*(3), 430–449. https://doi.org/10.1177/21677026211024644

Kube, T., Kirchner, L., Rief, W., Gärtner, T., & Glombiewski, J. A. (2019). Belief updating in depression is not related to increased sensitivity to unexpectedly negative information. *Behaviour Research and Therapy, 123*, Article 103509. https://doi.org/10.1016/j.brat.2019.103509

Kube, T., Rief, W., & Glombiewski, J. A. (2017). On the maintenance of expectations in major depression – Investigating a neglected phenomenon. *Frontiers in Psychology, 8*, Article 232630. https://doi.org/10.3389/fpsyg.2017.00009

Kube, T., Rief, W., Gollwitzer, M., Gartner, T., & Glombiewski, J. A. (2019). Why dysfunctional expectations in depression persist – Results from two experimental studies investigating cognitive immunization. *Psychological Medicine, 49*(9), 1532–1544. https://doi.org/10.1017/S0033291718002106

Kube, T., Schwarting, R., Rozenkrantz, L., Glombiewski, J. A., & Rief, W. (2020). Distorted cognitive processes in major depression: A predictive processing perspective. *Biological Psychiatry, 87*(5), 388–398. https://doi.org/10.1016/j.biopsych.2019.07.017

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). *lmerTest: Tests in linear mixed effects models* (R Package, Version 2.0-734) [Computer software]. Datacamp: New York, NY, USA.

Lam, D., & Cheng, L. (2001). Cognitive behavior therapy approach to assessing dysfunctional thoughts. *Counselling Psychology Quarterly, 14*(3), 255–265. https://doi.org/10.1080/09515070110092325

Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2018). *Emmeans: Estimated marginal means, aka least-squares means* (R package version, 1(1), 3) [Computer software].

Lightsey, R. (1999). Albert Bandura and the exercise of self-efficacy. *Journal of Cognitive Psychotherapy, 13*(2), 158–166. https://doi.org/10.1891/0889-8391.13.2.158

Liknaitzky, P., Smillie, L. D., & Allen, N. B. (2017). Out-of-the-blue: Depressive symptoms are associated with deficits in processing inferential expectancy-violations using a novel cognitive rigidity task. *Cognitive Therapy and Research, 41*(5), 757–776. https://doi.org/10.1007/s10608-017-9853-x

Liknaitzky, P., Smillie, L. D., & Allen, N. B. (2018). The low and narrow: A preliminary test of the association between depressive symptoms and deficits in producing divergent inferences. *Creativity Research Journal, 30*(1), 67–77. https://doi.org/10.1080/10400419.2018.1411459

Lintvedt, O. K., Griffiths, K. M., Sørensen, K., Østvik, A. R., Wang, C. E. A., Eisemann, M., & Waterloo, K. (2013). Evaluating the effectiveness and efficacy of unguided internet-based self-

help intervention for the prevention of depression: A randomized controlled trial. *Clinical Psychology & Psychotherapy, 20*(1), 10–27. https://doi.org/10.1002/cpp.770

Lüdecke, D. (2024). *sjPlot: Data Visualization for Statistics in Social Science* (R package) [Computer software]. https://cran.r-project.org/web/packages/sjPlot/index.html

Margraf, J., & Ehlers, A. (2007). *Beck Angst-Inventar (BAI) – Deutsche Ausgabe* [Beck Anxiety Inventory (BAI) – German Version]. Harcourt.

McMillan, D., & Lee, R. (2010). A systematic review of behavioral experiments vs. exposure alone in the treatment of anxiety disorders: A case of exposure while wearing the emperor's new clothes? *Clinical Psychology Review, 30*(5), 467–478. https://doi.org/10.1016/j.cpr.2010.01.003

Meiran, N., Diamond, G. M., Toder, D., & Nemets, B. (2011). Cognitive rigidity in unipolar depression and obsessive compulsive disorder: Examination of task switching, Stroop, working memory updating and post-conflict adaptation. *Psychiatry Research, 185*(1-2), 149–156. https://doi.org/10.1016/j.psychres.2010.04.044

Pinquart, M., Rothers, A., Gollwitzer, M., Khosrowtaj, Z., Pietzsch, M., & Panitz, C. (2021). Predictors of coping with expectation violation: An integrative review. *Review of General Psychology, 25*(3), 321–333. https://doi.org/10.1177/10892680211024123

Posit PBC. (2009-2020). *RStudio* (Version 1.2.5042) [Computer software]. https://www.rstudio.com

Puspitarini, Y. D., Akhyar, M., & Djono. (2019). Development of video media based on Powtoon in social sciences. *International Journal of Educational Research Review, 4*(2), 198–205. https://doi.org/10.24331/ijere.518054

Richmond, A. S., Broussard, K. A., Sterns, J. L., Sanders, K. K., & Shardy, J. C. (2015). Who are we studying? Sample diversity in teaching of psychology research. *Teaching of Psychology, 42*(3), 218–226. https://doi.org/10.1177/0098628315587619

Rief, W., & Glombiewski, J. A. (2016). Erwartungsfokussierte Psychotherapeutische Interventionen (EFPI) [Expectation-Focused Psychological Interventions (EFPI)]. *Verhaltenstherapie, 26*(1), 47–54. https://doi.org/10.1159/000442374

Rief, W., & Glombiewski, J. A. (2017). The role of expectations in mental disorders and their treatment. *World Psychiatry, 16*(2), 210–211. https://doi.org/10.1002/wps.20427

Rief, W., Glombiewski, J. A., Gollwitzer, M., Schubö, A., Schwarting, R., & Thorwart, A. (2015). Expectancies as core features of mental disorders. *Current Opinion in Psychiatry, 28*(5), 378–385. https://doi.org/10.1097/YCO.0000000000000184

Rief, W., & Joormann, J. (2019). Revisiting the cognitive model of depression: The role of expectations. *Clinical Psychology in Europe, 1*(1), Article e32605. https://doi.org/10.32872/cpe.v1i1.32605

Sander, L., Rausch, L., & Baumeister, H. (2016). Effectiveness of internet-based interventions for the prevention of mental disorders: A systematic review and meta-analysis. *JMIR Mental Health, 3*(3), Article e41. https://mental.jmir.org/2016/3/e41/

Servatius, R. J. (2016). Editorial: Avoidance: From basic science to psychopathology. *Frontiers in Behavioral Neuroscience, 10*, Article 15. https://doi.org/10.3389/fnbeh.2016.00015

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3), 1–67. https://doi.org/10.18637/jss.v045.i03

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270–277. https://doi.org/10.1002/wps.20238

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge.

Wampold, B. E., Minami, T., Tierney, S. C., Baskin, T. W., & Bhati, K. S. (2005). The placebo is powerful: Estimating placebo effects in medicine and psychotherapy from randomized clinical trials. *Journal of Clinical Psychology, 61*(7), 835–854. https://doi.org/10.1002/jclp.20129

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine, 30*(4), 377–399. https://doi.org/10.1002/sim.4067

Wilhelm, M., Bleichhardt, G., Ewen, A.-C., & Rief, W. (2022). Erwartungsfokussierte psychotherapeutische Interventionen bei Depression: Ein Behandlungsmanual [Expectation-Focused Psychotherapeutic Interventions for Depression: A treatment manual]. *Psychotherapeut, 67*(3), 204–212. https://doi.org/10.1007/s00278-022-00590-1

Wright, J. H., & Davis, D. (1994). The therapeutic relationship in cognitive-behavioral therapy: Patient perceptions and therapist responses. *Cognitive and Behavioral Practice, 1*(1), 25–45. https://doi.org/10.1016/S1077-7229(05)80085-9

Zhou, T., Li, X., Gao, J., & Kong, J. (2016). Internet-based cognitive behavioral therapy for subthreshold depression: A systematic review and meta-analysis. *BMC Psychiatry, 16*, Article 356. https://doi.org/10.1186/s12888-016-1061-9

EACLIPT

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

zpid
Leibniz-Institut für Psychologie

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