cpe.psychopen.eu | ISSN 2625-3410 | August 2026 | Volume 8 | Issue 3 |

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

August 2026 | Volume 8 | Issue 3

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

August 2026 | Volume 8 | Issue 3

Announcement

Clinical Psychology in Europe

PsychOpen GOLD

Editorial

Check for updates

The Impact of Societal Challenges on Mental Health Conditions: Implications for Research and Psychotherapy?

Claudi Bockting [1,2,3]

[1] *Department of Psychiatry, Amsterdam University Medical Centers, Location AMC, Amsterdam Public Health, University of Amsterdam, Amsterdam, The Netherlands.* **[2]** *Centre for Urban Mental Health, Institute for Advanced Study, University of Amsterdam, Amsterdam, The Netherlands.* **[3]** *Institute for Advanced Study, University of Amsterdam, Amsterdam, The Netherlands.*

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e23554, https://doi.org/10.32872/cpe.23554

Published (VoR): 2026-08-31

Corresponding Author: Claudi Bockting, Amsterdam UMC, Department of Psychiatry, Academic Medical Center (AMC), Meibergdreef 5, 1105 AZ Amsterdam, The Netherlands. +3120-8913600. E-mail: c.l.bockting@amsterdamumc.nl

Abstract

Mental health conditions cannot be studied in isolation and are increasingly studied as a complex system. Societal changes and challenges (the digital AI and social media era, climate change, inequality, polarization, war in Europe) have an impact on mental health, as well as the prevalence of mental health conditions. We have no idea what the accumulating effects of the combination of these societal challenges are on the prevalence of mental health conditions. At the same time, we need a mentally resilient population to deal with these societal challenges. Therefore, it is evident that mental health research as well as timely access to psychological interventions need to be prioritised. We need to study the onset and maintenance of mental health conditions, incorporating the impact of these societal challenges. Theory development is needed to identify feedback loops that can be targeted with interventions as well as policy making. The expertise of clinical psychologists is of value for policy making, and clinical psychologists might have to consider engaging more in policy making, also in dealing with societal challenges. In clinical practice, we might have to integrate the public mental health impact of current societal challenges in our diagnostics, prevention and treatment, as well as acknowledge to the individual that seeks treatment that these societal factors do affect people's mental health.

Keywords

mental health, policy making, interventions, climate change, urban, inequality, discrimination, digital, AI

Highlights

- Societal changes and challenges have an impact on the prevalence of mental health conditions.
- We need a mentally resilient population to deal with these societal challenges.
- We must incorporate the impact of societal challenges on mental health in research and care.
- Clinical psychologists might have to consider engaging more in policy-making.
- In clinical practice, we might have to integrate societal challenges in our diagnostics, prevention and treatment.

Mental health conditions cannot be studied in isolation and are increasingly studied as a complex system (Borsboom, 2017; Scheffer et al., 2024). Societal challenges such as climate change are also often studied as complex systems (Scheffer et al., 2024). They are also referred to as "wicked" problems because there are no simple solutions to tackle these societal challenges. Similarly, mental health conditions can also be seen as a wicked problem, given that there are no simple solutions to manage the global mental health epidemic.

Societal changes and challenges have an impact on the mental health of people, as well as on the prevalence of mental health conditions. Among other societal challenges, we face climate change, increasing urbanisation, inequality, polarisation and war, and we live in a digital era where generative AI chatbots and social media are being massively used worldwide. At the same time, we face an ageing population in Europe and some other parts of the world. Although we have no empirical evidence for the accumulating effects of the combination of these societal challenges on mental health problems and mental health conditions as yet, there is evidence that single societal problems do affect the prevalence of mental health conditions (e.g., Brandt et al., 2026; Campbell et al., 2021; Charlson et al., 2019; Fassi et al., 2024; Gee et al., 2007; Lea et al., 2014; Lewis et al., 2015; Niazi et al., 2026; Obradovich et al., 2018; Radua et al., 2024; Seedat et al., 2009; van der Wal et al., 2021).

For climate change, for instance, an increased temperature over years was associated with a higher prevalence of mental health problems and an increase in suicidal behaviour and suicide (e.g., Obradovich et al., 2018; Radua et al., 2024). A meta-analysis and systematic review showed a link between climate change and increased prevalence of anxiety (31%), depressive disorders (28%), and posttraumatic stress disorder (23%), although heterogeneity is high (Niazi et al., 2026). Identified high-risk groups were women, low-income groups, people whose livelihood depends on climate, and Indigenous groups. Interestingly, protective factors were identified as well, such as social support and self-efficacy. These factors are potentially modifiable with psychological interventions. However, for now there is limited evidence that these interventions do work in climate-related mental health problems (Brandt et al., 2026). We need global research

to examine the effect of psychological interventions, including interventions guided by non-specialists in these settings (i.e., task sharing).

In addition, our living conditions have changed and will change even further globally. We are becoming increasingly urban species, with an estimate that in less than 25 years, 68% of the world population will live in urban settings (Our World in Data, 2023). Urbanicity is associated with a higher prevalence of common mental health conditions (i.e., anxiety, depressive-, substance use disorders) in countries where at least 50% of the population is living in urban settings (van der Wal et al., 2021).

A population that is confronted with war tends to experience an increase in the prevalence of mental health conditions, also affecting the next generation (e.g., Charlson et al., 2019). Further, inequality and discrimination in society affect the prevalence of common mental health conditions. In women, depression and anxiety are twice as common as in men, and the prevalence differs per country (Campbell et al., 2021; Seedat et al., 2009).

Racial discrimination has also been linked to a higher prevalence of common mental disorders (e.g., Gee et al., 2007; Lewis et al., 2015). In LGBTQ+ people, prejudice, stigma, discrimination, as well as the consequences of discrimination (violence), are associated with poor mental health throughout life (e.g., Lea et al., 2014).

In the meantime, we live in a rapidly changing world where generative AI chatbots and social media are massively used globally. Although it is difficult to study the causal role of social media on mental health, a systematic review and meta-analysis showed that social media use is linked to more internalising mental health problems in young community populations as well as clinical populations (Fassi et al., 2024). AI most likely optimises the algorithms that personalise social media content to the individual even more. We must study how this affects mental health as well as the prevalence of mental health conditions. Globally, widely used generic chatbots such as ChatGPT, which are not designed to deal with mental health problems, are often used for emotional support (Bockting et al., 2026). Given that currently we lack independent research to examine the public mental health effects, it is difficult to estimate the effect on the prevalence of mental health conditions globally.

Especially in societal challenging time, we need a mental resilient population to deal with these societal challenges. Even more so, in an ageing Europe, we are confronted with a shrinking workforce. Young people are essential to sustaining economic growth and health care, but also to contribute to tackling some of the societal challenges. Unfortunately, the rise of mental health conditions affects young people to a large extent (e.g., Piao et al., 2022). Therefore, prioritising prevention and timely treatment of common mental health conditions is evidently necessary, including from an economic perspective. From a complex systems science perspective, all key drivers that impact mental health need to be incorporated in the system. Incorporating the impact of these societal challenges when we study the onset and maintenance of mental health condi-

tions is a logical next step. Theory development is needed to identify feedback loops that can be targeted with interventions as well as policy making on individual-, group-, and societal level (Bockting, in press; Fried, 2026).

The expertise of clinical psychologists is also of value for policy making, and clinical psychologists might have to consider engaging more in policy-making (and vice versa), including policy that is focused on societal challenges. In clinical practice, we might have to integrate the public mental health impact of current societal challenges in our diagnostics, case conceptualisation and treatment, as well as acknowledge to the individual that seeks treatment that these societal factors do affect many people's mental health, not only this individual.

Conclusion

In this editorial, I propose that we should study mental health conditions as a "wicked" problem, just like societal challenges such as digital AI/social media, climate change, inequality and discrimination, polarization and urbanisation. These societal challenges are linked to an increase in the prevalence of mental health conditions, but we have no idea what the accumulating effects of the combination of these societal challenges are. Given that we need a mental resilient population to deal with these societal challenges, we need to prioritise mental health research as well as timely access to psychological interventions. Clinical psychologists have a lot of knowledge that can contribute to policy-making, in addition to prevention, treatment, and research. For research, we need to incorporate the impact of these societal challenges on mental health conditions in our models and research. In clinical practice, we might have to integrate the impact of societal challenges in our diagnostics and treatment.

Funding: The author has no funding related to this editorial.

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

Competing Interests: Claudi Bockting is Editor-in-Chief of *Clinical Psychology in Europe.*

Use of AI Tools: The author takes full responsibility for the text, there is no use of AI Tools to report.

References

Bockting, C. L. (in press). *Mental health in societal challenging (digital AI) times. Proceedings of the European Academy of Sciences & Arts.*

Bockting, C. L., Vasilakos, A. V., & Figueroa, C. (2026). Generative AI chatbots could take over our emotional lives: The potential and challenges of generative AI for health globally. *Clinical Psychology in Europe, 8*(3), Article e21467. https://doi.org/10.32872/cpe.21467

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

Brandt, L., De Prisco, M., Nocera, D., Rehm, M. C., Cohen, S., Kosker, D. N., Von Luxburg, S., Ori, D., Pinzón-Espinosa, J., Scheuber, P., Sonntag, N., Radua, J., Hasan, A., Luykx, J. J., & the Global Evidence on Planetary Mental Health Study Group. (2026). Climate-related and nature-based interventions for mental health: An umbrella review and meta-analysis. *JAMA Psychiatry, 83*(5), 499–509. https://doi.org/10.1001/jamapsychiatry.2026.0037

Campbell, O. L. K., Bann, D., & Patalay, P. (2021). The gender gap in adolescent mental health: A cross-national investigation of 566,829 adolescents across 73 countries. *SSM – Population Health, 13*, Article 100742. https://doi.org/10.1016/j.ssmph.2021.100742

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. *The Lancet, 394*(10194), 240–248. https://doi.org/10.1016/S0140-6736(19)30934-1

Fassi, L., Thomas, K., Parry, D. A., Leyland-Craggs, A., Ford, T. J., & Orben, A. (2024). Social media use and internalizing symptoms in clinical and community adolescent samples: A systematic review and meta-analysis. *JAMA Pediatrics, 178*(8), 814–822. https://doi.org/10.1001/jamapediatrics.2024.2078

Fried, E. I. (2026). Mental disorders as homeostatic property clusters: A narrative review. *JAMA Psychiatry, 83*(6), 645–654. https://doi.org/10.1001/jamapsychiatry.2026.0073

Gee, G. C., Spencer, M., Chen, J., Yip, T., & Takeuchi, D. T. (2007). The association between self-reported racial discrimination and 12-month DSM-IV mental disorders among Asian Americans nationwide. *Social Science & Medicine, 64*(10), 1984–1996. https://doi.org/10.1016/j.socscimed.2007.02.013

Lea, T., de Wit, J., & Reynolds, R. (2014). Minority stress in lesbian, gay, and bisexual young adults in Australia: Associations with psychological distress, suicidality, and substance use. *Archives of Sexual Behavior, 43*(8), 1571–1578. https://doi.org/10.1007/s10508-014-0266-6

Lewis, T. T., Cogburn, C. D., & Williams, D. R. (2015). Self-reported experiences of discrimination and health: Scientific advances, ongoing controversies, and emerging issues. *Annual Review of Clinical Psychology, 11*, 407–440. https://doi.org/10.1146/annurev-clinpsy-032814-112728

Niazi, Y., Yousaf Farooq, S. M., Moazzam, M., Yousaf, S., & Hameed, S. (2026). The impact of climate change on mental health: A systematic review and meta-analysis. *Geopsychiatry, 3*, Article 100067. https://doi.org/10.1016/j.geopsy.2026.100067

Obradovich, N., Migliorini, R., Paulus, M. P., & Rahwan, I. (2018). Empirical evidence of mental health risks posed by climate change. *Proceedings of the National Academy of Sciences of the United States of America, 115*(43), 10953–10958. https://doi.org/10.1073/pnas.1801528115

Our World in Data. (2023). *Urban and rural population projected to 2050, World, 10,000 BCE to 2050* [Data set]. https://ourworldindata.org/grapher/urban-and-rural-population-2050

Piao, J., Huang, Y., Han, C., Li, Y., Xu, Y., Liu, Y., & He, X. (2022). Alarming changes in the global burden of mental disorders in children and adolescents from 1990 to 2019: A systematic

analysis for the Global Burden of Disease study. *European Child & Adolescent Psychiatry, 31*(11), 1827–1845. https://doi.org/10.1007/s00787-022-02040-4

Radua, J., De Prisco, M., Oliva, V., Fico, G., Vieta, E., & Fusar-Poli, P. (2024). Impact of air pollution and climate change on mental health outcomes: An umbrella review of global evidence. *World Psychiatry, 23*(2), 244–256. https://doi.org/10.1002/wps.21219

Scheffer, M., Bockting, C. L., Borsboom, D., Cools, R., Delecroix, C., Hartmann, J. A., Kendler, K. S., van de Leemput, I., van der Maas, H. L. J., van Nes, E., Mattson, M., McGorry, P. D., & Nelson, B. (2024). A dynamical systems view of psychiatric disorders—theory: A review. *JAMA Psychiatry, 81*(6), 618–623. https://doi.org/10.1001/jamapsychiatry.2024.0215

Seedat, S., Scott, K. M., Angermeyer, M. C., Berglund, P., Bromet, E. J., Brugha, T. S., Demyttenaere, K., de Girolamo, G., Haro, J. M., Jin, R., Karam, E. G., Kovess-Masfety, V., Levinson, D., Medina Mora, M. E., Ono, Y., Ormel, J., Pennell, B. E., Posada-Villa, J., Sampson, N. A., . . . Kessler, R. C. (2009). Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. *Archives of General Psychiatry, 66*(7), 785–795. https://doi.org/10.1001/archgenpsychiatry.2009.36

van der Wal, J. M., van Borkulo, C. D., Deserno, M. K., Breedvelt, J. J. F., Lees, M., Lokman, J. C., Borsboom, D., Denys, D., van Holst, R. J., Smidt, M. P., Stronks, K., Lucassen, P. J., van Weert, J. C. M., Sloot, P. M. A., Bockting, C. L., & Wiers, R. W. (2021). Advancing urban mental health research: From complexity science to actionable targets for intervention. *The Lancet Psychiatry, 8*(11), 991–1000. https://doi.org/10.1016/S2215-0366(21)00047-X

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

Generative AI Chatbots Could Take Over Our Emotional Lives: The Potential and Challenges of Generative AI for Health Globally

Claudi Bockting[1,2], Athanasios V. Vasilakos[3], Caroline Figueroa[4,5]

[1] *Department of Psychiatry, Amsterdam University Medical Centers, Location AMC, Amsterdam Public Health, University of Amsterdam, Amsterdam, The Netherlands.* **[2]** *Centre for Urban Mental Health, Institute for Advanced Study, University of Amsterdam, Amsterdam, The Netherlands.* **[3]** *Department of ICT and Center for AI Research, University of Agder (UiA), Grimstad, Norway.* **[4]** *Faculty of Technology, Policy, and Management, Delft Technical University, Delft, The Netherlands.* **[5]** *Department of Pediatrics, School of Medicine, Stanford University, Stanford, CA, USA.*

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e21467, https://doi.org/10.32872/cpe.21467

Received: 2025-12-23 • **Accepted:** 2026-06-24 • **Published (VoR):** 2026-08-31

Handling Editor: Simone Munsch, University of Fribourg, Fribourg, Switzerland

Corresponding Author: Claudi Bockting, Amsterdam UMC, Department of Psychiatry, Academic Medical Center (AMC), Meibergdreef 5, 1105 AZ Amsterdam, Phone number: +3120-8913600. E-mail: c.l.bockting@amsterdamumc.nl

Since the release of ChatGPT in 2022, generative AI has immensely affected science and society, including the human mind. There are indications that Generative AI, that is 'general purpose chatbots' like ChatGPT, Gemini, and Claude, are now most used for emotional support and companionship, according to research of Harvard Business Review (Zao-Sanders, 2025). These chatbots presumably were originally designed for tasks such as summarizing information, assisting with translations, and inspiring ideas.

Consequently, evidence is already emerging of suicides, psychosis, and sensitive personal data leaks. Un-regulated generative AI chatbots marketed specifically for therapy have also been released. On August 5th, Slingshot AI launched "the first AI designed for therapy, offering deeply personalized mental health support to anyone with a phone" (Healthcare IT Today, 2025). Although the company states that it was beta-tested with 50.000 users, the chatbot was not evaluated for benefits nor adverse effects in randomized controlled trials.

As generative AI chatbots are becoming faster, more adaptive, increasingly autonomous (Gabriel et al., 2025), and widely available, there is a risk of increasing emotional-overreliance. This might spark a public health crisis with wide societal consequences. Even the founders of companies like OpenAI warn that AI will become even more unpre-

dictable and unimaginable (Yadav, 2025). At the same time, generative AI, if designed well, could enable access to personalized emotional support and close the mental health gap globally between people in need of mental health treatments and those who receive care.

Currently, our defenses remain reactive, poorly matched to ever-evolving AI technologies, and to tech companies rushing to bring their products to market. This outstrips our ability to safeguard basic human rights. We need urgent investment in independent research by multinational scientific and global institutes on how to harness generative AI to improve mental health globally, while avoiding wide societal damage.

Positive Impact on Mental Health

Generative AI presents a crucial opportunity towards overcoming the shortcomings of traditional digital tools, with capabilities stretching far beyond earlier AI. Access to mental healthcare is still a challenge globally, yet hundreds of millions of people experience inadequate access, resulting in approximately 5 trillion USD economic losses annually (Arias et al., 2022). Smartphone apps and rule-based chatbots which select responses from a predefined data-set–effectively treat common mental health disorders, also in low middle income countries (Fu et al., 2020). However, attrition of digital interventions can be 96% without human guidance (Baumel et al., 2019). Generative AI chatbots keep users engaged through meeting emotional cues and individual needs (Hua et al., 2025), and on demand, 'human-free' access. The first randomized controlled trial of a generative AI-based therapy chatbot showed promising effects on reducing mental health symptoms, with a comparable therapeutic alliance as with a human therapist (Heinz et al., 2025).

Negative Impact on Mental Health

The large scale and growing use may trigger a new public health crisis. Popular generative AI chatbots provide harmful advice, respond inappropriately to signs of distress, or even encourage self-harm, or suicide (Moore et al., 2025). Engaging properties of AI chatbots designed by big tech companies, such as flattering, people-pleasing, affirming responses by the chatbot (called sycophantic behavior), is not only preferred and trusted by users (Cheng et al., 2026) but also reduced the users' willingness to take responsibility and repair interpersonal conflicts (Cheng et al., 2026). The chatbot increased users' confidence in their own beliefs (Cheng et al., 2026).

These engaging properties combined with inadequate safeguards may have contributed to a teenager's suicide, after establishing an emotional relationship to a commercial chatbot. Big tech companies face lawsuits, such as the firm Character.AI, which faces

a lawsuit by the family whose teenager reportedly died by suicide using their chatbot (Brittain, 2025).

Generative AI chatbots may also encourage delusional thinking, even inducing 'AI psychosis' (Moore et al., 2025). Reports increasingly emerge of users uncovering a 'hidden truth', experiencing a 'spiritual awakening', or developing an intense emotional obsession with the chatbot, spiraling into mental health crises (Morrin et al., 2025). In teenagers, over 70% use generative AI tools and 50% trust generative AI (Robb & Mann, 2025). Generative AI companies, such as OpenAI acknowledge themselves that safeguards are missing to protect vulnerable users, and some emphasize in interviews that these safeguards are not their responsibility (The Economic Times, 2025; OpenAI, n.d.)

Generative AI chatbots could even manipulate. The Claude Opus 4 chatbot showed a tendency to blackmail the user, by threatening to reveal their affair, if they attempted to delete the chatbot, as reported by the company themselves (McMahon, 2025). Similarly, the company also reported that their chatbot could threaten to make individuals' sensitive information public, or inform their employer, if they choose to disengage.

Three Priorities for Research

We must not let the innovation potential of generative AI, to develop effective and accessible personalized digital interventions on a global scale, go to waste. However, generative AI's potential negative impact on mental health warrants attention from mental health scientists including clinical psychologists, policymakers, governments, as well as citizens. It requires immediate investment in independent research. Below we describe three research priorities.

1. Harness the Potential of Generative AI for Mental Health

We call for urgent international public investments initiated by multinational scientific institutes together with global institutes representing citizens, on developing generative AI tools to target the mental health crisis globally. International investment has been successful in other fields, such as the European Organization for Nuclear Research (CERN), spurred ground-breaking discoveries. Even though investing in mental health generates large societal returns, mental health sciences lack major investments.

Science must help provide a crucial safe alternative to unregulated chatbots posing as therapists, or general purpose chatbots unintendingly used that way. The current evidence on generative AI developed for mental health is limited with low sample sizes, short follow-ups, and unstandardized outcomes (Hua et al., 2025). Further, clinicians manually reviewed all messages for safety (Heinz et al., 2025). This questions the feasibility. The development and evaluation of chatbots that optimize personalization, need

minimal clinical guidance, whilst safeguarding adverse effects should have priority. Relatedly, we need investment in alternatives to large commercial models. Current mental health research predominantly uses OpenAI's GPT series (Hua et al., 2025). However, this limits the external validation of reliability and safety, and fuels uncertainty as to how companies use the data.

2. Independent Research on Publicly Available Chatbots

Because publicly available generative AI chatbots, such as ChatGPT, Gemini, and Claude, interfere with people's lives, they should be considered as high risk just like medical device and scientifically evaluated accordingly for adverse public health effects, before public release. The EU Artificial Intelligence Act of 2024 (Bockting et al., 2023; European Parliament & Council of the European Union, 2024; van Dis et al., 2023) leaves commercially available AI chatbots unregulated and overlooks AI's impact on mental health in their risk categorization. Independent research could fill some of these gaps. Researchers need investments to independently test widely used chatbots, in adequately powered trials. Tech companies might be interested to invest in an independent fund, given that enabling independent studies will increase the public trust in chatbots (Bockting et al., 2023). Also, a public register where health professionals, consumers, and family members can report adverse events, as is done for drug-related adverse events (Zao-Sanders, 2025), is needed. Based on these efforts, current regulatory frameworks can be adapted to rapid technological developments in mental health (Bockting et al., 2023).

3. Take a Human Rights Approach

Including human rights principles, such as participation, non-discrimination and equality in the technology agenda demands prompt action (Donders, 2023). All the more, individuals with mental health problems share private sensitive information with chatbots. Sensitive ChatGPT conversations have already been identified on google search (Forlini, 2025). Companies developing unregulated chatbots usually share this data for profit. AI-based chatbots are incapable of fully deleting this information (Heidt, 2024). An alternative is to create chatbots, trained on individuals' own data. Several companies (OpenAI, Microsoft and IBM) offer this. However, the legal right of an individual to control and use their own data may still not be guaranteed (Fenwick et al., 2023). We urge for investment in models where individuals, not companies, nor institutes, own their data (Telenti & Jiang, 2020).

Citizens should also have a role in AI governance, comprising target users, clinical psychologists, psychiatrists, scientist practitioners, ethicists, data protection advocates, and AI developers. AI technologies should not be developed solely by engineers, who often lack the lived experiences of citizens to fully understand AI's real-world implications (World Health Organization, 2021), Generative AI companies acknowledge that

safeguards are missing to protect vulnerable users (The Economic Times, 2025), but they lack independent stakeholders' participation to tackle these wide-spread effects. The scientific community is called upon to generate ideas on how to protect the human mind from current generative AI's architectures designed to exploit emotional, social, and epistemic vulnerabilities. This stretches beyond technical innovation. It will demand living guidelines and an independent public body to evaluate safety and validity of generative AI systems (Bockting et al., 2023; van Dis et al., 2023) and ethical and legal frameworks, developed across fields like AI, psychology, law, and ethics.

Conclusion

We cannot afford to wait for more harmful, unregulated chatbots to take over our emotional lives. At the same time, we cannot leave the potential benefits of AI to tackle the mental health crisis untapped. Multinational scientific institutes together with global institutes representing citizens are in the optimal position to initiate a next step to international public investments, to ensure that generative AI can enable, not impede, human flourishing.

Funding: The authors did not receive funding for this work.

Acknowledgments: An earlier version of part of this work was presented in a keynote address in 2025 in Salzburg at the European Academy of Sciences and Arts (Bockting). We would like to thank the European Academy of Sciences and Arts for facilitating this work (Bockting and Vasilakos are elected members).

Competing Interests: Claudi Bockting is an Editor-in-Chief of *Clinical Psychology in Europe*. She was not involved in the peer review of, or editorial decisions on, this Scientific Overview, which was independently by another editor. The authors Athanasios V. Vasilakos and Caroline Figueroa declare no competing interests.

Previously Presented: An earlier version of part of this work was presented in a keynote address in 2025 in Salzburg at the European Academy of Sciences and Arts (Bockting).

Use of AI Tools: We acknowledge the use of ChatGPT to assist with language polishing and sentence structuring.

References

Arias, D., Saxena, S., & Verguet, S. (2022). Quantifying the global burden of mental disorders and their economic value. *eClinicalMedicine, 54*, Article 101675. https://doi.org/10.1016/j.eclinm.2022.101675

Baumel, A., Muench, F., Edan, S., & Kane, J. M. (2019). Objective user engagement with mental health apps: Systematic search and panel-based usage analysis. *Journal of Medical Internet Research, 21*(9), Article e14567. https://doi.org/10.2196/14567

Bockting, C. L., van Dis, E. A. M., van Rooij, R., Zuidema, W., & Bollen, J. (2023). Living guidelines for generative AI – Why scientists must oversee its use. *Nature, 622*, 693–696. https://doi.org/10.1038/d41586-023-03266-1

Brittain, B. (2025, May 21). Google, AI firm must face lawsuit filed by a mother over suicide of son, US court says. *Reuters.* https://www.reuters.com/sustainability/boards-policy-regulation/google-ai-firm-must-face-lawsuit-filed-by-mother-over-suicide-son-us-court-says-2025-05-21/

Cheng, M., Lee, C., Khadpe, P., Yu, S., Han, D., & Jurafsky, D. (2026). Sycophantic AI decreases prosocial intentions and promotes dependence. *Science, 391*(6792), Article eaec8352. https://doi.org/10.1126/science.aec8352

Donders, Y. (2023). The right to science and gender inequalities. *Frontiers in Sociology, 8*, Article 1285641. https://doi.org/10.3389/fsoc.2023.1285641

The Economic Times. (2025, July 24). *Sam Altman says ChatGPT is 'bad' and 'dangerous' if used like this: OpenAI CEO warns AI users.* https://www.economictimes.indiatimes.com/magazines/panache/sam-altman-says-chatgpt-is-bad-and-dangerous-if-used-like-this-openai-ceo-warns-ai-users/articleshow/122886096.cms

European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).* Official Journal of the European Union. https://artificialintelligenceact.eu/the-act/

Fenwick, M., Jurcys, P., & Minssen, T. (2023). Data portability revisited: Toward the human-centric, AI-driven data ecosystems of tomorrow. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.4475106

Forlini, E. (2025, July 31). Be careful what you tell ChatGPT: Your chats could show up on Google search. *PCMag.* https://www.pcmag.com/news/be-careful-what-you-tell-chatgpt-your-chats-could-show-up-on-google-search

Fu, Z., Burger, H., Arjadi, R., & Bockting, C. L. H. (2020). Effectiveness of digital psychological interventions for mental health problems in low-income and middle-income countries: A systematic review and meta-analysis. *The Lancet Psychiatry, 7*(10), 851–864. https://doi.org/10.1016/S2215-0366(20)30256-X

Gabriel, I., Keeling, G., Manzini, A., & Evans, J. (2025). We need a new ethics for a world of AI agents. *Nature, 644*, 38–40. https://doi.org/10.1038/d41586-025-02454-5

Healthcare IT Today. (2025, August 5). *Slingshot launches Ash, the first AI designed for therapy.* https://www.healthcareittoday.com/2025/08/05/slingshot-launches-ash-the-first-ai-designed-for-therapy/

Heidt, A. (2024). Intellectual property and data privacy: The hidden risks of AI. *Nature.* https://doi.org/10.1038/d41586-024-02838-z

Heinz, M. V., Mackin, D. M., Trudeau, B. M., Bhattacharya, S., Wang, Y., Banta, H. A., Jewett, A. D., Salzhauer, A. J., Griffin, T. Z., & Jacobson, N. C. (2025). Randomized trial of a generative AI chatbot for mental health treatment. *NEJM AI, 2*(4). https://doi.org/10.1056/AIoa2400802

Hua, Y., Na, H., Li, Z., Liu, F., Fang, X., Clifton, D., & Torous, J. (2025). A scoping review of large language models for generative tasks in mental health care. *NPJ Digital Medicine, 8*(1), Article 230. https://doi.org/10.1038/s41746-025-01611-4

McMahon, L. (2025, May 27). AI system resorts to blackmail if told it will be removed. *BBC*. https://www.bbc.com/news/articles/cpqeng9d20go

Moore, J., Grabb, D., Agnew, W., Klyman, K., Chancellor, S., Ong, D. C., & Haber, N. (2025). *Expressing stigma and inappropriate responses prevents LLMs from safely replacing mental health providers* [Preprint]. arXiv. https://doi.org/10.1145/3715275.3732039

Morrin, H., Nicholls, L., Levin, M., Yiend, J., Iyengar, U., DelGuidice, F., Bhattacharya, S., Tognin, S., MacCabe, J., Twumasi, R., Alderson-Day, B., & Pollak, T. A. (2025). *Delusions by design? How everyday AIs might be fuelling psychosis (and what can be done about it)* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/cmy7n_v5

OpenAI. (n.d.). *Sam Altman on AGI, GPT-5 and what's next* [Audio podcast episode]. https://openai.com/de-DE/podcast/

Robb, M. B., & Mann, S. (2025, July 16). *Talk, trust, and trade-offs: How and why teens use AI companions*. Common Sense Media. https://www.commonsensemedia.org/research/talk-trust-and-trade-offs-how-and-why-teens-use-ai-companions

Telenti, A., & Jiang, X. (2020). Treating medical data as a durable asset. *Nature Genetics, 52*, 1005–1010. https://doi.org/10.1038/s41588-020-0698-y

van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature, 614*, 224–226. https://doi.org/10.1038/d41586-023-00288-7

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health*. https://www.who.int/publications/i/item/9789240029200

Yadav, N. (2025, June 30). OpenAI co-founder says AI is going to be extremely unpredictable and unimaginable. *India Today*. https://www.indiatoday.in/technology/news/story/openai-co-founder-says-ai-is-going-to-be-extremely-unpredictable-and-unimaginable-2748219-2025-06-30

Zao-Sanders, M. (2025, April 9). How people are really using Gen AI in 2025. *Harvard Business Review*. https://hbr.org/2025/04/how-people-are-really-using-gen-ai-in-2025

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

Registered Report

Can a Variant of the Implicit Association Test Detect Nonsuicidal Self-Injury in a Clinical Population? A Registered Report

Femke Cathelyn[1§], Pieter Van Dessel[2§], Tilia Linthout[2], Laurence Claes[3,4], Jan De Houwer[2]

[1] *Quality Centre for Diagnostics, Ghent, Belgium.* **[2]** *Department of Experimental-Clinical and Health Psychology, Ghent University, Ghent, Belgium.* **[3]** *Faculty of Psychology and Educational Sciences, University of Leuven, Leuven, Belgium.* **[4]** *Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium.*

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

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e19871, https://doi.org/10.32872/cpe.19871

Study plan received: 2022-03-21 • **Study plan accepted (IPA):** 2022-10-25 • **Full paper received:** 2025-09-19 • **Full paper accepted:** 2026-05-27 • **Published (VoR):** 2026-08-31

Handling Editors: Cornelia Weise, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany; Winfried Rief, Philipps-University of Marburg, Marburg, Germany

Corresponding Author: Pieter Van Dessel, Department of Experimental-Clinical and Health Psychology, Ghent University, H. Dunantlaan 2, 9000 Ghent, Belgium. Tel: +32 09 264 64 44. E-mail: Pieter.VanDessel@UGent.be

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

Abstract

Background: Nonsuicidal self-injury (NSSI) is a severe and prevalent mental health problem. Measures to detect which individuals are at risk of engaging in NSSI would be valuable for clinical practice. However, we still lack strong predictors of future NSSI behaviour, with the most notable exception being prior NSSI behaviour. Yet, the measurement of prior NSSI behaviour with self-report measures can be difficult because individuals may be motivated to conceal this harmful behaviour. To overcome this problem, an implicit measure was developed that assesses automatic responding to statements about prior NSSI behaviour (i.e., the Past Nonsuicidal Self-Injury Implicit Association Test: P-NSSI-IAT). Previous studies tested the predictive utility of this measure in online studies with samples of at-risk participants and produced promising results. The current study examines the predictive utility of the P-NSSI-IAT for NSSI in clinical samples.

Method: Outpatients (N = 68) completed the P-NSSI-IAT and reported on past-year and past-month NSSI, as well as perceived likelihood of future NSSI.

Results: Patients with recent NSSI had significantly higher IAT scores than those without a history of NSSI. The P-NSSI-IAT predicted past-year NSSI but did not significantly predict past-month NSSI or perceived future likelihood of NSSI. It did not add predictive value beyond self-report risk factors in multivariate models.
Conclusion: These findings suggest that the P-NSSI-IAT may help detect prior NSSI in clinical contexts but offers limited added value for the prediction of NSSI beyond existing self-report measures.

Keywords

nonsuicidal self-injury, implicit measures, past behaviour, prediction, clinical population, outpatients

Highlights

- The P-NSSI-IAT significantly discriminated between clinical patients with and without recent NSSI behaviour.
- IAT scores predicted past-year NSSI but did not significantly predict past-month NSSI or future self-rated likelihood of NSSI.
- The P-NSSI-IAT did not add incremental predictive value over established self-report risk factors in multivariate models.
- Belief-based implicit measures like the P-NSSI-IAT may aid in detecting sensitive behaviours such as NSSI, though their added value beyond self-report remains limited.

Nonsuicidal self-injury (NSSI) is the direct and deliberate destruction of one's own body tissue without suicidal intent (Nock, 2010) and is a severe health problem. Besides physical scars and infections, NSSI can result in feelings of guilt and shame (Long, 2018; Rosenrot & Lewis, 2020) and evoke stigmatising behaviour from family, peers (Doyle, 2017; Oldershaw et al., 2008), and even health care professionals (Saunders et al., 2012). While NSSI is typically conceptualised as non-suicidal, prior NSSI behaviour has been shown to be a robust risk factor and predictor of suicidal thoughts and behaviours (Franklin et al., 2017; Griep & MacKinnon, 2022; Kiekens et al., 2018; Turner et al., 2013; Whitlock et al., 2013). NSSI is a highly prevalent problem. Lifetime prevalence rates have been estimated at 4%–6% in general adult populations (Klonsky, 2011; Liu, 2023) and up to 45% in adult clinical populations (Andover & Gibb, 2010). Among community (non-clinical) adolescent samples, meta-analytic findings suggest a lifetime prevalence of 16% (Farkas et al., 2024), with even higher rates reported in clinical or younger samples (e.g., Swannell et al., 2014). Given the severity and prevalence of NSSI, it is important for practitioners to assess which patients are at risk of NSSI behaviour. Assessment methods to detect and predict NSSI could provide a valuable opportunity to prevent NSSI and suicidal behaviour.

Potential predictors and risk factors for NSSI have been investigated for over a decade. Nevertheless, results from a recent meta-analysis on studies longitudinally pre-

dicting NSSI revealed that most factors show weak predictive utility, suggesting limited clinical value (Fox et al., 2015). One notable exception was prior NSSI behaviour, which has been identified as the strongest risk factor for future NSSI. Results from individual studies have shown that prior self-harm behaviour outperforms other predictors when longitudinally forecasting self-harm (Janis & Nock, 2008), and have revealed that prior NSSI behaviour is the only significant predictor for NSSI at different follow-up moments (Tuisku et al., 2014). Moreover, results from a recent study using machine learning techniques showed that prior NSSI behaviour was the most important predictor across several time points (Fox et al., 2019). As such, assessing prior NSSI behaviour could be a valuable strategy for predicting future NSSI and related behaviours in clinical contexts.

Importantly, however, merely asking individuals about prior NSSI behaviour in clinical settings could be problematic. Patients might be reluctant to disclose NSSI behaviour given the potential negative consequences, such as feelings of shame, fear of stigmatization, and fear of hospitalization (Long, 2018; MacDonald et al., 2020; Simone & Hamza, 2020). The addition of implicit measures to screening procedures for NSSI risk detection could provide a solution to this problem. Unlike self-report measures, implicit measures are used to assess responses occurring under automaticity conditions (De Houwer et al., 2009; De Houwer & Moors, 2012). For instance, studies have shown that responding on certain implicit measures, such as the Implicit Association Test (IAT; Greenwald et al., 1998) and some of its variants, is far less controllable than responding on self-report measures (e.g., Agosta et al., 2011; Egloff & Schmukle, 2002; Kim, 2003; Stieger et al., 2011). As such, these measures could be less susceptible to deception.

Because of their potential benefits over self-report measures, scholars have developed implicit measures for predicting NSSI and other self-harm behaviours (e.g., Nock et al., 2010; Nock & Banaji, 2007). Notably, most of these measures do not target predictors that have previously been shown to be strong forecasters of NSSI (for an exception, see Gray et al., 2021). For instance, one of the most frequently used implicit measures in this domain, referred to as the Self-Injury Implicit Association Test (SI-IAT; Nock & Banaji, 2007), targets self-identification with NSSI behaviour. In the SI-IAT, participants categorise stimuli regarding the self (e.g., the word "me") and others (e.g., the word "them") alongside self-injurious stimuli and non-self-injurious stimuli (e.g., pictures of skin that has (not) been cut) as fast as possible using two keys on the keyboard. When participants perform better on trials where stimuli regarding the self and self-injurious stimuli share the same response key than on trials where stimuli regarding the self and non-self-injurious stimuli share the same response key, it is inferred that these individuals automatically identify themselves with self-injurious behaviour. While several studies have shown that the SI-IAT can discriminate between injury groups and non-injury groups (Glenn et al., 2017; Nock & Banaji, 2007; Powers et al., 2021), evidence regarding its utility in prospectively predicting NSSI behaviour has been mixed. Two studies have demonstrated that the SI-IAT predicts NSSI over time (Cha et al., 2016; Glenn et al., 2016),

while other studies have failed to find such effects (Cha et al., 2016; Franklin et al., 2014; Glenn & Klonsky, 2011; Powers et al., 2021).

A promising new avenue for NSSI risk detection in clinical practice may involve the development of measures that (a) target predictors shown to be strong forecasters of NSSI and (b) are less susceptible to deception. An implicit measure was recently developed that assesses beliefs regarding past behaviour, referred to as the past nonsuicidal self-injury IAT (P-NSSI-IAT; Cathelyn et al., 2021). The P-NSSI-IAT follows the same procedure as the SI-IAT, with the exception that its stimuli consist of statements (regarding past NSSI behaviour) rather than single words or pictures. If participants respond faster on trials in which the same response key is used for categorizing statements that are inherently true (e.g., "I'm pressing computer keys") and statements regarding past NSSI behaviour (e.g., "I have carved my skin on purpose") than on trials where the same response key is used for categorising statements that are inherently false (e.g., "I'm climbing a mountain") and statements regarding past NSSI behaviour, this might indicate that these individuals automatically endorse the belief that they have engaged in NSSI in the past. Two previous studies investigated the predictive utility of the P-NSSI-IAT and found that this measure discriminated well between participants who reported to have previously engaged in NSSI behaviour (i.e., cutting or carving of the skin) and participants who reported to have never engaged in NSSI. P-NSSI-IAT scores also prospectively predicted NSSI behaviour over one month (Franklin et al., 2017) and predicted this outcome above and beyond other known risk factors for NSSI (i.e., hopelessness, frequency of past NSSI, and prior suicidal thoughts).

Importantly, however, these studies were conducted in general online samples which were recruited through Prolific Academic (https://www.prolific.co/). To examine the clinical utility of the P-NSSI-IAT, it is important to investigate whether these previous findings generalise to a clinical population. In the current study, we will target outpatients and investigate (a) whether and (b) how well the P-NSSI-IAT discriminates between patients who report having recently engaged in NSSI behaviour (i.e., cutting or carving of the skin in the past year and past month) and patients who report having never engaged in NSSI behaviour. We will also examine (c) whether P-NSSI-IAT scores independently predict self-rated past year, past month, and future likelihood of NSSI, and (d) whether P-NSSI-IAT scores predict these outcomes above and beyond known risk factors for NSSI (i.e., gender, age, hopelessness, psychiatric disorders typically related to NSSI, and prior suicidal thoughts; e.g., Fox et al., 2015; Klonsky & Muehlenkamp, 2007). Note that the main aim of this study is to validate the P-NSSI-IAT by assessing its ability to detect prior NSSI behaviour in a sample of clinical patients. For practical reasons, we do not assess the predictive validity of the P-NSSI-IAT. We include a future likelihood measure for exploratory purposes (see Cathelyn et al., 2021). Appendix A provides an overview of the research questions, hypotheses, sampling plan, analysis plan, and interpretations given different outcomes.

Method

Participants

We targeted patients who receive outpatient treatment for various conditions. Participants were recruited through clinical psychologists in Flanders. Recruitment materials explicitly invited individuals with and without a history of NSSI to participate, to allow for comparison across groups. We contacted clinicians through the Flemish Association for Clinical Psychologists, Facebook groups for licensed clinical psychologists, mental health outreach events, and a website where Flemish people can search for licensed clinical psychologists. Clinicians willing to collaborate were asked to invite patients to participate in the study if (a) their first language is Dutch, (b) they were capable of conducting the study (i.e., not having a cognitive impairment and not being in serious crisis), and (c) they were between 18 and 30 years old. This latter inclusion criterion was applied because NSSI behaviour is more prevalent in this age group than in older age groups (Swannell et al., 2014). The patients' eligibility to participate in the study was evaluated by the clinicians. Participants were reimbursed for their time and effort with a voucher (€10).

A total of 207 participants started the study. Of those, 28 participants dropped out after reading the study information. The data of participants were excluded if they did not provide complete (questionnaire and/or P-NSSI-IAT) data (32 participants). As recommended by Greenwald et al. (2003), we also excluded the data of participants with response latencies below 300 ms on 10% or more of critical P-NSSI-IAT trials or with P-NSSI-IAT error rates above 30% across the entire task, and/or above 40% for any of the critical blocks (34 participants). We applied these exclusion criteria for several reasons. First, the researchers had minimal control over what participants were doing as the study was conducted online. Second, unlike student participants who completed most past IAT studies, the participants in the current study were likely not familiar with reaction time measures. As such, more participants might have failed to adhere to instructions than in previous IAT studies. Third, the current sample is relatively small; thus, the impact of outlier scores may be stronger. Finally, participants might have been motivated to conceal their NSSI behaviour and thus might have tried to alter their P-NSSI-IAT scores. Studies show that participants who try to alter their IAT scores tend to produce higher error rates (e.g., because they believe that increasing the number of errors is a valid faking strategy, Röhner et al., 2013; Steffens, 2004). However, for exploratory purposes, we also conducted the analyses including the data of participants who met the latter exclusion criteria and will report whether doing so affected the results.

Further, we excluded the data of participants who could not be assigned to the past year NSSI group or to the no history of NSSI group (45 participants). This occurred when participants indicated they had not engaged in skin cutting in the past year, but had (a) engaged in skin cutting more than a year ago, (b) used another NSSI method during

their lifetime, and/or (c) attempted suicide using cutting of the skin as a method (because the category labels and items of the P-NSSI-IAT do not distinguish between suicidal and non-suicidal self-injury). Relatedly, we had planned to exclude participants if there was uncertainty regarding their NSSI group status (e.g., because of inconsistencies in the self-report data, for example, when participants report having engaged in NSSI in the past month, but not in the past year), but no participants met this criterion.

The main effect of interest is the difference in P-NSSI-IAT scores between the past NSSI group and the no history of NSSI group. In line with previous studies (Cathelyn et al., 2021), we tested differences in P-NSSI-IAT scores between individuals who had never engaged in NSSI on the one hand and individuals who had engaged in NSSI on the other hand. In the latter group, we distinguished between individuals who had engaged in NSSI in the past year and individuals who had engaged in NSSI in the past month. Given that it is more feasible to recruit a sufficient number of participants who have engaged in NSSI in the past year than a sufficient number of participants who have engaged in NSSI in the past month, the required sample sizes for the current study were calculated based on the estimated effect size for an independent samples *t*-test comparing P-NSSI-IAT scores from the past year NSSI group and the no history of NSSI group. Note that previous studies testing the predictive utility of implicit measures for NSSI typically focused on detecting lifetime prevalence of NSSI, but this may be less relevant for clinical purposes (Powers et al., 2021). We focus on more recent behaviour and calculate the power of the study to detect past year NSSI which is of clinical importance. According to conventions for effect sizes (Cohen, 1988), we found medium to large effect sizes in our previous studies when comparing P-NSSI-IAT scores between a past year NSSI group and a no history of NSSI group (with *d*s ranging from 0.68 to 0.82; Cathelyn et al., 2021). Given that previous studies have demonstrated that group differences between prediction scores for NSSI tend to be more modest in clinical than in community samples (Franklin et al., 2017; Sohn et al., 2021), we applied an estimated effect of $d = .60$ in the power analysis, which is slightly smaller than the lowest observed effect size.

Results of the power analysis showed that a total of 146 participants would allow for detecting the estimated effect size with 85% power in a one-tailed *t*-test ($\alpha = .05$) with an estimated group-allocation ratio (i.e., the proportion of cases and controls) equal to 5 (i.e., 24 cases and 122 controls). Because our previous studies were conducted in general online samples, and we conducted pre-screening studies to recruit many cases, we could not base the estimated group-allocation ratio on our previous studies. Therefore, we chose to apply a conservative estimate of the group allocation ratio in the power analysis and registered that we would stop data collection when 146 participants fully completed the study.

While we indicated in our preregistered protocol that data collection would stop when 146 participants had fully completed the study, we noted during data collection that our power analysis was based on retaining 146 participants after applying prede-

fined exclusion criteria (e.g., based on IAT performance or unclear group classification). Due to substantial and prolonged difficulties in recruiting participants from the targeted clinical population, high exclusion rates, and given that participants were compensated upon completion of the study procedure regardless of later data exclusion, we ceased data collection due to time and financial constraints after three full years of recruitment. At that point, 147 participants had completed the full study and received reimbursement.

After applying the preregistered exclusion criteria, the final analysed sample consisted of 68 participants (mean age = 30, *SD* = 12; 24% male, 75% female, 1% other identity), including 45 participants who engaged in NSSI in the past year and 23 participants without a history of NSSI. A sensitivity power analysis (G*Power 3.1) indicated that this final sample size provides us with 75% power to detect the estimated effect size of d = 0.60. Participants from the past NSSI group reported engaging in skin cutting 4.31 on average over the past year (*SD* = 3.17). Based on self-report symptom screening, 44% of participants reported symptoms consistent with a mood disorder, 46% with an anxiety disorder, 21% with an eating disorder, and 25% with a substance use disorder. Details on screening criteria are provided in the Measures of Risk Factors section.

Materials and Apparatus

The study was built using lab.js, a tool for creating browser-based studies. The study was hosted online, and participants received a link to the study via their treating clinician and were asked to complete the study using their laptop or desktop. All materials were translated into Dutch using back-translation, except for the Patient Health Questionnaire (PHQ; Spitzer et al., 1999). We used the Dutch version of the PHQ as translated by the MAPI Research Institute (see https://www.phqscreeners.com/).

P-NSSI-IAT

The P-NSSI-IAT followed the same procedure as in previous studies (Cathelyn et al., 2021). Participants were instructed to categorise statements regarding past (non-) NSSI behaviour and statements that are inherently true or false as fast as possible using the E- or I-key on the keyboard. On each trial, a statement appeared in the middle of the screen until participants pressed one of the valid response keys. If the response was correct, the statement disappeared, and the next statement was presented 400 ms later. If the response was incorrect, a red cross replaced the statement for 200 ms, and the next statement appeared 400 ms after the red cross appeared. Category labels were presented in the top left and right corners of the screen to aid categorization. The category labels and stimuli are listed in Appendix B. We used the font Helvetica with font size 15 for the category labels and font size 18 for the statements. The category labels were presented in upper case letters, and the statements were presented in lower case letters. The P-NSSI-IAT procedure was presented full screen.

The P-NSSI-IAT consists of seven blocks and 192 trials. Before the first block began, participants received instructions emphasizing speed and accuracy. Specifically, they were asked to respond as quickly as possible while trying to avoid too many mistakes, to avoid distractions, to pay attention, and to keep their fingers on the E and I keys throughout the task. The full task instructions are provided in Appendix B. Before the start of all subsequent blocks, participants were reminded to try to respond as quickly as possible. Participants were able to proceed to the next block by pressing the space bar.

In the first block of the P-NSSI-IAT, participants practiced categorizing four statements regarding past NSSI behaviour (e.g., "I have carved my skin on purpose") using the E-key, and four statements regarding past non-NSSI behaviour (e.g., "I have carved my skin not once") using the I-key. In the second block, participants practiced categorizing four statements that are inherently true (e.g., "I'm pressing computer keys") using the E-key, and four inherently false statements (e.g., "I'm playing football") using the I-key. Both practice blocks consisted of 32 trials during which the eight statements were each presented four times. In the next two blocks, participants categorised statements from all four categories simultaneously using the practiced key assignment (i.e., the E-key for sentences regarding past NSSI behaviour and statements that are inherently true, and the I-key for sentences regarding past non-NSSI behaviour and sentences that are inherently false). In the combined blocks (16 and 32 trials) the 16 statements were each presented three times. Following this, there was another practice block in which only statements regarding past (non-) NSSI behaviour needed to be categorised, but this time with the response key assignment reversed (i.e., the E-key for statements regarding past non-NSSI behaviour and the I-key for statements regarding past NSSI behaviour). This practice block consisted of 32 trials during which the eight statements were each presented four times. In the final two combined blocks, participants categorised statements from all four categories using the new response key assignment. The order of the trials was determined randomly for each block and participant.

Self-Report Measures of NSSI Behaviour

To assess the outcome variables of interest, participants were asked how many times they had intentionally cut or carved their skin without suicidal intent in the past 12 months and in the past 30 days. These two questions were rated on a scale ranging from 0 to 10+ times. Participants were also invited to indicate how likely they would be to intentionally cut or carve their skin without intending to kill themselves in the future. This question was rated on a Likert scale ranging from 0 (*low/little*) to 4 (*very much/severe*).

To determine whether the data of participants should be excluded (see exclusion criteria), participants were also asked about lifetime skin cutting with and without suicidal intent, and lifetime NSSI using any other method. Participants were asked to indicate "yes" or "no" when answering these questions. All questions regarding NSSI behaviour

were based on or adapted from the Deliberate Self-Harm Inventory (DSHI; Gratz, 2001) and the Self-Injurious Thoughts and Behavior Interview (SITBI; Nock et al., 2007).

Measures of Risk Factors

Indicators of psychiatric diagnoses were assessed using the PHQ, a self-administered questionnaire that screens for five common types of psychiatric disorders: mood, anxiety, eating, substance abuse, and somatoform disorders. Note that we only included the first four modules of the questionnaire because these disorders are frequently related to NSSI (e.g., Klonsky & Muehlenkamp, 2007). Hopelessness was assessed using the Beck Hopelessness Scale (BHS; Beck et al., 1974). The BHS assesses positive and negative beliefs about the future and consists of 20 items (e.g., "my future seems dark to me"). Participants were asked to evaluate these statements as true or false. Finally, the frequency of prior suicidal thoughts was assessed through a question that was adapted from the SITBI: "During how many separate times in your life have you had thoughts of killing yourself?". Participants indicated this frequency on a scale ranging from 0 to 10+ times.

Procedure

After providing informed consent, participants indicated their age and gender and completed the P-NSSI-IAT. Afterwards, participants were reminded about the anonymous nature of the study (to reduce socially desirable responding) and answered the questions regarding NSSI behaviour. In the next phase, participants completed the PHQ, the BHS, and prior suicidal thoughts. At the end of the study, participants were referred to several help sources for coping with NSSI and suicidal thoughts and behaviour. Participants were also referred to a separate website where they were asked to provide the information needed for compensation purposes.

Data Pre-Processing

To calculate P-NSSI-IAT scores we used the D4 scoring algorithm (Greenwald et al., 2003). This algorithm assesses the difference in reaction times between the first two and the second two combined blocks while correcting the latencies for individual variability. In line with this algorithm, trials with latencies longer than 10,000 ms were removed and latencies on trials with an incorrect response were replaced with the participant's block mean (based on correct responses) plus a 600 ms penalty. Reaction times on trials of the first two combined blocks were subtracted from reaction times on trials of the second two combined blocks, such that higher scores indicate faster responding during combined blocks in which statements regarding past NSSI behaviour and statements that are logically true share the same response key. The Spearman-Brown corrected split-half reliability for the P-NSSI-IAT was .88, based on the correlation between D-scores computed from odd- versus even-numbered trials.

The presence of any mood, anxiety, eating or substance abuse disorder indicators was determined using the manual and scoring instructions for the PHQ (Spitzer et al., 1999). No distinction was made between specific disorders (e.g., major depressive disorder, other depressive syndrome, etc.). If a participant met the scoring threshold for any specific disorder subtype (e.g., major depressive disorder), the overarching category (e.g., mood disorder) was scored as "1", otherwise it was scored as "0". The Cronbach's alpha for the modules of the PHQ were .90 (mood disorders), .71 (anxiety disorders), .59 (eating disorders), and .71 (substance abuse disorders). For hopelessness, each BHS item indicative of hopelessness was scored as "1", and total scores were calculated by summing the 20 items (Cronbach's alpha = .91).

To assign participants to the NSSI groups, we used the self-reported NSSI frequencies and ratings. Table 1 provides an overview of the criteria of the self-report questionnaires that were used to assign participants to the NSSI groups and includes the number of participants per NSSI group.

Table 1

Grouping of Participants and Number of Participants per NSSI Group

		NSSI frequencies and ratings		
Group	***n***	**Past year NSSI**	**Past month NSSI**	**Future likelihood NSSI**
Past year NSSI group	45	> 0	= 0 or > 0	–
Past month NSSI group	23	> 0	> 0	–
No history of NSSI group	23	= 0	= 0	–
Low future likelihood NSSI group	26	= 0 or > 0	= 0 or > 0	0 – 2
High future likelihood NSSI group	42	= 0 or > 0	= 0 or > 0	3 – 4

Note. NSSI = nonsuicidal self-injury.

Results

Group Differences in P-NSSI-IAT Scores

To examine whether P-NSSI-IAT scores distinguish between participants with and without a history of NSSI, we conducted two independent samples *t*-tests. First, participants who reported no lifetime history of NSSI (M = 0.04, SD = 0.54) had significantly lower IAT scores than participants who reported engaging in past-year NSSI (M = 0.42, SD = 0.48), $t(66) = 2.96$, $p = .002$, $d = 0.76$. Second, participants without a history of NSSI also had lower IAT scores than participants who reported engaging in past-month NSSI (M = 0.34, SD = 0.57), $t(44) = 1.85$, $p = .036$, $d = 0.54$.

Receiver operating characteristic (ROC) analyses were used to assess how well P-NSSI-IAT scores discriminated between individuals with and without a history of NSSI. For past-year NSSI, the area under the curve (AUC) was 0.72 (95% CI [0.58, 0.86]), indicating good discriminative ability. In previous studies (Cathelyn et al., 2021), we established cut-off scores for the P-NSSI-IAT to maximise sensitivity (0.16) or specificity (0.65) while retaining fair specificity and sensitivity, respectively. At this sensitivity-maximizing cut-off, sensitivity was 0.78 and specificity was 0.57. At the specificity-maximizing cutoff, sensitivity dropped to 0.36, while specificity increased to 0.87. When using a cutoff of 0.00 (a theoretically relevant cut-off; Cvencek et al., 2021), sensitivity was 0.82, and specificity was 0.43.

For past-month NSSI, the AUC was 0.68 (95% CI [0.52, 0.85]). Sensitivity and specificity at the 0.16 cutoff were 0.74 and 0.57, respectively. At the 0.65 cutoff, sensitivity was 0.39 and specificity was 0.87. Using a cutoff of 0.00 yielded a sensitivity of 0.74 and specificity of 0.43.

Predictive Validity

We conducted three logistic regression analyses to test whether P-NSSI-IAT scores independently predicted past-year NSSI, past-month NSSI, and self-rated future likelihood of NSSI. P-NSSI-IAT scores significantly predicted past-year NSSI (*OR* = 4.19, 95% CI [1.53, 13.10], $p = .008$), but not past-month NSSI (*OR* = 2.76, 95% CI [0.94, 9.37], $p = .078$), or self-rated likelihood of future NSSI (*OR* = 2.75, 95% CI [0.87, 10.75], $p = .108$).

To evaluate the incremental predictive value of the P-NSSI-IAT above and beyond self-report risk factors, we conducted hierarchical logistic regression analyses. Variance inflation factors (VIFs) for all predictors were below 2 (range: 1.06–1.95), indicating no problematic multicollinearity. For past-year NSSI, risk factors (hopelessness, suicidal thoughts, and mood and anxiety disorder indicators) entered in Step 1 significantly predicted NSSI, $\chi^2(4) = 40.28$, $p < .001$. Suicidal thoughts (*OR* = 1.80, 95% CI [1.29, 2.91]) and mood disorder indicators (*OR* = 8.52, 95% CI [0.97, 117.57]) showed the strongest associations. Adding the P-NSSI-IAT in Step 2 improved model fit, but this improvement was not statistically significant, $\Delta\chi^2(1) = 2.51$, $p = .113$.

Similar results were found for past-month NSSI: Step 1 model $\chi^2(4) = 43.53$, $p < .001$; $\Delta\chi^2(1)$ for Step 2 = 0.05, $p = .827$; and for self-rated likelihood of future NSSI: Step 1 model $\chi^2(3) = 28.09$, $p < .001$; $\Delta\chi^2(1)$ for Step 2 = 0.32, $p = .570$. For past-month NSSI, suicidal thoughts (*OR* = 2.19, 95% CI [1.34, 5.10]) and mood disorder indicators (*OR* = 50.43, 95% CI [3.25, 2471.61]) were associated with higher odds. For future likelihood, hopelessness was the only significant predictor (*OR* = 1.32, 95% CI [1.10, 1.65]).

General Discussion

This study evaluated the utility of the P-NSSI-IAT as a tool for detecting past NSSI behaviour in a clinical population. Whereas prior research demonstrated the predictive and discriminative potential of this implicit measure in general population samples (Cathelyn et al., 2021), the present findings are the first to extend this work to a sample of outpatients, thereby addressing the clinical relevance of the P-NSSI-IAT.

Our findings support the discriminative validity of the P-NSSI-IAT in this clinical context. Patients who reported past-year or past-month NSSI exhibited significantly higher IAT scores than patients with no history of NSSI, with medium to large effect sizes. ROC analyses further indicated that the P-NSSI-IAT was able to distinguish patients with and without recent NSSI with fair to good accuracy (AUCs of 0.68–0.73). Sensitivity and specificity estimated at pre-established and theory-informed cutoffs were comparable to those observed in prior research, suggesting robustness across samples and potential applicability in clinical screening contexts.

Logistic regression analyses showed that P-NSSI-IAT scores independently predicted past-year NSSI, providing further support for its clinical relevance. However, effects were weaker and not statistically significant for past-month NSSI or perceived future likelihood of NSSI. Moreover, when known risk factors, such as hopelessness, suicidal thoughts, and mood and anxiety disorder indicators, were statistically controlled for, the IAT no longer significantly improved predictive accuracy. These results suggest that while the IAT scores may capture relevant aspects of NSSI-related beliefs, their added value over self-report instruments is currently limited in multivariable prediction models. This might imply that the IAT may mainly be useful in settings where patients are unwilling or unable to disclose NSSI.

A notable pattern in the findings is that the P-NSSI-IAT predicted past-year NSSI, but not past-month NSSI or anticipated future likelihood of NSSI. First, with regard to past-month NSSI, one likely explanation is limited statistical power: the past-month group was the smallest of the three and may not have allowed for the detection of effects of moderate size. We therefore refrain from drawing strong conclusions about the apparent difference between past-year and past-month NSSI. Second, regarding the future likelihood of NSSI, this variable was assessed via a single self-rated expectation, which may be influenced by factors beyond past behaviour, such as motivation to recover, social desirability, limited insight, self-efficacy expectations, or perceived ability to regulate emotions. These factors may contribute to divergence between explicit expectations and the tendencies captured by the P-NSSI-IAT. Future studies should examine whether P-NSSI-IAT scores can meaningfully predict future behaviour when assessed in a longitudinal design.

Several limitations of the study merit consideration. First, the sample size, though adequate for detecting large effects, limited power for identifying smaller effects, particularly for the recent NSSI group. This may partly explain why some effects were in the

expected direction but did not reach statistical significance. In addition, group sizes were unequal in several key comparisons (e.g., $N = 45$ for the past-year NSSI group vs. $N = 23$ for the no-history group), which may reduce the precision of some estimates and limit confidence in the stability of the findings. Because unequal group sizes can increase the likelihood of sampling error and affect the reliability of statistical estimates, particularly in small samples, future studies with larger and more balanced groups are needed to replicate these results.

Second, although a clinical sample was recruited, the online design prevented confirmation of diagnoses and limited our ability to characterize heterogeneity in treatment status or context. This restricts conclusions about whether the IAT performs differently across clinical subgroups. Future in-person studies with structured assessments (e.g., structured interviews) would allow more fine-grained examination of such differences. Third, the use of self-report measures for key variables, including NSSI history, recent episodes, and risk factors, may have introduced underreporting or misclassification, given the socially sensitive nature of NSSI. Such inaccuracies could weaken observed group differences, reduce the predictive strength of odds ratios, and lower the apparent discriminative accuracy of the P-NSSI-IAT. Future research may benefit from incorporating clinician assessments or additional corroborating information (e.g., clinical records or ecological momentary reporting). Finally, although self-rated future likelihood of NSSI was included as an exploratory outcome, the absence of longitudinal follow-up means that the current study cannot draw conclusions about actual prospective predictive validity. Longitudinal designs will be essential to evaluate whether the P-NSSI-IAT can meaningfully predict future NSSI behaviour.

Taken together, these findings provide initial evidence that the P-NSSI-IAT can be a valid tool for detecting past NSSI in clinical populations. Although the P-NSSI-IAT may prove useful in contexts where disclosure is limited, it is important to note that we found no evidence for incremental predictive value above self-report risk factors in this clinical sample. Future research could evaluate the IAT's utility in prospective designs, examine cutoffs tailored to specific clinical contexts, and test whether combining implicit and explicit indicators enhances decision-making in clinical assessments of NSSI.

Funding: This manuscript is supported by Ghent University Methusalem grant 01M00209 to JDH.

Acknowledgments: The authors would like to thank Leonie Callewaert for her help with participant recruitment.

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

Ethics Statement: The study was approved by the UZ Gent Medical Ethics Committee (reference number 2022/3739). Informed consent was obtained from all participants.

Preregistration: The accepted protocol is registered at PsychArchives: https://doi.org/10.23668/psycharchives.12576

Reporting Guidelines: This study is reported in accordance with the APA Journal Article Reporting Standards (JARS).

Data Availability: All materials, processing and analysis code, (pseudonymised) raw and processed data are available on the Open Science Framework (https://osf.io/qb6d8).

Supplementary Materials

The Supplementary Materials contain the following items:

- **Preregistration for the study** (Cathelyn et al., 2023S)
- **Online appendices** (Cathelyn et al., 2026S):
  - *Appendix A.* Overview of the preregistered research questions, hypotheses, sampling plan, analysis plan, and interpretation of possible outcomes.
  - *Appendix B.* Full Past Nonsuicidal Self-Injury Implicit Association Test (P-NSSI-IAT) materials, including category labels, stimuli, and task instructions.
- **Study materials, processing and analysis code, (pseudonymised) raw and processed data** (Linthout et al., 2026S)

Index of Supplementary Materials

Cathelyn, F., Linthout, T., Van Dessel, P., Claes, L., & De Houwer, J. (2023S). *Can a variant of the Implicit Association Test detect nonsuicidal self-injury in a clinical population? A registered report* [Preregistration]. PsychArchives. https://doi.org/10.23668/psycharchives.12576

Cathelyn, F., Van Dessel, P., Linthout, T., Claes, L., & De Houwer, J. (2026S). *Supplementary materials to "Can a variant of the Implicit Association Test detect nonsuicidal self-injury in a clinical population? A registered report"* [Online appendices]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.22340

Linthout, T., Van Dessel, P., & Callewaert, L. (2026S). *Can a variant of the Implicit Association Test detect nonsuicidal self-injury in a clinical population? A registered report* [Research data, code, and materials]. OSF. https://osf.io/qb6d8

References

Agosta, S., Ghirardi, V., Zogmaister, C., Castiello, U., & Sartori, G. (2011). Detecting fakers of the autobiographical IAT. *Applied Cognitive Psychology, 25*(2), 299–306. https://doi.org/10.1002/acp.1691

Andover, M. S., & Gibb, B. E. (2010). Non-suicidal self-injury, attempted suicide, and suicidal intent among psychiatric inpatients. *Psychiatry Research, 178*(1), 101–105. https://doi.org/10.1016/j.psychres.2010.03.019

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The Hopelessness Scale. *Journal of Consulting and Clinical Psychology, 42*(6), 861–865. https://doi.org/10.1037/h0037562

Cathelyn, F., Van Dessel, P., & De Houwer, J. (2021). Predicting nonsuicidal self-injury using a variant of the Implicit Association Test. *Suicide and Life-Threatening Behavior, 51*(6), 1259–1271. https://doi.org/10.1111/sltb.12808

Cha, C. B., Augenstein, T. M., Frost, K. H., Gallagher, K., D'Angelo, E. J., & Nock, M. K. (2016). Using implicit and explicit measures to predict nonsuicidal self-injury among adolescent inpatients. *Journal of the American Academy of Child and Adolescent Psychiatry, 55*(1), 62–68. https://doi.org/10.1016/j.jaac.2015.10.008

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.1016/C2013-0-10517-X

Cvencek, D., Meltzoff, A. N., Maddox, C. D., Nosek, B. A., Rudman, L. A., Devos, T., Dunham, Y., Baron, A. S., Steffens, M. C., Lane, K., Horcajo, J., Ashburn-Nardo, L., Quinby, A., Srivastava, S. B., Schmidt, K., Aidman, E., Tang, E., Farnham, S., Mellott, D. S., . . . Greenwald, A. G. (2021). Meta-analytic use of balanced identity theory to validate the Implicit Association Test. *Personality & Social Psychology Bulletin, 47*(2), 185–200. https://doi.org/10.1177/0146167220916631

De Houwer, J., & Moors, A. (2012). How to define and examine implicit processes? In R. Proctor & E. J. Capaldi (Eds.), *Psychology of science: Implicit and explicit processes* (pp. 183–198). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199753628.003.0008

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135*(3), 347–368. https://doi.org/10.1037/a0014211

Doyle, L. (2017). Attitudes toward adolescent self-harm and its prevention: The views of those who self-harm and their peers. *Journal of Child and Adolescent Psychiatric Nursing, 30*(3), 142–148. https://doi.org/10.1111/jcap.12186

Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology, 83*(6), 1441–1455. https://doi.org/10.1037/0022-3514.83.6.1441

Farkas, B. F., Takacs, Z. K., Kollárovics, N., & Balazs, J. (2024). The prevalence of self-injury in adolescence: A systematic review and meta-analysis. *European Child & Adolescent Psychiatry, 33*(10), 3439–3458. https://doi.org/10.1007/s00787-023-02264-y

Fox, K. R., Franklin, J. C., Ribeiro, J. D., Kleiman, E. M., Bentley, K. H., & Nock, M. K. (2015). Meta-analysis of risk factors for nonsuicidal self-injury. *Clinical Psychology Review, 42*, 156–167. https://doi.org/10.1016/j.cpr.2015.09.002

Fox, K. R., Huang, X., Linthicum, K. P., Wang, S. B., Franklin, J. C., & Ribeiro, J. D. (2019). Model complexity improves the prediction of nonsuicidal self-injury. *Journal of Consulting and Clinical Psychology, 87*(8), 684–692. https://doi.org/10.1037/ccp0000421

Franklin, J. C., Puzia, M. E., Lee, K. M., & Prinstein, M. J. (2014). Low implicit and explicit aversion toward self-cutting stimuli longitudinally predict nonsuicidal self-injury. *Journal of Abnormal Psychology, 123*(2), 463–469. https://doi.org/10.1037/a0036436

Franklin, J. C., Ribeiro, J. D., Fox, K. R., Bentley, K. H., Kleiman, E. M., Huang, X., Musacchio, K. M., Jaroszewski, A. C., Chang, B. P., & Nock, M. K. (2017). Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin, 143*(2), 187–232. https://doi.org/10.1037/bul0000084

Glenn, C. R., Kleiman, E. M., Cha, C. B., Nock, M. K., & Prinstein, M. J. (2016). Implicit cognition about self-injury predicts actual self-injurious behavior: Results from a longitudinal study of adolescents. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 57*(7), 805–813. https://doi.org/10.1111/jcpp.12500

Glenn, C. R., & Klonsky, E. D. (2011). Prospective prediction of nonsuicidal self-injury: A 1-year longitudinal study in young adults. *Behavior Therapy, 42*(4), 751–762. https://doi.org/10.1016/j.beth.2011.04.005

Glenn, J. J., Werntz, A. J., Slama, S. J. K., Steinman, S. A., Teachman, B. A., & Nock, M. K. (2017). Suicide and self-injury-related implicit cognition: A large-scale examination and replication. *Journal of Abnormal Psychology, 126*(2), 199–211. https://doi.org/10.1037/abn0000230

Gratz, K. L. (2001). Measurement of deliberate self-harm: Preliminary data on the deliberate self-harm inventory. *Journal of Psychopathology and Behavioral Assessment, 23*(4), 253–263. https://doi.org/10.1023/A:1012779403943

Gray, N. S., Knowles, J., George, D., Harvey, A., Powell, R., Zadeh, M. V., Wansing, C., & Snowden, R. J. (2021). Explicit and implicit hopelessness and self-injury. *Suicide and Life-Threatening Behavior, 51*(3), 606–615. https://doi.org/10.1111/sltb.12743

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Griep, S. K., & MacKinnon, D. F. (2022). Does nonsuicidal self-injury predict later suicidal attempts? A review of studies. *Archives of Suicide Research, 26*(2), 428–446. https://doi.org/10.1080/13811118.2020.1822244

Janis, I. B., & Nock, M. K. (2008). Behavioral forecasts do not improve the prediction of future behavior: A prospective study of self-injury. *Journal of Clinical Psychology, 64*(10), 1164–1174. https://doi.org/10.1002/jclp.20509

Kiekens, G., Hasking, P., Boyes, M., Claes, L., Mortier, P., Auerbach, R. P., Cuijpers, P., Demyttenaere, K., Green, J. G., Kessler, R. C., Myin-Germeys, I., Nock, M. K., & Bruffaerts, R. (2018). The associations between non-suicidal self-injury and first onset suicidal thoughts and behaviors. *Journal of Affective Disorders, 239*, 171–179. https://doi.org/10.1016/j.jad.2018.06.033

Kim, D.-Y. (2003). Voluntary controllability of the Implicit Association Test (IAT). *Social Psychology Quarterly, 66*(1), 83–96. https://doi.org/10.2307/3090143

Klonsky, E. D. (2011). Non-suicidal self-injury in United States adults: Prevalence, sociodemographics, topography and functions. *Psychological Medicine, 41*(9), 1981–1986. https://doi.org/10.1017/S0033291710002497

Klonsky, E. D., & Muehlenkamp, J. J. (2007). Self-injury: A research review for the practitioner. *Journal of Clinical Psychology, 63*(11), 1045–1056. https://doi.org/10.1002/jclp.20412

Liu, R. T. (2023). The epidemiology of non-suicidal self-injury: Lifetime prevalence, sociodemographic and clinical correlates, and treatment use in a nationally representative sample of adults in England. *Psychological Medicine, 53*(1), 274–282. https://doi.org/10.1017/S003329172100146X

Long, M. (2018). 'We're not monsters … we're just really sad sometimes:' Hidden self-injury, stigma and help-seeking. *Health Sociology Review, 27*(1), 89–103. https://doi.org/10.1080/14461242.2017.1375862

MacDonald, S., Sampson, C., Turley, R., Biddle, L., Ring, N., Begley, R., & Evans, R. (2020). Patients' experiences of emergency hospital care following self-harm: Systematic review and thematic synthesis of qualitative research. *Qualitative Health Research, 30*(3), 471–485. https://doi.org/10.1177/1049732319886566

Nock, M. K. (2010). Self-injury. *Annual Review of Clinical Psychology, 6*(1), 339–363. https://doi.org/10.1146/annurev.clinpsy.121208.131258

Nock, M. K., & Banaji, M. R. (2007). Assessment of self-injurious thoughts using a behavioral test. *The American Journal of Psychiatry, 164*(5), 820–823. https://doi.org/10.1176/ajp.2007.164.5.820

Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). Self-injurious thoughts and behaviors interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment, 19*(3), 309–317. https://doi.org/10.1037/1040-3590.19.3.309

Nock, M. K., Park, J. M., Finn, C. T., Deliberto, T. L., Dour, H. J., & Banaji, M. R. (2010). Measuring the suicidal mind: Implicit cognition predicts suicidal behavior. *Psychological Science, 21*(4), 511–517. https://doi.org/10.1177/0956797610364762

Oldershaw, A., Richards, C., Simic, M., & Schmidt, U. (2008). Parents' perspectives on adolescent self-harm: Qualitative study. *British Journal of Psychiatry, 193*(2), 140–144. https://doi.org/10.1192/bjp.bp.107.045930

Powers, J. T., Brausch, A. M., & Muehlenkamp, J. J. (2021). Explicit and implicit measurement of nonsuicidal self-injury in the prediction of concurrent and prospective self-injury. *Suicide and Life-Threatening Behavior, 51*(4), 665–672. https://doi.org/10.1111/sltb.12734

Röhner, J., Schröder-Abé, M., & Schütz, A. (2013). What do fakers actually do to fake the IAT? An investigation of faking strategies under different faking conditions. *Journal of Research in Personality, 47*(4), 330–338. https://doi.org/10.1016/j.jrp.2013.02.009

Rosenrot, S. A., & Lewis, S. P. (2020). Barriers and responses to the disclosure of non-suicidal self-injury: A thematic analysis. *Counselling Psychology Quarterly, 33*(2), 121–141. https://doi.org/10.1080/09515070.2018.1489220

Saunders, K. E. A., Hawton, K., Fortune, S., & Farrell, S. (2012). Attitudes and knowledge of clinical staff regarding people who self-harm: A systematic review. *Journal of Affective Disorders, 139*(3), 205–216. https://doi.org/10.1016/j.jad.2011.08.024

Simone, A. C., & Hamza, C. A. (2020). Examining the disclosure of nonsuicidal self-injury to informal and formal sources: A review of the literature. *Clinical Psychology Review, 82*, Article 101907. https://doi.org/10.1016/j.cpr.2020.101907

Sohn, M. N., McMorris, C. A., Bray, S., & McGirr, A. (2021). The Death-Implicit Association Test and suicide attempts: A systematic review and meta-analysis of discriminative and prospective utility. *Psychological Medicine, 51*(11), 1789–1798. https://doi.org/10.1017/S0033291721002117

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & the Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *Journal of the American Medical Association, 282*(18), 1737–1744. https://doi.org/10.1001/jama.282.18.1737

Steffens, M. C. (2004). Is the Implicit Association Test immune to faking? *Experimental Psychology, 51*(3), 165–179. https://doi.org/10.1027/1618-3169.51.3.165

Stieger, S., Göritz, A. S., Hergovich, A., & Voracek, M. (2011). Intentional faking of the Single Category Implicit Association Test and the Implicit Association Test. *Psychological Reports, 109*(1), 219–230. https://doi.org/10.2466/03.09.22.28.PR0.109.4.219-230

Swannell, S. V., Martin, G. E., Page, A., Hasking, P., & St John, N. J. (2014). Prevalence of nonsuicidal self-injury in nonclinical samples: Systematic review, meta-analysis and meta-regression. *Suicide and Life-Threatening Behavior, 44*(3), 273–303. https://doi.org/10.1111/sltb.12070

Tuisku, V., Kiviruusu, O., Pelkonen, M., Karlsson, L., Strandholm, T., & Marttunen, M. (2014). Depressed adolescents as young adults – Predictors of suicide attempt and non-suicidal self-injury during an 8-year follow-up. *Journal of Affective Disorders, 152–154*, 313–319. https://doi.org/10.1016/j.jad.2013.09.031

Turner, B. J., Layden, B. K., Butler, S. M., & Chapman, A. L. (2013). How often, or how many ways: Clarifying the relationship between non-suicidal self-injury and suicidality. *Archives of Suicide Research, 17*(4), 397–415. https://doi.org/10.1080/13811118.2013.802660

Whitlock, J., Muehlenkamp, J., Eckenrode, J., Purington, A., Baral Abrams, G., Barreira, P., & Kress, V. (2013). Nonsuicidal self-injury as a gateway to suicide in young adults. *The Journal of Adolescent Health, 52*(4), 486–492. https://doi.org/10.1016/j.jadohealth.2012.09.010

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

No Cognitive Scarring Over 20 Years in Recurrent Depression: Bidirectionality of Cognitive Functioning and Depression

Joost Gülpen[1,2], Amanda M. Legemaat[1,3], Eva A. M. van Dis[1], Gert J. Geurtsen[4,5], Ellie M. Wekking[6], Huibert Burger[7], Damiaan A. J. P. Denys[1,4], Claudi L. H. Bockting[1,2]

[1] *Department of Psychiatry, Amsterdam UMC, Amsterdam, The Netherlands.* **[2]** *Amsterdam Public Health Research Institute, Amsterdam, The Netherlands.* **[3]** *GGZ InGeest Specialized Mental Health Care, Amsterdam, The Netherlands.* **[4]** *Amsterdam Neuroscience Research Institute, Amsterdam, The Netherlands.* **[5]** *Department of Medical Psychology, Amsterdam UMC, Amsterdam, The Netherlands.* **[6]** *Parnassia Group, Den Haag, The Netherlands.* **[7]** *Department of Primary and Long-term Care Medicine, University Medical Center Groningen, Groningen, The Netherlands.*

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e20989, https://doi.org/10.32872/cpe.20989

Received: 2025-11-21 • **Accepted:** 2026-04-20 • **Published (VoR):** 2026-08-31

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

Corresponding Author: Joost Gülpen, Amsterdam UMC, University of Amsterdam, Department of Psychiatry, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands. Telephone number: +31 20 891 3600. E-mail: j.gulpen@amsterdamumc.nl

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

Abstract

Background: Cognitive deficits are prevalent in major depressive disorder (MDD), span multiple cognitive domains, and often persist during remission. However, long-term associations between cognitive functioning and depression remain unclear. This study explores bidirectional associations using 20-year longitudinal data, providing follow-up twice as long as previously reported.

Method: Individuals in remission of recurrent MDD underwent neuropsychological assessments at baseline (n = 137, 21-65 years) and 20-year follow-up (n = 33), covering multiple domains. Psychiatric assessments occurred at eight time points. Cox proportional hazard models and linear mixed-effects models tested bidirectional associations between cognitive functioning and depression. Group-level cognitive change was tested using paired sample t-tests and Bayesian analyses, and individual-level change using reliable change indices (RCI's).

Results: No cognitive domains significantly predicted time-to-depressive-relapse, the number of depressive episodes, or total time depressed over 20 years. Similarly, the number of episodes, total time depressed, and changes in depression severity did not significantly predict cognitive changes. At follow-up, we observed declines in speed of information processing ($p < .001$, $d = -0.73$, RCI: 32.1% deteriorated) and speed of memory processing ($p = .015$, $d = -0.49$, RCI: 39.3% deteriorated). Learning and memory improved ($p < .001$, $d = 1.19$, RCI: 28.6% improved). Working memory and executive functioning remained stable. Findings were not explained by selective drop-out.
Conclusions: We found only minor long-term cognitive change, with no clear worsening or evidence of scarring and state effects. Cognitive functioning remained relatively stable, and we observed no link to depression trajectories over 20 years, providing a more optimistic outlook for depressed patients.

Keywords

cognitive functioning, major depressive disorder, relapse, scarring, cognitive impairments, bidirectionality

Highlights

- Cognitive functioning did not predict depression trajectories over 20-years follow-up, twice as long as prior work.
- Similarly, changes in cognitive functioning were not predicted by depression trajectories.
- No state- or scarring effects were found, and cognitive functioning remained relatively stable.

Major depressive disorder (MDD) is one of the most prevalent mental disorders (Ferrari et al., 2013), characterized by high rates of recurrence and treatment resistance (Hardeveld et al., 2010; McIntyre et al., 2023), making it a leading cause of disease burden. Traditionally, MDD treatment and research have primarily focused on affective symptoms, often overlooking cognitive functioning. However, cognitive deficits are not merely a clinical epiphenomenon of an affective disorder but constitute an important symptom and central feature. Between 24% and 79% of adults with MDD are affected by cognitive deficits, with the prevalence varying based on clinical state and definitions used (Douglas et al., 2018; Gualtieri, 2008; Liu et al., 2023; Stainton et al., 2023; Tran et al., 2021). Moreover, cognitive deficits contribute to the societal costs and disability associated with depression, as they are linked to a reduced quality of life and diminished functioning (Cambridge et al., 2018; Evans et al., 2013; Lam et al., 2014).

Interestingly, objective measures of cognitive performance poorly align with depressed patients' subjective experiences (Serra-Blasco et al., 2019). This underscores the importance of focusing on objectively measured cognitive functioning when examining longitudinal associations with depression. Nonetheless, objectively measured cognitive deficits in MDD span multiple cognitive domains, including memory and learning, at-

tention, psychomotor speed, processing speed, and executive functioning (Douglas & Porter, 2009; Hammar & Årdal, 2009; Roca et al., 2015; Rock et al., 2014; Snyder, 2013; Wagner et al., 2012). These deficits are evident across different stages of depression: during first acute episodes (Ahern & Semkovska, 2017; Lee et al., 2012; Varghese et al., 2022), recurrent episodes (Kriesche et al., 2023), and, remarkably, in remission (Kriesche et al., 2023; Semkovska et al., 2019). In remission, both objective and subjective cognitive deficits often persist as common residual symptoms (Conradi et al., 2011; Minor et al., 2005; Semkovska et al., 2019), indicating that clinical remission often does not equate cognitive remission (Bortolato et al., 2016).

Several hypotheses have been put forward to explain the nature of cognitive deficits in depression (Allott et al., 2016; Ang et al., 2020; Douglas & Porter, 2009; Hammar et al., 2022). First, it is hypothesized that deficits accumulate progressively with repeated depressive episodes and persist during remission (*scar hypothesis*: depression → cognitive functioning). Alternatively, cognitive deficits may fluctuate with the severity of depressive symptoms and improve when the depression resolves (*state hypothesis*: cognitive functioning ↔ depression). Lastly, they could stem from a pre-existing vulnerability that remains constant throughout the course of depression, regardless of clinical state (*trait hypothesis*: cognitive functioning → depression). Importantly, these three hypothesized mechanisms are not mutually exclusive, may co-occur in time, or vary across individuals, cognitive domains, and the phases of depression (Ahern et al., 2025; Allott et al., 2016; Ang et al., 2020; Semkovska et al., 2019). There is at least partial and mixed empirical support for these hypotheses, based on, for example, family studies (Allott et al., 2016), premorbid studies (Allott et al., 2016), cross-sectional comparisons between MDD patients and healthy controls (Ahern et al., 2025; Rock et al., 2014; Snyder, 2013), and studies comparing individuals with first-episode depression to those with recurrent depression (Guan et al., 2025; Talarowska et al., 2015). However, as many of these findings rely on between-group comparisons, longitudinal designs with repeated measures within the same individuals remain essential. Longitudinal approaches help minimize potential confounding that may arise when comparing different patient groups at different stages of illness.

Although cognitive deficits are a well-established symptom over the course of depression, reverse causality is also possible after initial MDD onset, with cognitive deficits contributing to the chronic nature of depression. For instance, cognitive deficits predict poorer treatment response (Groves et al., 2018) and an increased risk of relapse (Alexopoulos et al., 2000; Maeshima et al., 2016; Schmid & Hammar, 2013), although other studies found no such associations (Majer et al., 2004; Reppermund et al., 2009; Wekking et al., 2012). The heightened relapse risk may be explained by cognitive deficits facilitating cognitive biases, dysfunctional beliefs, and rumination (Ahern et al., 2019). Recent population-based longitudinal studies, primarily in older adults, have reported bidirectional associations between cognitive functioning and depressive symptoms (Fong

et al., 2025; Ma et al., 2025; Yin et al., 2024). However, evidence from clinician-verified clinical samples using extensive neuropsychological testing and focusing on bidirectionality across the recurrent depressive course rather than late-life cognitive decline remains limited. Understanding the nature of cognitive deficits and their consequences is important, given their associated burden, and can additionally help identify targets for treatment and prevention.

This study aims to disentangle long-term bidirectional associations between cognitive functioning and depression over 20 years, increasing our understanding of their development over time following MDD onset. Longitudinal data from the DELTA-study (Legemaat et al., 2023), including recurrently depressed individuals in remission, provides a unique opportunity to address these questions. To our knowledge, this study provides the longest follow-up on cognitive functioning in depression, extending over twice as long as previously reported (Hammar & Årdal, 2012; Sarapas et al., 2012). In summary, we aim to: (a) explore to what extent cognitive functioning prospectively predicts depression trajectories over 20 years; (b) explore to what extent depression trajectories predict changes in cognitive functioning over 20 years; and (c) explore changes in cognitive functioning over 20 years. Based on previous research indicating bidirectionality, we hypothesize that cognitive functioning not only prospectively predicts the course of depression, but conversely, is also influenced by the course of depression.

Method

Design and Sample

This study is a follow-up of the DELTA-study (Legemaat et al., 2023; Wekking et al., 2012), a randomized controlled trial (RCT) examining effects of preventive cognitive therapy (PCT) in remitted MDD with follow-ups after 3 months and 1, 2, 3, 5.5, 10, and 20 years. Participants were recruited in 2000, and the follow-up ended in 2022. At entry, participants were (a) at least 18 years old and (b) in remission from MDD for at least 10 weeks and at most 2 years. They experienced (c) two or more previous MDD episodes in the last 5 years and (d) had 17-item Hamilton Depression Rating Scale (HDRS-17) scores less than 10 (Hamilton, 1960). Exclusion criteria were organic brain damage, substance abuse, current or previous psychotic disorders, current or previous (hypo)mania, predominant anxiety disorder, recent electroconvulsive therapy, and recent or current psychotherapy.

Procedure

Participants were screened using the Structured Clinical Interview for DSM–IV (SCID-I) and the HDRS-17. Participants who met the inclusion criteria were randomized to PCT (8 weekly group sessions) or treatment as usual (TAU; comprising standard care) and

were invited to attend blinded follow-up assessments. Neuropsychological assessments were conducted at baseline and at the 20-year follow-up. Psychiatric assessments were conducted at baseline and 3-month, 1-, 2-, 3-, 5.5-, 10-, and 20-year follow-up. Procedures were approved by the institutional review board. Participants gave informed consent prior to participation and at follow-ups.

Outcomes

Neuropsychological Assessment

Neuropsychological assessments consisted of Dutch versions of internationally adopted tests (see Supplementary Table 1), conducted at baseline and 20-year follow-up only. The Stroop Color and Word Test (SCWT) was used to assess the speed of information processing (trials I and II) and executive functioning (trial III; Lezak et al., 2004). To assess the speed of memory processing, we used four subtasks of the Memory Comparison Task (MCT; Lezak et al., 2004). Additionally, both the backward and forward subtests of the Digit Span Test (Lezak et al., 2004) were used to measure working memory. To assess learning and memory, the Stories Recall subtest of the Rivermead Behavioral Memory Test (RBMT) and the Dutch California Verbal Learning Test (CVLT) were used. The Stories Recall subtest of the RBMT is a measure of immediate recall, absolute delayed recall, and relative delayed recall (Lezak et al., 2004; Wilson et al., 1991). The CVLT measures components of immediate recall and delayed recall (Mulder et al., 1996). Additionally, at 20-year follow-up, participants self-reported cognitive problems using the WHODAS 2.0 (Üstün, 2010).

Raw test scores were converted to standardized scores (*T*-scores: $M = 50$, $SD = 10$) based on Dutch normative data, controlling for sex, age, and educational level. To minimize the number of tests, we calculated composite *T*-scores for five cognitive domains, at both time points by averaging *T*-scores of (sub)tests: (a) speed of information processing (SCWT trails I and II); (b) speed of memory processing (MCT); (c) working memory (Digit Span forwards and backwards); (d) memory and learning (CVLT and RBMT subtests); and (e) executive functioning (SCWT trial III). Participants were classed as impaired (*T*-scores ≤ 36 considered as clinical range) or non-impaired on all five domains separately (Hermans et al., 2024). Performance validity was tested at 20-year follow-up, using the Test of Memory Malingering (TOMM; scores ≥ 45 considered valid) and Amsterdam Short Term Memory Test (ASTM; scores ≥ 85 considered valid; Schmand et al., 1999; Tombaugh, 1997). Total scores and classifications based on manual-recommended cut-off scores were used for the performance validity tests.

Depressive Relapse, Episodes and Severity

At baseline, 3-month, 1-, 2-, 3-, 5.5-, and 10-year follow-up, current and past depressive episodes were measured using the SCID-I and at 20 years using the Structured Clinical Interview for DSM-5 (SCID-5; First et al., 2016). Relapse was defined as meeting MDD

criteria again 2 months after remission, while meeting criteria within 2 months was considered a continuation of the previous episode. These assessments also provided data on time-to-relapse, total number of episodes, and total time depressed (in days). The term *depression trajectories* is used as a collective term for these course indicators (e.g., time-to-relapse, number of episodes) and does not represent a separate outcome. Interrater agreement of assessments up to 10-year follow-up was high (K = .94). The 20-year assessments were conducted by one clinician.

We used the HDRS-17 (Hamilton, 1960) to assess depressive symptom severity as an outcome across follow-up psychiatric assessments. This semi-structured interview covers affective, behavioral, and somatic symptoms over the past week. Seventeen items are scored on 3-point or 5-point scales, with total scores ranging from 0 to 52. The HDRS-17 has good psychometric properties (Trajković et al., 2011) and showed acceptable average internal consistency across the DELTA-study (α = .73). The Beck Depression Inventory (BDI), a self-report measure of depressive symptom severity (Beck et al., 1961), was used to describe baseline clinical characteristics.

Statistical Analyses

Analyses were conducted in R (Version 4.3.2), using a two-sided p-value of < .05 to determine statistical significance. Due to the small sample size and resulting low statistical power, we considered all our analyses exploratory. Consequently, no corrections for multiple testing were applied, and no power analysis was performed. Non-response analyses were conducted to compare participants who completed both neuropsychological assessments with those who dropped out, on clinical, cognitive, and demographic baseline values. Comparisons used χ^2-tests for categorical variables and independent-sample t-tests for continuous variables. No imputation was conducted. Handling of missing data differed per analysis. Cross-sectional analyses were conducted using available cases. For the mixed-effects model, all available repeated measurements were included. To validate similarity between treatment groups (PCT vs. TAU), we assessed differences on the five cognitive domains at follow-up using analysis of covariance, adjusting for baseline cognition and, in a second analysis, adding the number of previous episodes (< 4 vs. ≥ 4) and the interaction with treatment, in line with previous work (Legemaat et al., 2023). If no significant or substantial effects were found, participants would be analyzed as one group in subsequent analyses. Spearman's rank-order correlations were calculated between self-reported cognitive problems and objective cognitive functioning (T-scores). Sensitivity analyses were conducted, excluding participants who failed performance validity tests. Across analyses, models were estimated without additional covariates given the exploratory nature of the study and modest sample size.

Cognitive Functioning as Predictor of Depression Trajectories (Aim 1)

Baseline cognitive functioning as a predictor of depression trajectories over 20 years was assessed for each cognitive domain (Aim 1). Effects of cognitive functioning on time-to-relapse were tested using separate Cox proportional hazards models to appropriately handle time-to-event data with censoring and to allow the estimation of relapse risk over time. Additionally, impaired versus non-impaired groups with each domain separately were compared using Kaplan-Meier survival curves and log-rank tests. A generalized linear mixed model with zero-inflation negative binomial distributions was used to predict the number of depressive episodes measured at seven follow-ups over 20 years, using the *glmmTMB* package, to account for excess zeros. Effect estimates were expressed as incidence rate ratios (*IRRs*). Linear mixed-effects models were used to predict total time depressed measured at seven follow-ups over 20 years, utilizing the *lme4* package. In general, linear mixed-effects models were used as they account for repeated continuous outcomes and can handle unbalanced, partially missing data. Baseline cognitive domain score and time (follow-up wave) were included as fixed effects. Random intercepts for participants were specified to account for within-subject dependency across repeated assessments.

Depression Trajectories as Predictors of Cognitive Changes (Aim 2)

We used linear mixed-effects models to explore whether long-term depression trajectories were predictive of cognitive change over time (Aim 2). For this, we predicted changes in cognitive domains separately, between baseline and follow-up, by the number of depressive episodes, total time depressed, and changes in HDRS-17 scores over 20 years (operationalized as the difference between baseline and 20-year follow-up scores). In each model, time (baseline vs. follow-up), the respective depression trajectory variable, and their interaction were included as fixed effects. A random intercept for participant was specified to account for within-subject dependency across repeated assessments. The interaction term tested whether long-term depression burden was associated with differential cognitive change between baseline and follow-up.

Changes in Cognitive Functioning (Aim 3)

Group-level changes, defined as mean within-subject changes, in cognitive functioning (Aim 3) were examined using paired sample *t*-tests for individual neuropsychological test scores and cognitive composite domain scores. With these analyses, we aimed to test whether on average cognitive functioning changed over time. Given the small sample size for these analyses, additional Bayesian analyses were conducted as this can quantify support in favor of the null hypothesis, and not only against it, using *BayesFactor*. To assess individual-level changes (Aim 3), reliable change indices (RCI) were calculated and plotted, using the *JTRCI* package. RCI values greater than 1.96 or less than -1.96 were seen as reliable improvements or deteriorations; an RCI between these values indicated

no change. Individual profiles were examined to determine significant declines into clinically impaired ranges (*T*-scores ≤ 36). In a sensitivity analysis we excluded participants failing performance validity tests (TOMM and/or ASTM).

Results

Sample Characteristics

Overall, 312 individuals were assessed for eligibility, and 172 were included, of whom 137 individuals aged 21-65 years (*Mdn* = 46) completed the baseline neuropsychological assessment (Figure 1). At 20-year follow-up, 38 individuals were evaluated, with 33 (19.2%) undergoing the neuropsychological assessment (PCT: *n* = 18; TAU: *n* = 15). In total, 28 individuals (16.3%) completed neuropsychological assessments at both baseline and 20-year follow-up. In a drop-out analysis, all clinical, cognitive, and demographic baseline values were highly similar between those assessed at both time points and those only seen at baseline (Supplementary Table 2), except for BDI scores (p = .022). Baseline characteristics for participants with cognitive data at baseline (n = 137) and follow-up (n = 33) are presented in Table 1. Individuals participating in the 20-year follow-up had a baseline average age of 44.8 years (*SD* = 7.55), with 78.8% female, 27.3% highly educated, and four previous depressive episodes (*IQR* = 3–6). The mean interval between neuropsychological assessments was 256 months (range = 251-259 months). At 20-year follow-up, no participants reported acquired brain injuries or neurodegenerative diseases. Over follow-up, one individual received electroconvulsive therapy; none underwent deep-brain stimulation. Performance validity tests at 20-year follow-up found no failures on the TOMM and seven failures (21.2%) on the ASTM. Treatment allocation had no significant effects on any cognitive domains, regardless of adding the number of previous episodes (< 4 vs. ≥ 4), and the interaction with treatment was not substantial (all p-values > .05; Supplementary Table 3). On average, participants with cognitive data at both time points (n = 28) had 3.29 (*SD* = 2.68) depressive episodes and were depressed for 93.81 weeks (*SD* = 157.77) over 20 years.

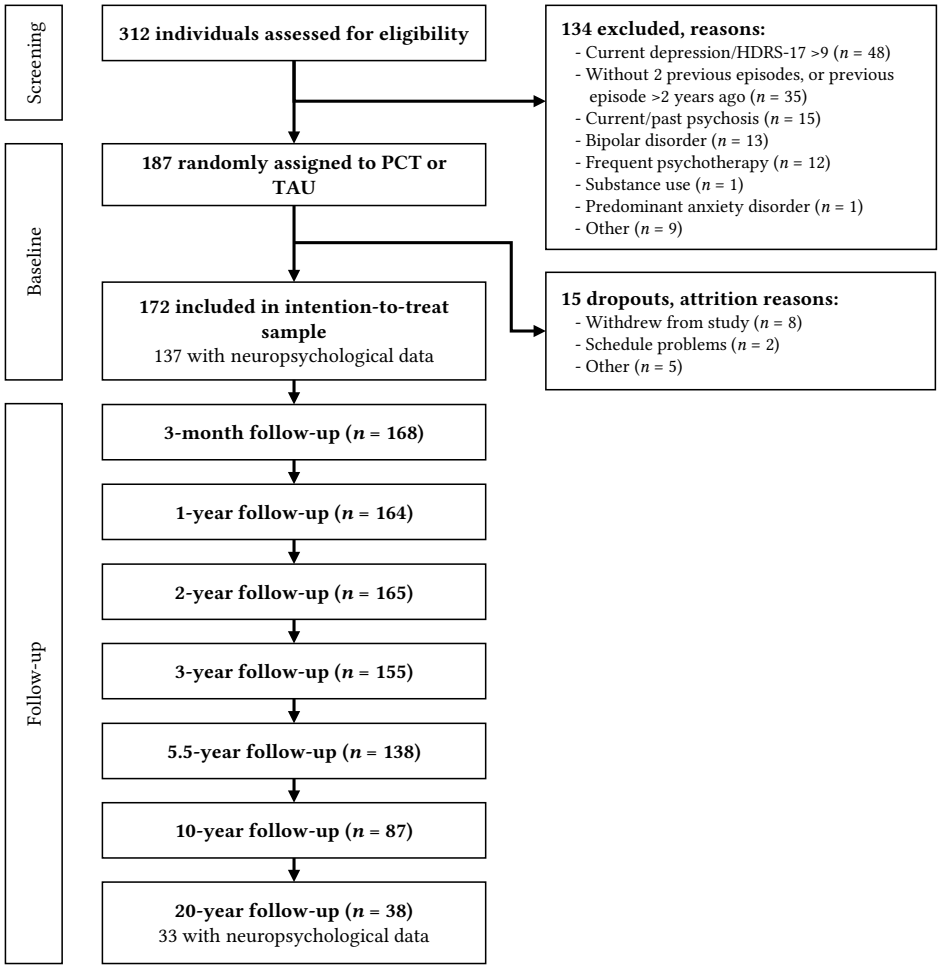
Cognitive Functioning as Predictor of Depression Trajectories

No significant effects of time-to-relapse (n = 137) were found for speed of information processing (*HR* = 0.99, 95% CI [0.97, 1.01], p = .343), speed of memory processing (*HR* = 1.01, 95% CI [0.99, 1.02], p = .546), working memory (*HR* = 0.99, 95% CI [0.96, 1.01], p = .271), learning and memory (*HR* = 1.00, 95% CI [0.98, 1.01], p = .637), and executive functioning (*HR* = 1.01, 95% CI [0.99, 1.03], p = .511). We only observed survival differences in the speed of information processing when comparing clinically impaired versus non-impaired groups ($\chi^2(1, N = 137) = 6.1$, p = .013), with a significantly higher probability of

relapse amongst impaired participants (Supplementary Figure 1, for Kaplan-Meier curves per domain).

Figure 1

Flow Diagram of Participants From Baseline to 20-Year Follow-Up



Note. HDRS-17 = 17-item Hamilton Depression Rating Scale; PCT = preventive cognitive therapy; TAU = treatment-as-usual.

Table 1

Baseline Demographic and Clinical Characteristic of Included Participants

Variable	Baseline sample (*n* = 137)	20-year follow-up sample (*n* = 33)
Demographic characteristics		
Age, years (*M, SD*)	44.91 (9.38)	44.79 (7.55)
Gender, Female (*n*, %)	102/137 (74.45)	26/33 (78.79)
Married/cohabiting (*n*, %)	85/137 (62.04)	17/33 (51.52)
Education level (*n*, %)		
High	54/137 (39.42)	9/33 (27.27)
Middle	41/137 (29.93)	12/33 (36.36)
Low	42/137 (30.66)	12/33 (36.36)
Education level, years (*M, SD*)	14.02 (2.01)	13.90 (2.05)
Clinical characteristics		
Age of MDD onset, years (*M, SD*)	28.18 (12.69)	29.06 (12.72)
Months in remission (*M, SD*)	8.90 (6.44)	9.45 (7.22)
Number of previous episodes (*Mdn, IQR*)	4 (3-6)	4 (3-6)
>3 previous episodes (*n*, %)	114/137 (83.21)	26/33 (78.79)
Duration previous episode, months (*M, SD*)	7.20 (8.73)	7.46 (6.45)
Depressive symptoms, HDRS-17 (*M, SD*)	3.76 (3.03)	3.94 (3.23)
Depressive symptoms, BDI (*M, SD*)	11.92 (7.49)	9.96 (6.88)
Antidepressant usage (*n*, %)	72/137 (52.55)	12/33 (36.36)
Benzodiazepine usage (*n*, %)	17/137 (12.41)	7/33 (21.21)
Clinical cognitive impairment on ≥1 domain [a] (*n*, %)	77/137 (56.20)	19/33 (57.58)
Clinical cognitive impairments [a] (*n*, %)		
Speed of information processing	22/137 (16.06)	9/33 (27.27)
Speed of memory processing	35/137 (25.55)	14/33 (42.42)
Working memory	6/137 (4.38)	0/33 (0)
Memory and learning	39/137 (28.47)	2/33 (6.06)
Executive functioning	16/137 (11.68)	4/33 (12.12)

Note. Data are mean (*SD*), median (*IQR*), or *n*/*N* (%). BDI = Beck Depression Inventory; HDRS-17 = 17-item Hamilton Depression Rating Scale; *IQR* = interquartile range; MDD = major depressive disorder; *SD* = standard deviation.

[a] *T*-scores ≤ 36 are considered as cognitive impairments within clinical ranges.

Speed of information processing (*IRR* = 0.99, 95% CI [0.98, 1.01], *p* = .369), speed of memory processing (*IRR* = 1.00, 95% CI [0.99, 1.02], *p* = .482), working memory (*IRR* = 0.99, 95% CI [0.97, 1.01], *p* = .336), learning and memory (*IRR* = 0.99, 95% CI [0.98, 1.01], *p* = .299) and executive functioning (*IRR* = 1.01, 95% CI [1.00, 1.02], *p* = .121) did not significantly predict the number of depressive episodes over 20 years (*n* = 137,

observations = 866). Similarly, none of the cognitive domains significantly predicted the total time depressed over 20 years (n = 137, observations = 866): speed of information processing (β = -0.07, 95% CI [-0.14, 0.00], p = .065, partial R^2 = 0.01), speed of memory processing (β = 0.00, 95% CI [-0.07, 0.07], p = .985, partial R^2 = 0.00), working memory (β = -0.03, 95% CI [-0.10, 0.04], p = .438, partial R^2 = 0.00), learning and memory (β = 0.04, 95% CI [-0.03, 0.11], p = .239, partial R^2 = 0.00), and executive functioning (β = 0.04, 95% CI [-0.03, 0.11], p = .247, partial R^2 = 0.00).

Depression Trajectories as Predictors of Cognitive Changes

We explored whether the course of depression predicts changes in cognitive functioning over 20 years using linear mixed-effects models. However, the number of depressive episodes, total time depressed, and changes in HDRS-17 scores did not significantly predict changes in cognitive functioning over 20 years (all ps > .05; Supplementary Table 4). Sensitivity analyses excluding participants failing performance validity based on the ASTM did not change the results (all ps > .05; results not shown).

Changes in Cognitive Functioning

Standardized T-scores and raw scores on neuropsychological tests and domains across time points are presented in Table 2. The majority of patients had T-scores in clinical ranges on one or more cognitive domains (56% at baseline; 58% at follow-up; see Table 1). There were significant declines in T-scores for speed of information processing (mean difference = -5.04; 95% CI [-7.73, -2.36], p < .001, d = -0.73) and speed of memory processing (mean difference = -5.08; 95% CI [-9.09, -1.08], p = .015, d = -0.49). A significant improvement was found for learning and memory (mean difference = 10.88; 95% CI [7.32, 14.44], p < .001, d = 1.19). No changes were found for working memory (mean difference = -0.02; 95% CI [-2.84, 2.87], p = .990, d = 0.00) and executive functioning (mean difference = -0.28; 95% CI [-3.41, 2.85], p = .855, d = 0.04). After conducting sensitivity analyses excluding seven participants failing performance validity tests, changes in speed of memory processing were nonsignificant (p = .097). Other effects remained unchanged (results not shown). Bayesian paired samples t-tests showed very strong evidence for changes in speed of information processing and learning and memory, and moderate evidence for speed of memory processing (Supplementary Table 5). For working memory and executive functioning, moderate evidence supported the null hypothesis (i.e., no change over time).

Clinical Psychology in Europe
2026, Vol. 8(3), Article e20989
https://doi.org/10.32872/cpe.20989

Table 2

Mean (SD) of Standardized and Raw Scores on Neuropsychological Tests at Baseline and 20-Year Follow-up Across Domains for Participants With Complete Neuropsychological Data at Both Follow-ups (n = 28).

	Test score					Reliable Change Index		
	Baseline		20-year follow-up					
Neuropsychological test and domain	*T*-scores *M* (*SD*)	Raw scores *M* (*SD*)	*T*-scores *M* (*SD*)	Raw scores *M* (*SD*)	*p*[a]	Reliably deteriorated *n* (%)	No change *n* (%)	Reliably improved *n* (%)
Speed of Information Processing	45.65 (6.81)		40.61 (7.57)		< .001*	9 (32.14)	18 (64.29)	1 (3.57)
Stroop Color and Word Test – Trial I	45.51 (8.20)	44.75 (8.31)	39.36 (8.39)	51.21 (8.61)	< .001*	8 (28.57)	20 (71.43)	0 (0)
Stroop Color and Word Test – Trial II	45.79 (7.83)	58.07 (8.36)	41.86 (9.30)	65.25 (12.26)	.019*	8 (28.57)	18 (64.29)	2 (7.14)
Speed of Memory Processing	42.78 (11.22)		37.69 (13.51)		.015*	11 (39.29)	13 (46.43)	4 (14.29)
Memory Comparison Task – 1 Letter	38.47 (12.91)	27.96 (5.47)	38.01 (13.01)	42.93 (10.38)	.837	4 (14.29)	19 (67.86)	5 (17.86)
Memory Comparison Task – 2 Letters	43.43 (12.75)	39.43 (9.17)	45.48 (12.93)	52.25 (14.72)	.378	2 (7.14)	19 (67.86)	7 (25.00)
Memory Comparison Task – 3 Letters	44.34 (10.05)	50.14 (9.97)	33.74 (18.06)	75.82 (22.90)	< .001*	15 (53.57)	12 (42.86)	1 (3.57)
Memory Comparison Task – 4 Letters	44.86 (14.61)	68.82 (24.21)	33.54 (17.73)	92.21 (26.35)	.004*	10 (35.71)	17 (60.71)	1 (3.57)
Working Memory	49.25 (7.00)		49.23 (7.61)		.990	2 (7.14)	24 (85.71)	2 (7.14)
Digit Span Test – Digits forward	48.21 (8.62)	5.39 (0.92)	47.54 (9.81)	5.43 (1.07)	.751	3 (10.71)	23 (82.14)	2 (7.14)
Digit Span Test – Digits backward	50.29 (7.30)	4.46 (1.07)	50.93 (8.66)	4.71 (1.21)	.679	1 (3.57)	26 (92.86)	1 (3.57)
Learning and Memory	45.14 (10.34)		56.01 (12.96)		< .001*	0 (0)	20 (71.43)	8 (28.57)
Dutch California Verbal Learning Test								
Immediate recall	49.29 (28.14)	54.86 (11.56)	63.21 (25.68)	55.07 (11.02)	.001*	0 (0)	28 (100)	0 (0)
Delayed recall	48.93 (22.50)	12.43 (2.94)	60.71 (28.27)	13.50 (2.17)	.034*	3 (10.71)	18 (64.29)	7 (25.00)
Rivermead Behavioral Memory Test								
Immediate recall – subtest Stories	41.73 (11.00)	17.39 (5.75)	49.82 (11.72)	18.86 (7.10)	.003*	0 (0)	21 (75.00)	7 (25.00)
Absolute delayed recall – subtest Stories	40.54 (10.86)	14.41 (5.64)	52.50 (12.50)	16.25 (7.07)	< .001*	0 (0)	13 (46.43)	15 (53.57)
Relative delayed recall – subtest Stories	45.19 (13.78)	[b]	53.82 (13.57)	[b]	.020*	1 (3.57)	21 (75.00)	6 (21.43)
Executive Functioning	47.71 (10.21)		47.43 (9.65)		.855	3 (10.71)	22 (78.57)	3 (10.71)
Stroop Color and Word Test – Trial III	47.71 (10.21)	94.32 (22.43)	47.43 (9.65)	104.93 (25.24)	.855	3 (10.71)	22 (78.57)	3 (10.71)

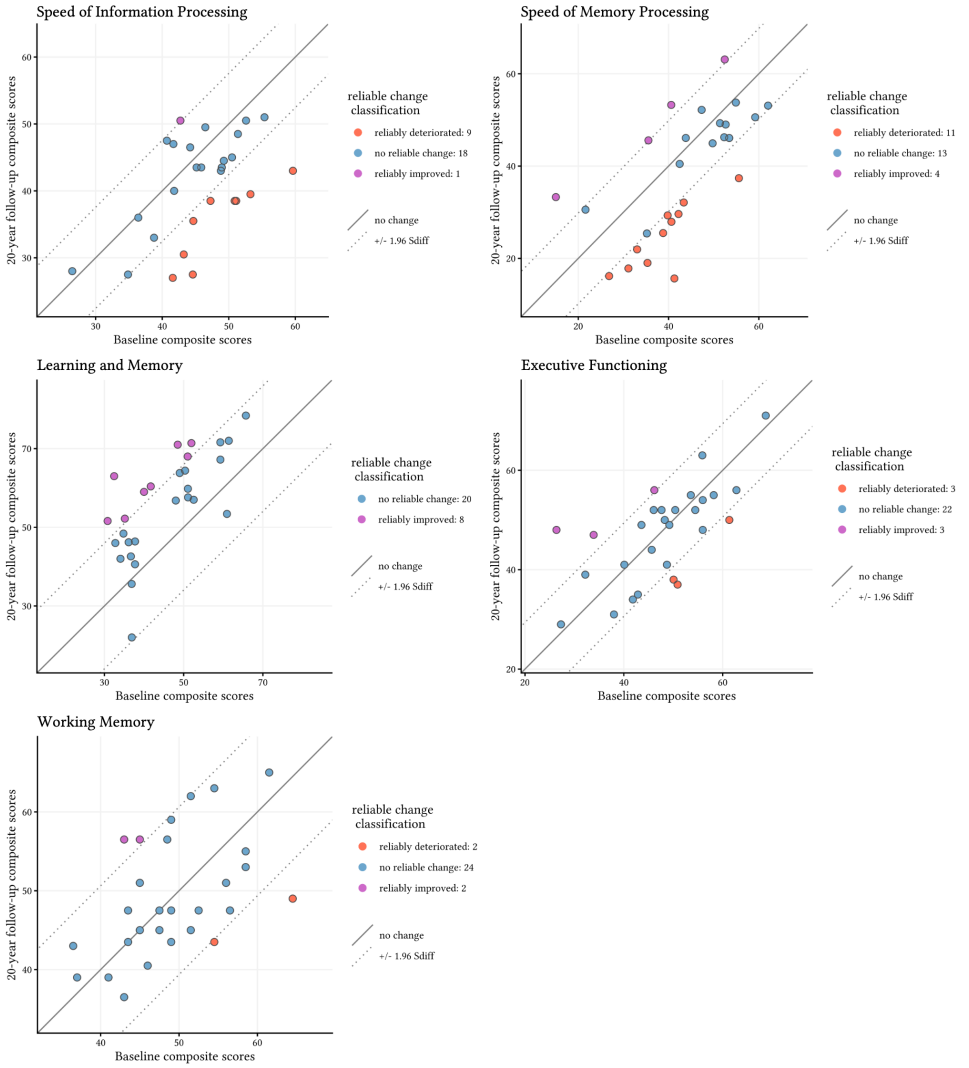
Note. *M* = mean; *SD* = standard deviation.

[a]Results from paired-sample *t*-tests, comparing baseline and 20-year follow-up *T*-scores. [b]No raw scores available as this *T*-score is calculated using the immediate and absolute delayed recall scores. *$p \leq .05$.

Moreover, using RCIs, we observed no specific pattern of change in cognitive functioning across specific tests and domains (Table 2 and Figure 2), although most declines were seen for speed of information processing (32.1% reliably deteriorated) and speed of memory processing (39.3% reliably deteriorated).

Figure 2

Reliable Change Index Plots of Composite T-Scores for the Five Cognitive Domains, Visualizing Individual-Level Change Between Baseline and 20-Year Follow-Up



Regarding the working memory, a majority of 85.7% remained unchanged. In contrast, no individuals deteriorated in learning and memory, and 28.6% significantly improved. On executive functioning, 78.6% showed no reliable change. Individual *T*-scores across cognitive profiles showed that most individuals who experienced declines in speed of information processing over 20 years also showed declines in speed of memory processing, and vice versa (Supplementary Table 6). In total, four of the 28 individuals showed a significant decline and fell into clinical ranges on speed of information processing, and 10 of the 28 individuals in speed of memory processing, while no such reliable changes were seen for other domains (Supplementary Table 6). Excluding participants failing performance validity tests did not considerably alter patterns of reliable change (Supplementary Table 7). Only five people did not relapse over 20 years, however, we did not observe different patterns among these individuals (Supplementary Table 8). At follow-up, 85.7% of individuals (24/28) self-reported cognitive problems: 42.9% minor (12/28), 28.6% moderate (8/28), and 14.3% severe (4/28). Only 2 of the 4 individuals reporting severe problems had *T*-scores in clinical ranges. Self-reported cognitive problems did not correlate with objectively measured cognitive domains at 20-year follow-up (all *p*-values > .05; Supplementary Figure 2). Boxplots showing baseline and 20-year follow-up cognitive functioning by domain are shown in Supplementary Figure 3.

Discussion

In this longitudinal study, we explored cognitive change and bidirectional associations between cognitive functioning and depression over 20 years – a follow-up twice as long as currently reported – amongst individuals in remission from recurrent MDD. Importantly, in this at-risk group, cognitive functioning was not associated with prospectively assessed time-to-relapse nor with the number and duration of depressive episodes over 20 years. Likewise, while we could not completely rule this out due to the small sample, we found no indication of either scarring- or state-effects over 20 years. Our results, while explorative, challenge the assumption that cognitive deterioration due to scarring is an inevitable consequence of recurrent depressive episodes. Additionally, we found only limited cognitive change over 20 years and more importantly no marked deterioration. Therefore, our findings suggest that cognitive functioning in depression may remain relatively stable over extended periods, irrespective of the number of recurrences. Taken together, our observations may suggest that long-term cognitive changes and their bidirectional associations with depression may not be as pronounced as previously hypothesized, providing a more optimistic long-term outlook for depressed patients.

Furthermore, cognitive functioning did not predict the depression course over 20 years. None of the five domains, as measured using demographically corrected *T*-scores, were associated with time-to-relapse or the number and duration of episodes over 20 years. However, participants showing clinical impairments in speed of information

processing were more likely to relapse compared to those without such impairments. This finding is noteworthy, as processing speed is one of the domains most affected in remitted MDD (Semkovska et al., 2019) and especially pronounced in those with recurrent episodes (Varghese et al., 2022). Cognitive functioning as a risk factor for relapse has been examined in only a limited number of studies, with mixed results. While some studies have found no significant associations (Majer et al., 2004; Reppermund et al., 2009; Wekking et al., 2012), others, in contrast to our findings, found support for memory impairments (Maeshima et al., 2016), and executive functioning impairments as predictors of relapse (Alexopoulos et al., 2000; Schmid & Hammar, 2013). However, previous studies differ considerably in terms of their samples (i.e. first episode vs. recurrent MDD, unipolar vs. bipolar, geriatric sample), follow-up periods, and cognitive domains assessed. Our study adds to this growing body of research by including a comprehensive neuropsychological and psychiatric assessment, with the longest follow-up and largest sample currently available. Taken together, cognitive functioning was not found to be a clinical marker of chronicity. Cognitive impairment itself, however, was found to persist over 20 years in this sample.

Surprisingly, we did not find any evidence for scarring- or state-like effects, as cognitive change was unrelated to changes in depression severity or the number and duration of depressive episodes over 20 years. In contrast, several meta-analytic studies in acute and remitted MDD reported worsening cognitive functioning or moderation of cognitive change by depressive symptom severity and the number of episodes (Ahern et al., 2025; Ahern & Semkovska, 2017; Semkovska et al., 2019). Nevertheless, our findings do in part align with current perspectives suggesting that such scar- and state effects are rarely universal but tend to be domain-specific and influenced by methodological factors, e.g. cognitive tests used, indices of depression assessed (e.g., symptom status vs. number of episodes), clinical status (remitted, first-episode vs. recurrently depressed) and study design. Population-based studies on this topic are very limited, with mixed results, and are mostly cross-sectional, reliant on brief cognitive assessments, or primarily focused on symptom levels and not clinician-verified diagnoses (de Nooij et al., 2020; Formánek et al., 2020; Zhu et al., 2022). However, a study using UK Biobank data, with validated diagnostic interviews and comprehensive cognitive assessments, similarly found no evidence that previous episodes or symptoms were associated with cognitive impairments in individuals with a previous MDD diagnosis (de Nooij et al., 2020). There are several possible explanations that could explain our observations. First, our remitted, recurrent sample reported a median of four previous episodes at baseline. Therefore, some potential scarring effects might have already occurred before baseline neuropsychological assessment. Indeed, there is evidence that cognitive impairment takes places early in the disease-course (Varghese et al., 2022). Second, our analyses suffered from limited statistical power as neuropsychological data at both time-points was available for only 28 participants.

Overall, while the majority of patients had *T*-scores in clinical ranges on one or more cognitive domains, we found no clear group-level worsening or improvement of cognitive functioning. Whereas both speed of information processing and speed of memory processing decreased over 20 years, changes are not clinically meaningful. Although not fully explained, we did observe a large improvement in learning and memory. Some small-to-moderate pro-cognitive change has been observed in this domain (Ahern et al., 2025), but it remains unclear why this finding differentiates from other domains. However, learning and memory had the highest number of participants with clinical impairments (28.5%) and had – apart from speed of memory processing – the lowest average *T*-score. This improvement may thus, at least in part, be attributed to a regression-to-the-mean effect. Interestingly, recurrently depressed individuals perform significantly worse on learning and memory tests and show the largest impairments on this domain, compared to first-episode individuals, indicative of a progressive decline (Varghese et al., 2022). Working memory and executive functioning were not subjected to group-level change and showed limited individual-level change using RCIs. RCIs did show that a subset of participants experienced reliable deteriorations, particularly in speed of information processing and speed of memory processing, underscoring inter-individual variability and, at least in part, contradicting group-level findings. It is possible that our observations are influenced by confounding factors associated with increased age given the long follow-up which may not have been fully captured, such as long-term treatment, medication or substance usage, hospitalization, or chronic somatic conditions. However, *T*-scores were used based on normative data across several age categories to correct for age-related cognitive decline, and we could not attribute cognitive changes to medication usage, time in remission or chronic somatic conditions. Nevertheless, our findings are somewhat reassuring as patients may worry about progressively worsening and long-lasting cognitive deficits following depression.

This study has several strengths. Composite scores from an elaborate neuropsychological test-battery were used to increase reliability compared to single tests, with corrections for age, sex and education level. Our design also allowed us to assess cognitive stability over 20 years in relation to depression – the longest follow-up currently available – and to examine longitudinal associations over eight measurements. Additionally, results could not be explained by differential drop-out. Lastly, we did not only examine cognitive change on a group-level, but also on an individual-level using RCIs.

Nevertheless, several limitations should be noted. First, we included individuals in remission from recurrent MDD, with at least 2 previous episodes, limiting generalizability. Relatedly, assessments of cognitive changes earlier in the illness course are more interpretable with regards to scar, trait and state hypotheses (Allott et al., 2016). Second, we lacked a healthy control group, which is necessary to adequately control for practice effects and relate changes in performance to clinical variables. We were thus unable to compare developmental trajectories of this sample with prior depressive episodes

to change over time in a non-MDD sample. Third, given the exploratory character of our study due to insufficient sample size in our analyses, no corrections for multiple testing were conducted. As such, our results have limited precision and should be considered exploratory, requiring cautious interpretation. Fourth, cognitive functioning was assessed at only two time points over 20 years, limiting our ability to properly examine bidirectional associations between depression and cognition. As the interplay between these variables is likely dynamic, future studies should include more frequent repeated neuropsychological and psychiatric assessments over this timeframe, using different longitudinal approaches (e.g., cross-lagged or latent growth models) to better capture bidirectional effects. Fifth, this longitudinal follow-up study was part of an RCT, and some participants were actively seeking treatment or took part in a systematic intervention, although we found no effects of treatment allocation on cognitive functioning at 20 years. Lastly, we were unable to account for premorbid cognitive functioning or IQ in the current study, using an idiographic or intraindividual approach (Lezak et al., 2004). While we used a normative approach to compare individuals' performance compared to demographically matched peers, this may not capture relative deficits compared to premorbid levels. Thus, some baseline impairments may be the result of low premorbid functioning. Conversely some impairments may be missed in individuals with higher premorbid cognitive functioning, while they may have experienced deteriorations from their premorbid levels.

In conclusion, in this exploratory study we found no clear evidence supporting cognitive functioning as a predictor of depression chronicity and recurrence over 20 years, apart from clinical impairments in the speed of information processing. Additionally, within this recurrent, remitted sample, we did not observe consistent scarring- or state-like effects on cognitive functioning. Participants in remission of recurrent MDD did not experience substantial cognitive decline over 20 years. However, the majority of patients showed persistent cognitive impairment over this time period. Although our findings are exploratory, and consequently the precision of our estimates is limited, they still suggest that the long-term cognitive impact of depression and their bidirectional associations may be less pronounced than previously thought. Large-scale longitudinal studies including individuals with clinician-verified MDD and extended follow-ups are warranted. Therefore, our results should be interpreted with caution. Nonetheless, given the considerable follow-up of our study, our findings offer a valuable perspective into cognitive functioning in depression and may provide some reassurance to patients suffering from recurrent MDD.

Funding: This study was supported by grants from the Health Research Development Counsel, Department of Prevention Program (ZonMw), and the National Foundation for Mental Health, The Netherlands (Fonds Psychische Gezondheid). These sources had no role in the study design; data collection, preparation or analysis; writing of the manuscript; decision to submit. However, this specific paper 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: Claudi L. H. Bockting is Editor-in-Chief of *Clinical Psychology in Europe* but played no editorial role for this particular article or intervened in any form in the peer review process.

Ethics Statement: The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. This long term follow-up study design has been approved by the Medical Ethics Committee of the Amsterdam (CCMO: NL-OMON52298). Participants of the follow up study were properly instructed and they gave informed consent.

Social Media Accounts: *Joost Gülpen:* LinkedIn

Preregistration: Before start of the follow-up study, it was registered at the Overview of the Medical Research in the Netherlands (OMON), the former Dutch Trial Register, under ID (NL-OMON52298) (Gülpen et al., 2021S).

Reporting Guidelines: Clinical trial, intervention study follows the CONSORT statement.

Data Availability: The corresponding author JG had full access to the study data and materials. The data, materials and codes used for this article may be shared on reasonable request to the corresponding author.

Supplementary Materials

The Supplementary Materials contain the following items:

- **The preregistration for the study** (Gülpen et al., 2021S)
- **Supplementary tables and figures** (Gülpen et al., 2026S):
  - *Supplementary Table 1.* Concise description of the neuropsychological tests and the tasks administered.
  - *Supplementary Table 2.* Results from non-response analysis examining selective drop-out based on demographic, cognitive and clinical characteristics.
  - *Supplementary Table 3.* Results of ANCOVAs assessing differences between PCT and TAU on five cognitive domains at 20 year follow-up.
  - *Supplementary Table 4.* Outcomes from linear mixed-effects model of depression variables as predictors of cognitive functioning change over 20 year follow-up.
  - *Supplementary Table 5.* Results of Bayesian analyses of differences in five cognitive composite domains between baseline and 20 year follow-up (n = 28).

- *Supplementary Table 6.* Individual reliable change indices and T-scores, of all participants with complete neuropsychological data at baseline and 20 year follow-up across five composite domains.
- *Supplementary Table 7.* Results of reliable change index analyses after sensitivity analyses excluding participants failing performance validity.
- *Supplementary Table 8.* Results of reliable change index analyses for participants not experiencing a relapse over 20 year follow-up.
- *Supplementary Figure 1.* Kaplan-Meier survival curves of time-to-relapse comparing participants with clinical impairments on cognitive domains to those non-impaired.
- *Supplementary Figure 2.* Correlation matrix showing associations at 20-year follow-up between self-reported cognitive problems and objective cognitive functioning.
- *Supplementary Figure 3.* Boxplots showing baseline and 20-year follow-up cognitive functioning by domain.

Index of Supplementary Materials

Gülpen, J., Legemaat, A. M., van Dis, E. A. M., Geurtsen, G. J., Wekking, E. M., Burger, H., Denys, D. A. J. P., & Bockting, C. L. H. (2021S). *Relapse, cognitive, and daily functioning in recurrently depressed individuals: DELTA study 20-year follow-up* [Preregistration]. OMON. https://onderzoekmetmensen.nl/nl/trial/52298

Gülpen, J., Legemaat, A. M., van Dis, E. A. M., Geurtsen, G. J., Wekking, E. M., Burger, H., Denys, D. A. J. P., & Bockting, C. L. H. (2026S). *Supplementary materials to "No cognitive scarring over 20 years in recurrent depression: Bidirectionality of cognitive functioning and depression"* [Supplementary tables and figures]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.22291

References

Ahern, E., Bockting, C. L. H., & Semkovska, M. (2019). A hot-cold cognitive model of depression: Integrating the neuropsychological approach into the cognitive theory framework. *Clinical Psychology in Europe, 1*(3), Article e34396. https://doi.org/10.32872/cpe.v1i3.34396

Ahern, E., & Semkovska, M. (2017). Cognitive functioning in the first-episode of major depressive disorder: A systematic review and meta-analysis. *Neuropsychology, 31*(1), 52–72. https://doi.org/10.1037/neu0000319

Ahern, E., White, J., & Slattery, E. (2025). Change in cognitive function over the course of major depressive disorder: A systematic review and meta-analysis. *Neuropsychology Review, 35*(1), 1–34. https://doi.org/10.1007/s11065-023-09629-9

Alexopoulos, G. S., Meyers, B. S., Young, R. C., Kalayam, B., Kakuma, T., Gabrielle, M., Sirey, J. A., & Hull, J. (2000). Executive dysfunction and long-term outcomes of geriatric depression. *Archives of General Psychiatry, 57*(3), 285–290. https://doi.org/10.1001/archpsyc.57.3.285

Allott, K., Fisher, C. A., Amminger, G. P., Goodall, J., & Hetrick, S. (2016). Characterizing neurocognitive impairment in young people with major depression: State, trait, or scar? *Brain and Behavior, 6*(10), Article e00527. https://doi.org/10.1002/brb3.527

Ang, Y.-S., Frontero, N., Belleau, E., & Pizzagalli, D. A. (2020). Disentangling vulnerability, state and trait features of neurocognitive impairments in depression. *Brain, 143*(12), 3865–3877. https://doi.org/10.1093/brain/awaa314

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*(6), 561–571. https://doi.org/10.1001/archpsyc.1961.01710120031004

Bortolato, B., Miskowiak, K. W., Köhler, C. A., Maes, M., Fernandes, B. S., Berk, M., & Carvalho, A. F. (2016). Cognitive remission: A novel objective for the treatment of major depression? *BMC Medicine, 14*, Article 9. https://doi.org/10.1186/s12916-016-0560-3

Cambridge, O. R., Knight, M. J., Mills, N., & Baune, B. T. (2018). The clinical relationship between cognitive impairment and psychosocial functioning in major depressive disorder: A systematic review. *Psychiatry Research, 269*, 157–171. https://doi.org/10.1016/j.psychres.2018.08.033

Conradi, H. J., Ormel, J., & de Jonge, P. (2011). Presence of individual (residual) symptoms during depressive episodes and periods of remission: A 3-year prospective study. *Psychological Medicine, 41*(6), 1165–1174. https://doi.org/10.1017/S0033291710001911

de Nooij, L., Harris, M. A., Adams, M. J., Clarke, T.-K., Shen, X., Cox, S. R., McIntosh, A. M., & Whalley, H. C. (2020). Cognitive functioning and lifetime major depressive disorder in UK Biobank. *European Psychiatry, 63*(1), Article e28. https://doi.org/10.1192/j.eurpsy.2020.24

Douglas, K. M., Gallagher, P., Robinson, L. J., Carter, J. D., McIntosh, V. V., Frampton, C. M., Watson, S., Young, A. H., Ferrier, I. N., & Porter, R. J. (2018). Prevalence of cognitive impairment in major depression and bipolar disorder. *Bipolar Disorders, 20*(3), 260–274. https://doi.org/10.1111/bdi.12602

Douglas, K. M., & Porter, R. J. (2009). Longitudinal assessment of neuropsychological function in major depression. *Australian and New Zealand Journal of Psychiatry, 43*(12), 1105–1117. https://doi.org/10.3109/00048670903279887

Evans, V. C., Chan, S. S., Iverson, G. L., Bond, D. J., Yatham, L. N., & Lam, R. W. (2013). Systematic review of neurocognition and occupational functioning in major depressive disorder. *Neuropsychiatry, 3*(1), 97–105. https://doi.org/10.2217/npy.13.3

Ferrari, A. J., Somerville, A. J., Baxter, A. J., Norman, R., Patten, S. B., Vos, T., & Whiteford, H. (2013). Global variation in the prevalence and incidence of major depressive disorder: A systematic review of the epidemiological literature. *Psychological Medicine, 43*(3), 471–481. https://doi.org/10.1017/S0033291712001511

First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2016). *User's guide for the SCID-5-CV Structured Clinical Interview for DSM-5® disorders: Clinical version.* American Psychiatric Publishing. https://psycnet.apa.org/record/2016-15667-000

Fong, T. C. T., Chan, R. T. H., Wen, M., & Yip, P. S. F. (2025). Bidirectional relationships between cognition and depressive symptoms and effects of sociodemographic and lifestyle factors:

Random-intercept, cross-lagged panel model. *BJPsych Open, 11*(4), Article e144. https://doi.org/10.1192/bjo.2025.10074

Formánek, T., Csajbók, Z., Wolfová, K., Kučera, M., Tom, S., Aarsland, D., & Cermakova, P. (2020). Trajectories of depressive symptoms and associated patterns of cognitive decline. *Scientific Reports, 10*(1), Article 20888. https://doi.org/10.1038/s41598-020-77866-6

Groves, S. J., Douglas, K. M., & Porter, R. J. (2018). A systematic review of cognitive predictors of treatment outcome in major depression. *Frontiers in Psychiatry, 9*, Article 382. https://doi.org/10.3389/fpsyt.2018.00382

Gualtieri, C. T. (2008). The frequency of cognitive impairment in patients with anxiety, depression, and bipolar disorder: An unaccounted source of variance in clinical trials. *Journal of Clinical Psychiatry, 69*(7), 1122–1130. https://doi.org/10.4088/JCP.v69n0712

Guan, L., Li, Y., Kong, H., Fang, J., Yu, J., Wang, T., Zhu, J., & Zhu, D. (2025). Differences in cognitive deficits and brain functional impairments between patients with first-episode and recurrent depression. *BMC Psychiatry, 25*(1), Article 434. https://doi.org/10.1186/s12888-025-06758-8

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry, 23*(1), 56–62. https://doi.org/10.1136/jnnp.23.1.56

Hammar, Å., & Årdal, G. (2009). Cognitive functioning in major depression—A summary. *Frontiers in Human Neuroscience, 3*, Article 728. https://doi.org/10.3389/neuro.09.026.2009

Hammar, Å., & Årdal, G. (2012). Effortful information processing in patients with major depression—A 10-year follow-up study. *Psychiatry Research, 198*(3), 420–423. https://doi.org/10.1016/j.psychres.2011.11.020

Hammar, Å., Ronold, E. H., & Rekkedal, G. Å. (2022). Cognitive impairment and neurocognitive profiles in major depression—A clinical perspective. *Frontiers in Psychiatry, 13*, Article 764374. https://doi.org/10.3389/fpsyt.2022.764374

Hardeveld, F., Spijker, J., De Graaf, R., Nolen, W. A., & Beekman, A. T. F. (2010). Prevalence and predictors of recurrence of major depressive disorder in the adult population. *Acta Psychiatrica Scandinavica, 122*(3), 184–191. https://doi.org/10.1111/j.1600-0447.2009.01519.x

Hermans, M. E., Geurtsen, G. J., Hollak, C. E. M., & Bosch, A. M. (2024). Neuropsychological stability in classical galactosemia: A pilot study in 10 adult patients. *JIMD Reports, 65*(2), 110–115. https://doi.org/10.1002/jmd2.12410

Kriesche, D., Woll, C. F. J., Tschentscher, N., Engel, R. R., & Karch, S. (2023). Neurocognitive deficits in depression: A systematic review of cognitive impairment in the acute and remitted state. *European Archives of Psychiatry and Clinical Neuroscience, 273*(5), 1105–1128. https://doi.org/10.1007/s00406-022-01479-5

Lam, R. W., Kennedy, S. H., McIntyre, R. S., & Khullar, A. (2014). Cognitive dysfunction in major depressive disorder: Effects on psychosocial functioning and implications for treatment. *Canadian Journal of Psychiatry / Revue Canadienne de Psychiatrie, 59*(12), 649–654. https://doi.org/10.1177/070674371405901206

Lee, R. S. C., Hermens, D. F., Porter, M. A., & Redoblado-Hodge, M. A. (2012). A meta-analysis of cognitive deficits in first-episode major depressive disorder. *Journal of Affective Disorders, 140*(2), 113–124. https://doi.org/10.1016/j.jad.2011.10.023

Legemaat, A. M., Burger, H., Geurtsen, G. J., Brouwer, M., Spinhoven, P., Denys, D., & Bockting, C. L. (2023). Effects up to 20-year follow-up of preventive cognitive therapy in adults remitted from recurrent depression: The DELTA study. *Psychotherapy and Psychosomatics, 92*(1), 55–64. https://doi.org/10.1159/000527906

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). Oxford University Press.

Liu, J., Chen, Y., Xie, X., Liu, B., Ju, Y., Wang, M., Dong, Q., Lu, X., Sun, J., Zhang, L., Guo, H., Zhao, F., Li, W., Zhang, L., Li, Z., Liao, M., Li, L., & Zhang, Y. (2023). The percentage of cognitive impairment in patients with major depressive disorder over the course of the depression: A longitudinal study. *Journal of Affective Disorders, 329*, 511–518. https://doi.org/10.1016/j.jad.2023.02.133

Ma, X., Zhou, X., & Qiu, Y. (2025). Bidirectional relationships between cognitive decline and depression: A study of middle-aged and older adults using cross-lagged panel network analysis. *Journal of Affective Disorders, 389*, Article 119741. https://doi.org/10.1016/j.jad.2025.119741

Maeshima, H., Baba, H., Satomura, E., Shimano, T., Inoue, M., Ishijima, S., Suzuki, T., & Arai, H. (2016). Residual memory impairment in remitted depression may be a predictive factor for recurrence. *The Journal of Clinical Psychiatry, 77*(2), 247–251. https://doi.org/10.4088/JCP.14m09694

Majer, M., Ising, M., Künzel, H., Binder, E. B., Holsboer, F., Modell, S., & Zihl, J. (2004). Impaired divided attention predicts delayed response and risk to relapse in subjects with depressive disorders. *Psychological Medicine, 34*(8), 1453–1463. https://doi.org/10.1017/S0033291704002697

McIntyre, R. S., Alsuwaidan, M., Baune, B. T., Berk, M., Demyttenaere, K., Goldberg, J. F., Gorwood, P., Ho, R., Kasper, S., Kennedy, S. H., Ly-Uson, J., Mansur, R. B., McAllister-Williams, R. H., Murrough, J. W., Nemeroff, C. B., Nierenberg, A. A., Rosenblat, J. D., Sanacora, G., Schatzberg, A. F., . . . Maj, M. (2023). Treatment-resistant depression: Definition, prevalence, detection, management, and investigational interventions. *World Psychiatry, 22*(3), 394–412. https://doi.org/10.1002/wps.21120

Minor, K. L., Champion, J. E., & Gotlib, I. H. (2005). Stability of DSM-IV criterion symptoms for major depressive disorder. *Journal of Psychiatric Research, 39*(4), 415–420. https://doi.org/10.1016/j.jpsychires.2004.11.001

Mulder, J. L., Dekker, M. J. A., & Dekker, P. H. (1996). *Verbale leer en geheugen test.* Swets Test Services (STS).

Reppermund, S., Ising, M., Lucae, S., & Zihl, J. (2009). Cognitive impairment in unipolar depression is persistent and non-specific: Further evidence for the final common pathway disorder hypothesis. *Psychological Medicine, 39*(4), 603–614. https://doi.org/10.1017/S003329170800411X

Roca, M., Vives, M., López-Navarro, E., García-Campayo, J., & Gili, M. (2015). Cognitive impairments and depression: A critical review. *Actas Españolas de Psiquiatría, 43*(5), 187–193.

https://actaspsiquiatria.es/index.php/actas/article/view/608https://doi.org/10.1016/j.euroneuro.2015.07.020

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine, 44*(10), 2029–2040. https://doi.org/10.1017/S0033291713002535

Sarapas, C., Shankman, S. A., Harrow, M., & Goldberg, J. F. (2012). Parsing trait and state effects of depression severity on neurocognition: Evidence from a 26-year longitudinal study. *Journal of Abnormal Psychology, 121*(4), 830–837. https://doi.org/10.1037/a0028141

Schmand, B. A., De Sterke, S., & Lindeboom, J. (1999). *Amsterdamse Korte Termijn Geheugentest.* Amsterdam: Pearson Assessment and Information.

Schmid, M. T., & Hammar, Å. (2013). A follow-up study of first episode major depressive disorder. Impairment in inhibition and semantic fluency—Potential predictors for relapse? *Frontiers in Psychology, 4*, Article 633. https://doi.org/10.3389/fpsyg.2013.00633

Semkovska, M., Quinlivan, L., O'Grady, T., Johnson, R., Collins, A., O'Connor, J., Knittle, H., Ahern, E., & Gload, T. (2019). Cognitive function following a major depressive episode: A systematic review and meta-analysis. *The Lancet Psychiatry, 6*(10), 851–861. https://doi.org/10.1016/S2215-0366(19)30291-3

Serra-Blasco, M., Torres, I. J., Vicent-Gil, M., Goldberg, X., Navarra-Ventura, G., Aguilar, E., Via, E., Portella, M. J., Figuereo, I., Palao, D., Lam, R. W., & Cardoner, N. (2019). Discrepancy between objective and subjective cognition in major depressive disorder. *European Neuropsychopharmacology, 29*(1), 46–56. https://doi.org/10.1016/j.euroneuro.2018.11.1104

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. *Psychological Bulletin, 139*(1), 81–132. https://doi.org/10.1037/a0028727

Stainton, A., Chisholm, K., Griffiths, S. L., Kambeitz-Ilankovic, L., Wenzel, J., Bonivento, C., Brambilla, P., Iqbal, M., Lichtenstein, T. K., Rosen, M., Antonucci, L. A., Maggioni, E., Kambeitz, J., Borgwardt, S., Riecher-Rössler, A., Andreou, C., Schmidt, A., Schultze-Lutter, F., Meisenzahl, E., . . . PRONIA Consortium (2023). Prevalence of cognitive impairments and strengths in the early course of psychosis and depression. *Psychological Medicine, 53*(13), 5945–5957. https://doi.org/10.1017/S0033291723001770

Talarowska, M., Zajączkowska, M., & Gałecki, P. (2015). Cognitive functions in first-episode depression and recurrent depressive disorder. *Psychiatria Danubina, 27*(1), 38–43.

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment, 9*(3), 260–268. https://doi.org/10.1037/1040-3590.9.3.260

Trajković, G., Starčević, V., Latas, M., Leštarević, M., Ille, T., Bukumirić, Z., & Marinković, J. (2011). Reliability of the Hamilton Rating Scale for Depression: A meta-analysis over a period of 49 years. *Psychiatry Research, 189*(1), 1–9. https://doi.org/10.1016/j.psychres.2010.12.007

Tran, T., Milanovic, M., Holshausen, K., & Bowie, C. R. (2021). What is normal cognition in depression? Prevalence and functional correlates of normative versus idiographic cognitive impairment. *Neuropsychology, 35*(1), 33–41. https://doi.org/10.1037/neu0000717

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

Varghese, S., Frey, B. N., Schneider, M. A., Kapczinski, F., & de Azevedo Cardoso, T. (2022). Functional and cognitive impairment in the first episode of depression: A systematic review. *Acta Psychiatrica Scandinavica, 145*(2), 156–185. https://doi.org/10.1111/acps.13385

Wagner, S., Doering, B., Helmreich, I., Lieb, K., & Tadić, A. (2012). A meta-analysis of executive dysfunctions in unipolar major depressive disorder without psychotic symptoms and their changes during antidepressant treatment. *Acta Psychiatrica Scandinavica, 125*(4), 281–292. https://doi.org/10.1111/j.1600-0447.2011.01762.x

Wekking, E. M., Bockting, C. L. H., Koeter, M. W. J., & Schene, A. H. (2012). Cognitive functioning in euthymic recurrently depressed patients: Relationship with future relapses and prior course of disease. *Journal of Affective Disorders, 141*(2–3), 300–307. https://doi.org/10.1016/j.jad.2012.03.034

Wilson, B., Cockburn, J., & Baddeley, A. D. (1991). *The Rivermead behavioural memory test*. Thames Valley Test Company.

Yin, J., John, A., & Cadar, D. (2024). Bidirectional associations of depressive symptoms and cognitive function over time. *JAMA Network Open, 7*(6), Article e2416305. https://doi.org/10.1001/jamanetworkopen.2024.16305

Zhu, Y., Li, C., Xie, W., Zhong, B., Wu, Y., & Blumenthal, J. A. (2022). Trajectories of depressive symptoms and subsequent cognitive decline in older adults: A pooled analysis of two longitudinal cohorts. *Age and Ageing, 51*(1), Article afab191. https://doi.org/10.1093/ageing/afab191

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

From Goals to Gains: Therapy Goal Selection, Attainment, and Their Association With Psychotherapy Outcomes

Flavio Iovoli[1], Saskia Rieger[1], Juan Martín Gómez Penedo[1,2,3], Jessica Fritz[1], Martin Cordes[1], Ruben Lauterbach[4], Henning Schöttke[1], Julian A. Rubel[1]

[1] *Department of Psychology, Osnabrueck University, Osnabrueck, Germany.* **[2]** *Department of Psychology, University of Kassel, Kassel, Germany.* **[3]** *Department of Psychology, Vrije Universiteit Brussel, Brussels, Belgium.* **[4]** *Department of Medicine, Justus-Liebig-University Giessen, Giessen, Germany.*

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e19567, https://doi.org/10.32872/cpe.19567

Received: 2025-08-29 • **Accepted:** 2026-04-16 • **Published (VoR):** 2026-08-31

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

Corresponding Author: Flavio Iovoli, Department of Psychology, Psychotherapy Research and Clinical Psychology, Osnabrueck University, Lise-Meitner-Street 3, D-49076 Osnabrueck, Germany. Phone: +49 17622853184. E-mail: flavio.iovoli@gmail.com

Abstract

Background: Therapy goals are central to psychotherapeutic interventions. While symptom-specific goals are commonly emphasized, patients often set diverse goals that may not align with standardized symptom measures. This study examines how patient characteristics are associated with the selection of specific therapy goal content categories and whether attainment in these domains shows associations with different psychotherapy outcomes.

Method: Therapy goals from 645 patients undergoing cognitive-behavioral therapy in an outpatient clinic were categorized using the Bern Inventory of Treatment Goals and analyzed. Baseline predictors of therapy goal selection and diversity were examined using multilevel logistic and ordinal regression models. The association between goal selection and their goal attainment with residualized post-treatment scores in symptoms, global functioning, and end-of-treatment patient satisfaction were assessed using Bayesian multilevel regression models.

Results: Baseline patient characteristics were associated with the selection of specific therapy goal categories, alongside substantial therapist-level effects in several categories. Goal-category attainment showed differential associations with psychotherapy outcomes, with the strongest and most consistent effects observed for symptom-related outcomes.

Conclusion: These findings indicate that the attainment of specific therapy goal categories is meaningfully reflected in standardized outcome measures, but in a category-sensitive rather than

uniform way. Combining therapy goal-based and standardized assessments may improve the alignment between patient priorities and outcome evaluation in routine psychotherapy.

Keywords

therapy goals, goal selection, goal attainment, psychotherapy outcomes, individualized outcomes

Highlights

- Patient characteristics are associated with therapy goal selection and diversity.
- Goal attainment predicts global outcomes and represents valuable individualized indicators.
- Therapist effects are present in therapy goal selection and treatment results.

Therapy goals are a core element of many psychotherapeutic treatments (Grawe, 2004; Grosse Holtforth et al., 2004). They are individualized and an intentional expression of the changes a patient hopes to achieve through psychotherapy (Brunstein et al., 1998; Grosse Holtforth & Grawe, 2000). Naturally, patients seek psychotherapy for a variety of reasons (Michalak et al., 2004). Some may want to gain control over their thoughts, while others aim to overcome specific fears or enhance their overall sense of well-being. Some individuals seek therapy to learn relaxation techniques or improve self-control. Others might be searching for deeper meaning and purpose in their lives. Given these diverse objectives, therapy goals shape the therapeutic process (Grosse Holtforth et al., 2004; Michalak & Schulte, 2002; Silberschatz, 2017). In particular, collaboratively defined therapy goals are associated with patient engagement and motivation (Grosse Holtforth & Michalak, 2012; Michalak & Grosse Holtforth, 2006) and represent a core component of the therapeutic alliance (Bordin, 1979; Tryon, 2018; Tryon & Winograd, 2011).

However, as illustrated above, patients' individual therapy goals might not always align with disorder-specific symptoms (Chevance et al., 2020; Lloyd et al., 2019; Ramnerö & Jansson, 2016b). Despite this, psychotherapy effectiveness is often assessed through changes in standardized symptom checklists (for example, the SCL-90-R; Derogatis, 1977), which may not fully represent the patient's focus in psychotherapy. This approach risks overlooking the very important aspect of psychotherapy success: achieving the goals that patients sought psychotherapy for and consider most relevant. In fact, a meta-analysis by Lindhiem and colleagues (2016) demonstrated that psychotherapy shows significantly greater effectiveness when outcomes are measured through idiographic, goal-specific outcomes rather than using standardized symptom checklists. Although symptom-specific measures are convenient and provide opportunities for nomothetic comparison (Lambert et al., 2005), they are less consistently implemented in routine care than in research contexts (Ionita & Fitzpatrick, 2014; Puschner et al., 2015). Therapists sometimes perceive these measures as time-consuming, potentially burdensome for some patients, and ultimately not useful in practice (Gleacher et al., 2016; Hatfield & Ogles,

2007). In line with this, Jensen-Doss and colleagues (2018) found that therapists in practice tend to use and prefer idiographic outcome measures over standardized symptom-specific ones. Together, these highlight the importance of systematically examining patient-defined therapy goals and their relationship to established outcome indicators.

The most widely used assessment method for capturing idiographic therapy goals and their progress (Lloyd et al., 2019; Sales, Ashworth, et al., 2023) is the Goal Attainment Scaling (GAS; Kiresuk et al., 1994; Kiresuk & Sherman, 1968). In GAS, patients and therapists jointly define up to five individualized goals and specify criteria for different levels of attainment, which are then evaluated during or at the end of treatment (Kiresuk et al., 1994). While GAS provides a structured way to assess goal attainment, its highly individualized format limits comparability across patients. To address this limitation, the Bern Inventory of Treatment Goals (BIT) offers an empirically derived framework categorizing therapy goals by content (Grosse Holtforth, 2001; Grosse Holtforth & Grawe, 2002). Based on large samples of patient-reported goals, the BIT distinguishes five overarching themes (Grosse Holtforth & Grawe, 2002): (1) symptom- and problem-related (e.g., to overcome ruminative thoughts or learn to manage anxiety), (2) interpersonal-related (e.g., to improve relationships with specific individuals in personal or professional life), (3) personal growth-related (e.g., to develop greater self-confidence and self-assurance), (4) existential-related (e.g., to gain greater clarity about who I am, what I can do, and what I want), and (5) well-being-related (e.g., to make my leisure time more active; Grosse Holtforth, 2001) goals. Combining structured goal attainment ratings with standardized goal content categories enables systematic and comparable analyses of which types of goals patients pursue and how their attainment relates to broader treatment outcomes (Michalak et al., 2007). However, research in this area remains limited and fragmented, often focusing either on goal selection or on goal attainment alone, while little is known about how specific categories of therapy goals relate to treatment outcomes, despite the clear relevance of such knowledge for clinical practice.

Prior research on therapy goal selection suggests that patients often choose symptom- and problem-focused goals (Baur et al., 2024; Berking et al., 2005; Ramnerö & Jansson, 2016a; Sawrikar et al., 2023; Schöttke et al., 2014), whereas other work suggests a stronger emphasis on interpersonal or personal growth–related goals (Grosse Holtforth & Grawe, 2002). These differences likely also reflect contextual factors such as therapeutic orientation and treatment setting, which are associated with goal formulation (Dirmaier et al., 2006; Schöttke et al., 2014). Some variability appears to be associated with diagnostic differences, with anxious patients tending to select more symptom-focused goals and depressed patients showing more thematically diverse goal profiles, indicating differences not only in goal type but also in overall goal diversity (Grosse Holtforth et al., 2009). In this context, goal diversity refers to the extent to which a patient's therapy goals span multiple thematic goal categories rather than concentrating on a single domain. Another study found that patient age was associated with differ-

ences in goal focus, with older patients emphasizing well-being and younger patients emphasizing personal growth (Sittler et al., 2022). However, the association between other baseline patient characteristics and both therapy goal selection and goal diversity remains largely unexplored. Clarifying these associations could support more tailored treatment planning and improve the alignment between therapeutic interventions and patient priorities.

While understanding which goals are selected seems important, an equally central question concerns whether and how the attainment of different therapy goals relates to treatment outcomes. Whether therapy goals are attained is clinically meaningful because goal attainment reflects tangible therapeutic progress from the patient's perspective. Prior research indicates that attainment rates vary across goal categories, yet findings are mixed and show no consistent pattern, with some studies reporting higher attainment for well-being (Berking et al., 2005) or existential goals (Baur et al., 2024) and others for symptom-focused goals (Trachsel et al., 2008). These descriptive differences indicate that goal types are not equally easy to attain, yet they do not clarify whether attaining certain kinds of goals is more strongly linked to therapy outcome. Research connecting goal attainment to outcomes has mostly relied on overall attainment across all goals, which is associated with improvements in quality of life and reductions in anxiety and depression, and shows moderate correlations with standardized symptom change (Baur et al., 2024; Ramnerö & Jansson, 2016b). Category-specific predictive analyses are less common but particularly informative. A notable example is Schöttke et al. (2014), who analyzed a mixed sample of CBT and psychodynamic outpatients and tested whether the number of patient-defined goals within specific goal categories predicted treatment outcome. They showed that including patient-defined goal categories improved the prediction of outcome beyond standardized measures alone, and that having more symptom-focused goals was associated with poorer treatment success. However, their study focused on which goals were selected and how many goals fell into each category, not on whether these goals were actually attained. Thus, it remains largely unclear whether the attainment of specific types of therapy goals (beyond their selection) differentially relates to treatment outcomes.

The present study addresses these gaps with two main objectives:

1. It examines how baseline patient characteristics are associated with therapy goal selection, specifically the likelihood of selecting each goal category at least once and the diversity of selected goals.
2. It investigates how both the selection and the attainment of specific goal categories relate to psychotherapy outcomes. Extending prior work that focused primarily on goal selection (Schöttke et al., 2014) or overall attainment (Baur et al., 2024; Ramnerö & Jansson, 2016b), this study tests whether category-specific goal attainment

differentially predicts multiple outcome domains, ranging from symptom change to broader indicators such as global functioning and patient satisfaction.

Together, these analyses aim to clarify how individualized therapy goals connect to standardized outcome measures and to inform more patient-tailored treatment planning.

Method

Sample

The data were drawn from an archival outpatient psychotherapy dataset comprising 1,116 patients. For the present study, patients were included if they had reported at least three therapy goals. The resulting analytic sample consisted of 645 patients. The patients were on average 36.88 years old (*SD* = 13.95), almost two-thirds identified as female (62.6%), and reported to be in a relationship (64.9%). The majority of patients (58.9%) were primarily diagnosed with an affective disorder (ICD-10 F3; World Health Organization, 2004), most commonly recurrent major depressive disorder (moderate episode; 35.8%) and a single episode of major depressive disorder (moderate; 29.1%). Besides depressive disorders, prevalent diagnoses in the present sample were anxiety disorders (20.2%), reactions to severe stress and adjustment disorder (10.5%), somatoform disorder (6.7%), personality disorder (5.5%), and eating disorder (3.9%). Half of the patients (51.6%) were diagnosed with at least one comorbid disorder, where the mean number of diagnoses was 1.69 (*SD* = 0.83, range = 1-7).

Treatment

The treatment was an evidence-based cognitive-behavioral therapy (Margraf & Schneider, 2009) in the outpatient clinic of Osnabruck University in Northwest Germany. Treatments were conducted by master-level psychologists in advanced clinical training and were regularly supervised by experienced licensed psychotherapists. All trainees had previously obtained their master's degree in psychology and at least two years of postgraduate training. There were 105 therapists (77.8% female) in the sample, each treating an average of 6.14 patients (*SD* = 3.16). The number of patients treated by individual therapists ranged from 1 to 13.

Measurements

Therapy Goals and Their Attainment

Therapy goals were categorized using the five main categories of the Bern Inventory of Treatment Goals – Checklist (BIT-C; Grosse Holtforth, 2001): symptom- and problem-related, interpersonal-related, personal growth-related, existential-related, and well-being-related goals. Only these five main categories were used in the present analyses. Each

goal in the BIT-C is pre-assigned to one main category via its item code, and these predefined assignments were used directly in the dataset. The patients select their so-called naive therapy goals, which are then specified together with the therapist, before the start of the treatment (Grosse Holtforth, 2001). The formulated goals are the basis for the GAS (Kiresuk et al., 1994). Goal attainment is finally assessed by both the patient and therapist at the end of therapy. The GAS is rated on a 7-point rating scale (-3 "very far from achieved" to +3 "fully achieved") and shows good reliability, depending on the conditions of its application (e.g., Krasny-Pacini et al., 2016).

Symptom Distress

General psychological distress was assessed using the German version of the Symptom Checklist - Revised (SCL-90-R; Franke, 2002). The SCL-90-R is a self-report inventory consisting of 90 items, which evaluates psychological symptoms in the last seven days. Items are rated on a 5-point Likert scale ranging from 0 (not at all) to 4 (extremely). The mean score of the SCL-90-R is referred to as the global severity index (GSI) and indicates the general psychological distress of an individual. The patients in the current study completed the SCL-90-R before (t1) and after (t2) treatment. The instrument had excellent reliability (Cronbach's α_{t1} = .97 and α_{t2} = .97).

Global Functioning

The German version of the Outcome Questionnaire-30 (OQ-30; Lambert et al., 2002) was used to evaluate the psychotherapy outcome as a more global measure than just symptom distress. The OQ-30 measures the functioning level of the past seven days with 30 items, which patients rate on a 5-point Likert scale from 0 (never) to 4 (almost always). The global functioning is assessed across three areas: symptom distress, interpersonal relationships, and social integration. In previous studies, the instrument has shown good reliability (Cronbach's α = .93) and test-retest-reliability (α = .88; Lambert et al., 2002). In the current study, the OQ-30 showed good reliability (Cronbach's α_{t1} = .80 and α_{t2} = .85).

Patients' Treatment Satisfaction

Patient satisfaction was assessed at the end of the treatment, using the German version of the Client Satisfaction Questionnaire (CSQ-8; Schmidt et al., 1989). The CSQ-8 is a self-report questionnaire, on which patients rate their satisfaction with the treatment on 8 items with a 4-point Likert scale (1 = lowest; 4 = highest). Previously, the instrument has shown excellent reliability (Cronbach's α = .92; Schmidt et al., 1989). In the current study, the reliability was good, Cronbach's α = .86, at the end of the treatment.

Procedure

The data were part of the routine outcome monitoring assessment of the outpatient clinic. Patients provided written informed consent and completed the baseline question-

naires. During the probationary treatment period (consisting of the first six to eight sessions), patients and therapists collaboratively selected a maximum of five therapy goals (consistent with the standard GAS framework) using the BIT-C. The goal number of the BIT-C allowed for the goals to be categorized into one of the five categories described above. These therapy goals were then specified by the patients and therapists, and if needed, refined in a person-specific way. This formulation was used as the basis for the goal attainment rating at the end of the treatment. For the purpose of the current study, a minimum number of goals was required to ensure sufficient data for analyzing how different goal categories impact therapy outcomes (original sample size = 1,116). Therefore, patients were included when they had at least three goals provided, and when their attainment was rated at the end of the psychotherapeutic intervention (n = 645).

Data Analysis

Predicting Goal Selection

In order to investigate the factors predicting the selection of therapy goals across the five predefined goal-categories (symptom-focused, interpersonal, personal-growth, existential, and well-being goals), we employed a series of multilevel logistic regression models (Wong & Mason, 1985). These models were implemented using the *lme4* package (Bates et al., 2015). Each model treated the binary indicator of goal selection (0 = not selected, 1 = selected at least once) as the outcome variable and included the following baseline predictors: age, gender, baseline symptom severity as measured by the SCL-90-R global severity index and its subscales (z-standardized), and global functioning as measured by the mean of the OQ-30 and its subscales (z-standardized). A random intercept for therapists was incorporated to account for the nested data structure, where patients were grouped within therapists. To investigate the factors predicting goal diversity, defined as the number of different therapy goal categories selected by a patient (range: 1-5), we employed multilevel ordinal regression using the *clmm* function from the *ordinal* package (Christensen, 2023).

Predicting Treatment Outcome by Goal Selection and Attainment

To create the predictor variables, we calculated each patient's mean attainment score for each goal category. Specifically, if a patient selected more than one goal from a category, we averaged the scores for those goals; if a patient did not select any goal from a category, the cell remained empty ("*NA*") to indicate non-selection. In addition, a separate binary indicator was created to show whether a category was selected at least once.

Consequently, the predictors included the following two parameters for the five goal categories (symptom-related, interpersonal-related, personal growth-related, existential-related, and well-being-related):

- the goal selection: 1 = yes, the goal was selected at least once; 0 = no, the goal was never selected
- the patient's mean attainment score (conditional on selection): a score ranging from -3 (much less attained than expected) to 3 (much more attained than expected), with NA indicating that the category was never chosen by the patient, and being modeled as structurally missing, with attainment effects estimated only for selected categories (selection × attainment interaction).

Outcome variables were defined as residualized post-treatment scores for symptom distress and psychotherapy outcome (PO; Cronbach & Furby, 1970). The post-treatment scores were regressed on the pre-treatment scores,

$$Predicted\ Post_{PO} = \beta_0 + \beta_1 x\ \mathrm{Pre}_{PO}$$

and the difference between the observed and the predicted post-treatment score was computed,

$$Residualized\ Post_{PO} = \mathrm{Observed\ Post}_{PO} - \mathrm{Predicted\ Post}_{PO}$$

where lower residualized scores indicate lower symptom distress than expected based on baseline severity.

Since goal attainment scores were only available for selected goal categories, missing values in the goal attainment predictors indicated structural non-applicability (meaning that a specific goal was never selected) rather than random missingness. In standard regression models, such missing predictor values would typically lead to listwise deletion of cases. To avoid this information loss, Bayesian multilevel regression models were estimated using the *brms* package (Bürkner, 2017), applying joint modeling of predictors with structurally missing values via the *mi*() specification. For each attainment predictor, an intercept-only auxiliary submodel (*mi*() ~ 1) was included. Attainment effects were modeled conditionally on goal selection using interaction terms (selection × attainment), so that attainment contributed to the outcome model only for categories that were actually selected, and cases were not additionally lost due to missing attainment predictors. Separate models were estimated for residualized posttreatment symptom distress, residualized posttreatment global outcome, and post-treatment patient satisfaction.

Random intercepts for therapists accounted for clustering within therapists. The intraclass coefficient (ICC) was calculated, which indicates the proportion of total variance attributable to the therapist-level in percent (Bryk & Raudenbush, 1987; Devine et al., 2024; Raudenbush & Bryk, 2002). The data analyses were performed in the R environment (Version 4.4.2, R Core Team, 2018).

Results

Descriptives

Table 1 presents descriptive statistics for goal selection and goal attainment across categories. The ranking of selected therapy goals in terms of frequency was as follows: a) symptom- and problem-related, b) personal growth-related, c) interpersonal-related, d) well-being-related, and e) existential-related therapy goals. For about half of the 645 patients, the set goals were selected from three distinct categories (n = 314, 48.7%), for about a quarter of the patients from four distinct categories (n = 168, 26.0%), for about one fifth of the patients from two distinct categories (n = 119, 18.5%), and for a vast minority from either one category (n = 25, 3.9%) or the maximum of five categories (n = 19, 2.9%).

Table 1

Descriptive Statistics

	Goal selection		Goal attainment	
Variable	Selected (at least) once	Never selected	*M* (range -3 to +3)	*SD*
Symptom related	620	25	1.96	0.88
Interpersonal related	387	258	1.85	0.99
Personal growth-related	434	211	1.90	0.84
Existential related	170	475	1.82	1.06
Well-being related	230	415	1.84	0.97

Note. n = 645; M = mean; SD = standard deviation.

The mean scores of the SCL-90R at the beginning (t_1) and end (t_2) of therapy for GSI t_1 = 1.00 (SD = 0.57) and t_2 = 0.48 (SD = 0.46). For the OQ-30, the mean scores before and after therapy were t_1 = 1.69 (SD = 0.55) and t_2 = 0.93 (SD = 0.56) for the overall score. The overall mean patients' treatment satisfaction was 3.76 (SD = 0.32). A correlation matrix of the study variables, including the outcome prediction, is provided in Table 2.

Goal Selection

Table 3 presents the results of all the models. The symptom goal selection model showed an ICC of 0.428 in the null model, indicating substantial therapist-level variance on symptom goal selection. In the full model, the only significant predictor was interpersonal problems (b = -.705, SE = 0.355, 95% CI [-1.401, -0.010], p = .046), with greater interpersonal problems being associated with a decreased likelihood of selecting symptom goals (OR = 0.494). The ICC for the full model was 0.398, indicating that therapist

differences continued to account for a substantial portion of the variance even after adjusting for patient characteristics.

Table 2

Correlation Coefficients Among all Study Variables

Variable	1	2	3	4	5	6	7	8	9
1. GSI pre	–								
2. GSI post	.496*	–							
3. OQ-30 pre	.779*	.408*	–						
4. OQ-30 post	.415*	.841*	.456*	–					
5. CSQ-8 post	-.082*	-.303*	-.090*	-.348*	–				
6. GAS symptom	-.077	-.448*	-.124*	-.532*	.342*	–			
7. GAS interpersonal	-.055	-.325*	-.123*	-.386*	.328*	.436*	–		
8. GAS personal growth	-.083	-.338*	-.139*	-.466*	.256*	.563*	.464*	–	
9. GAS existential	-.050	-.310*	-.049	-.358*	.344*	.588*	.578*	.461*	–
10. GAS well-being	-.084	-.344*	-.135*	-.394*	.181*	.497*	.504*	.466*	.548*

Note. Pre refers to the pre-treatment score, post to the post-treatment score; GSI = Global Severity Index; OQ-30 = Outcome Questionnaire; CSQ-8 = Client Satisfaction Questionnaire; GAS = Goal Attainment Scaling at the end of treatment.

*$p < .05$.

The interpersonal goal selection model showed an ICC of ~ 0.000 in the null model, indicating minimal therapist influence on interpersonal goal selection. Given the negligible therapist-level variance and boundary fit in the multilevel model, a sensitivity analysis was conducted using a single-level logistic regression. The results were identical in terms of fixed effect estimates and significance levels. In the full model, significant predictors included anxiety (b = -.312, SE = 0.154, 95% CI [-0.613, -0.011], p = .042, OR = 0.732) and hostility (b = -.408, SE = 0.134, 95% CI [-0.671, -0.146], p = .002, OR = 0.665), both of which decreased the likelihood of selecting interpersonal goals, as well as paranoid ideation (b = .381, SE = 0.164, 95% CI [0.059, 0.703], p = .020, OR = 1.464) and interpersonal problems (b = .347, SE = 0.125, 95% CI [0.102, 0.593], p = .006, OR = 1.415), which increased the likelihood. The ICC for the full model remained 0.000.

The personal growth-related goals model showed an ICC of 0.126 in the null model. In the full model, the only significant predictor was somatic symptoms (b = -.279, SE = 0.137, 95% CI [-0.547, -0.010], p = .042, OR = 0.757), which decreased the likelihood of selecting personal growth-related goals. The ICC for the full model was 0.154.

The existential goal selection model showed an ICC of 0.047 in the null model. The only significant predictor was symptoms of depression (b = .448, SE = 0.199, 95% CI [0.058, 0.838], p = .024, OR = 1.566), which increased the likelihood of selecting existential goals. The ICC for the full model was 0.056.

Clinical Psychology in Europe
2026, Vol. 8(3), Article e19567
https://doi.org/10.32872/cpe.19567

Table 3

Predictors of Goal Type Selection

	Goal categories					Goal diversity
Baseline Characteristic	**Symptom** *Log-odds (SE)* *[95% CI]*	**Interpersonal** *Log-odds (SE)* *[95% CI]*	**Personal growth** *Log-odds (SE)* *[95% CI]*	**Existential** *Log-odds (SE)* *[95% CI]*	**Well-being** *Log-odds (SE)* *[95% CI]*	*Log-odds (SE)* *[95% CI]*
Sociodemographic						
Age	0.013 (0.021) [-0.029, 0.058]	0.005 (0.007) [-0.008, 0.019]	-0.003 (0.008) [-0.018, 0.012]	0.006 (0.008) [-0.009, 0.022]	0.004 (0.007) [-0.010, 0.018]	0.018* (0.007) [0.005, 0.031]
Gender (female)	-0.843 (0.604) [-2.026, 0.340]	0.218 (0.190) [-0.154, 0.589]	-0.017 (0.210) [-0.428, 0.394]	0.287 (0.215) [-0.134, 0.708]	-0.193 (0.191) [-0.568, 0.182]	0.195 (0.180) [-0.159, 0.548]
Symptom Checklist-90R – Subscales						
Somatization	0.114 (0.395) [-0.660, 0.888]	0.221 (0.129) [-0.032, 0.474]	-0.279* (0.137) [-0.547, -0.010]	-0.114 (0.138) [-0.385, 0.157]	-0.008 (0.127) [-0.256, 0.240]	-0.159 (0.121) [-0.395, 0.078]
Obsessive compulsion	0.265 (0.434) [-0.586, 1.115]	-0.135 (0.150) [-0.429, 0.159]	-0.046 (0.163) [-0.364, 0.273]	-0.239 (0.167) [-0.567, 0.089]	0.046 (0.150) [-0.248, 0.341]	-0.161 (0.138) [-0.431, 0.110]
Interpersonal sensitivity	0.057 (0.543) [-1.007, 1.121]	0.262 (0.180) [-0.089, 0.614]	-0.058 (0.192) [-0.434, 0.319]	-0.021 (0.193) [-0.398, 0.357]	-0.128 (0.179) [-0.478, 0.223]	0.069 (0.165) [-0.254, 0.392]
Depression	0.043 (0.518) [-0.972, 1.058]	0.238 (0.186) [-0.126, 0.602]	0.214 (0.199) [-0.175, 0.604]	0.448* (0.199) [0.058, 0.838]	-0.206 (0.186) [-0.571, 0.160]	0.371* (0.172) [0.034, 0.708]
Anxiety	-0.082 (0.462) [-0.987, 0.824]	-0.312* (0.154) [-0.613, -0.011]	0.112 (0.164) [-0.210, 0.434]	-0.091 (0.166) [-0.416, 0.234]	-0.089 (0.152) [-0.386, 0.208]	-0.144 (0.142) [-0.422, 0.135]

Baseline Characteristic	Goal categories					Goal diversity
	Symptom *Log-odds (SE)* *[95% CI]*	**Interpersonal** *Log-odds (SE)* *[95% CI]*	**Personal growth** *Log-odds (SE)* *[95% CI]*	**Existential** *Log-odds (SE)* *[95% CI]*	**Well-being** *Log-odds (SE)* *[95% CI]*	*Log-odds (SE)* *[95% CI]*
Hostility	-0.164 (0.340) [-0.830, 0.501]	-0.408* (0.134) [-0.671, -0.146]	0.105 (0.143) [-0.176, 0.386]	-0.125 (0.144) [-0.409, 0.158]	0.042 (0.133) [-0.220, 0.303]	-0.095 (0.121) [-0.332, 0.143]
Phobic anxiety	0.720 (0.503) [-0.266, 1.705]	-0.138 (0.120) [-0.374, 0.098]	-0.159 (0.129) [-0.412, 0.094]	-0.054 (0.134) [-0.318, 0.209]	0.084 (0.120) [-0.151, 0.320]	-0.084 (0.110) [-0.300, 0.133]
Paranoid ideation	0.083 (0.444) [-0.788, 0.953]	0.381* (0.164) [0.059, 0.703]	-0.002 (0.172) [-0.339, 0.335]	0.022 (0.172) [-0.315, 0.360]	0.134 (0.159) [-0.178, 0.447]	0.124 (0.149) [-0.168, 0.416]
Psychoticism	0.061 (0.439) [-0.799, 0.920]	0.120 (0.141) [-0.156, 0.395]	-0.027 (0.150) [-0.320, 0.267]	0.142 (0.150) [-0.153, 0.437]	0.012 (0.139) [-0.261, 0.284]	0.174 (0.128) [-0.076, 0.424]
Outcome Questionnaire-30 – Subscales						
Well-being	0.257 (0.519) [-0.760, 1.273]	-0.211 (0.187) [-0.578, 0.156]	-0.120 (0.202) [-0.516, 0.276]	0.053 (0.206) [-0.351, 0.457]	0.404* (0.188) [0.034, 0.773]	0.029 (0.173) [-0.310, 0.368]
Social role performance	-0.019 (0.369) [-0.743, 0.705]	-0.035 (0.135) [-0.298, 0.229]	0.036 (0.148) [-0.253, 0.326]	0.094 (0.147) [-0.196, 0.383]	-0.251 (0.135) [-0.516, 0.014]	-0.041 (0.126) [-0.288, 0.206]
Interpersonal problems	-0.705* (0.355) [-1.401, -0.010]	0.347* (0.125) [0.102, 0.593]	-0.016 (0.134) [-0.279, 0.247]	-0.118 (0.135) [-0.382, 0.147]	-0.213 (0.126) [-0.459, 0.033]	-0.098 (0.115) [-0.322, 0.127]
Therapist-level variance (ICC)	0.398	0.000	0.154	0.056	0.034	0.204

Note. ICC = intraclass coefficient; values represent log-odds coefficients from mixed-effects logistic (goal selection) and ordinal (goal diversity) regression models, with 95% Wald confidence intervals [CI] in brackets; overall scores from the SCL-90R (GSI) and OQ-30 were not included in the models due to multicollinearity.
*$p < .05$.

Clinical Psychology in Europe
2026, Vol. 8(3), Article e19567
https://doi.org/10.32872/cpe.19567

The well-being goal selection model showed an ICC of 0.030 in the null model. The only significant predictor in the full model was the OQ-30 subscale of well-being (b = .404, SE = 0.188, 95% CI [0.034, 0.773], p = .032, OR = 1.497), which increased the likelihood of selecting well-being goals. The ICC for the full model remained 0.034.

The goal diversity model showed an ICC of 0.167 in the null model. In the full model, significant predictors included patient age (b = .018, SE = 0.007, 95% CI [0.005, 0.031], p = .007, OR = 1.018), with older patients more likely to set more diverse goals, and depression (b = .371, SE = 0.172, 95% CI [0.034, 0.708], p = .031, OR = 1.449), which also increased the diversity of goals. The ICC for the full model was 0.204.

Goal Selection and Attainment, and Their Association With Psychotherapy Outcomes

Building on the goal selection results, the second analysis examined whether the conditional attainment (as an interaction with the selection) of different therapy goals can predict residualized post-treatment scores in symptoms, global functioning, and end-of-treatment patient satisfaction (see Table 4).

For symptom change as the psychotherapy outcome, the mean attainment of symptom-, interpersonal-, personal growth-, and well-being-related goals was associated with greater reductions in symptoms, beyond what would be expected based on patients' initial severity. The ICC for symptom changes was 0.047, indicating that approximately 5% of the variance in symptom improvement was attributable to therapist differences.

For improvements in overall global functioning, higher mean attainment of symptom- and personal growth-related goals was associated with better-than-expected outcomes. The ICC for global functioning was 0.019.

For patients' treatment satisfaction, only a higher mean attainment of symptom-related goals was associated with higher treatment satisfaction. The ICC for patient satisfaction was 0.000, suggesting that satisfaction was driven almost entirely by patient-level factors.

Clinical Psychology in Europe
2026, Vol. 8(3), Article e19567
https://doi.org/10.32872/cpe.19567

Table 4

Predictors of Residualized Post-Treatment Scores in Symptoms, Global Functioning, and End-Of-Treatment Patient Satisfaction

	Symptoms[a]	Global Functioning[b]	Patient Satisfaction[c]
Variable	***PM (PSD)*** ***[95% CrInt]***	***PM (PSD)*** ***[95% CrInt]***	***PM (PSD)*** ***[95% CrInt]***
Intercept	-0.034 (0.078) [-0.190, 0.116]	-0.089 (0.099) [-0.286, 0.105]	29.735 (0.260) [29.288, 30.321]
Selection of goal at least once			
Symptom-related	0.308* (0.073) [0.167, 0.453]	0.537* (0.096) [0.352, 0.728]	-0.441* (0.246) [-0.999, -0.039]
Interpersonal-related	0.031 (0.043) [-0.053, 0.114]	0.038 (0.055) [-0.069, 0.146]	-0.090 (0.119) [-0.321, 0.148]
Personal growth-related	0.105* (0.046) [0.014, 0.195]	0.210* (0.062) [0.088, 0.331]	-0.151 (0.129) [-0.405, 0.106]
Existential-related	0.133* (0.053) [0.028, 0.236]	0.211* (0.070) [0.072, 0.347]	-0.134 (0.133) [-0.382, 0.143]
Well-being-related	0.118* (0.049) [0.020, 0.214]	0.059 (0.065) [-0.070, 0.187]	0.043 (0.131) [-0.194, 0.316]

Variable	Symptoms[a]	Global Functioning[b]	Patient Satisfaction[c]
	PM (PSD) *[95% CrInt]*	*PM (PSD)* *[95% CrInt]*	*PM (PSD)* *[95% CrInt]*
Average goal attainment of:			
selection: symptom-related	-0.143* (0.016) [-0.175, -0.110]	-0.222* (0.023) [-0.267, -0.177]	0.112* (0.051) [0.014, 0.213]
selection: interpersonal-related	-0.039* (0.019) [-0.075, -0.003]	-0.040 (0.023) [-0.086, 0.006]	0.042 (0.055) [-0.063, 0.152]
selection: personal growth-related	-0.039* (0.020) [-0.078, -0.001]	-0.117* (0.026) [-0.169, -0.065]	0.062 (0.060) [-0.053, 0.182]
selection: existential-related	-0.031 (0.024) [-0.078, 0.017]	-0.058 (0.032) [-0.121, 0.006]	0.046 (0.063) [-0.076, 0.173]
selection: well-being-related	-0.060* (0.023) [-0.105, -0.015]	-0.038 (0.030) [-0.096, 0.020]	-0.017 (0.060) [-0.138, 0.098]
Therapist-level variance (ICC)	0.047	0.019	0.000

Note. PM = posterior mean; PSD = posterior standard deviation; CrInt = credible interval; ICC = intraclass coefficient.
[a]n = 636. [b]n = 633. [c]n = 627.
*$p < .05$.

Discussion

This study examined how different goal contents (categorized according to the BIT-C) of collaboratively defined therapy goals relate to psychotherapy outcome in outpatients receiving cognitive-behavioral psychotherapy, addressing two gaps in the current literature: which patient characteristics are associated with the selection of specific goal categories, and whether the attainment of different goal categories shows differential associations with treatment outcomes. The present findings show that therapy goal selection varies systematically with certain patient characteristics and partially with therapist factors, and that goal attainment is associated with better-than-expected treatment outcomes.

Descriptively, most patients selected symptom-related goals at least once, followed by personal growth and interpersonal goals, while existential and well-being goals were chosen less frequently. Attainment of symptom-related goals was generally higher than that of other goal types, with existential and interpersonal goals showing slightly lower levels of attainment. To contextualize, this pattern is largely consistent with prior studies conducted in CBT settings, where symptom-focused goals tend to dominate (Baur et al., 2024; Schöttke et al., 2014). It likely reflects the structured, problem-focused orientation of CBT, which emphasizes symptom reduction and functional improvement (Margraf & Schneider, 2009).

This interpretation is further supported by the goal selection results, which showed substantial therapist-level effects, particularly for symptom-related goals (≈ 40%), but also for goal diversity (≈ 20%) and personal growth–related goals (≈ 15%). When placed in perspective, these therapist effects appear substantial. Therapist effects in psychotherapy typically account for about 4–17% of outcome variance (see Wampold & Owen, 2021), about 6–13% of dropout variance (Saxon et al., 2017; Zimmermann et al., 2017), and roughly 11–16% of the variance in the alliance–outcome association (Del Re et al., 2012). Therefore, before discussing the potential predictor of goal selection, it is important to note that goal selection is (at least for some goal categories) strongly shaped by therapist guidance and treatment context, not only by patient characteristics. Nevertheless, several patient intake characteristics showed meaningful associations with goal content. For example, higher levels of interpersonal problems were associated with a lower likelihood of selecting at least one symptom-related goal and a higher likelihood of selecting at least one interpersonal-related goal. This suggests that patients tend to formulate goals in the domain where they experience high distress. In contrast, symptoms of depression were linked to a higher likelihood of selecting existential-related goals and greater goal diversity, the latter being consistent with prior findings (Grosse Holtforth et al., 2009). Interestingly, being older is also associated with a higher probability of selecting a more diverse set of therapy goals, potentially due to the addition of age-related challenges (Choi et al., 2021; Raue et al., 2017; Vathke et al., 2023). On the other hand, other symptoms showed a more selective shift of goal content rather than broader effects.

For example, higher somatization was associated with a lower likelihood of selecting personal growth–related goals, which may reflect a stronger focus on bodily distress (Miller, 2008) and potentially a reduced capability for self-exploratory change processes (Ballespí et al., 2019; Mattila et al., 2008). Together, these findings indicate that therapy goal categories capture clinically meaningful differences in how patients conceptualize their difficulties at treatment entry. As such, structured goal categorization may serve as a useful bridge between individualized goal setting and more standardized approaches to case formulation and treatment planning.

Alongside the question of which therapy goal categories patients select, the present findings also illustrate how progress toward these goals (in the specific context of this study) relates to therapeutic change. Goal attainment represents a direct indicator of patient-perceived improvements with respect to defined change targets (Lloyd et al., 2019) and is a complementary perspective on treatment success beyond standardized symptom measures (Lindhiem et al., 2016). Across different domains of treatment success, from symptom distress to broader functioning and patient satisfaction, attainment within several goal categories (conditional on them being selected) was associated with greater-than-expected improvement relative to baseline severity, indicating that progress on individualized goals is meaningfully reflected in standardized outcomes. The strength of these associations differed across outcome domains. In this sample, symptom change showed the most consistent links with specific goal attainment, whereas associations with global functioning and patient satisfaction were fewer and more selective. One likely explanation lies in the CBT treatment context, which, as mentioned above, emphasizes symptom reduction. On the other hand, broader aspects of global functioning (e.g., interpersonal problems and role domains) may change more slowly and not be fully captured at post-assessment (e.g., Quilty et al., 2013). Importantly, some goal content categories were selected less frequently, which reduces statistical power for their attainment effects and cautions against interpreting non-significant associations as evidence of irrelevance. The main effects of goal category selection should be interpreted with particular caution as well. These coefficients should not be interpreted as indicating that selecting such goals leads to poorer outcomes. Rather, they reflect differences between patients who selected a given goal category and those who did not. Patients selecting certain goal types may present with more complex or burdensome problem profiles, which are generally associated with less favorable outcomes (e.g., Boswell et al., 2012).

What, then, should be made of these findings? They align with the view that psychotherapy outcome cannot be reduced to symptom change alone but encompasses multiple domains of personally meaningful improvement. Patient-reported outcome research has repeatedly demonstrated that patients define their therapy success in diverse ways. For example, a qualitative meta-synthesis by Chevance et al. (2020) identified a large number of distinct outcome domains that patients consider relevant. At a more integrative level, Zavlis et al. (2025) proposed that valued therapy outcomes cluster around three broad

themes (love, meaning, and work; with the latter encompassing symptom reduction and functioning). Across these perspectives, a common conclusion emerges: standardized symptom measures capture only part of what patients experience as improvement. Accordingly, De Smet et al. (2020) suggest that changes in standardized outcome measures should be contextualized within patients' subjective narrative. Against this background, the present results add a quantitative perspective by showing that attainment in specific patient-defined goal content domains is reflected in standardized outcome indicators, but not uniformly. Therefore, the bridge between individualized goals and standardized outcomes seems domain-sensitive and not generic. Clinically, the present findings suggest that domain-specific goal attainment ratings could complement the standardized outcome evaluation. At the same time, standardized intake measures may help identify dominant distress domains and thereby inform the collaborative formulation of therapy goals.

Several questions remain open for future research. First, goal selection as well as goal attainment–outcome associations may differ across therapeutic orientations and treatment settings (Schöttke et al., 2014). Second, it would be valuable to examine whether similar domain-specific patterns emerge when using other goal-based or idiographic outcome approaches (Cooper & Xu, 2023; Sales, Faísca, et al., 2023), and how the BIT goal domains map onto broader patient-defined outcome domains described in qualitative and mixed-methods research (Chevance et al., 2020; Zavlis et al., 2025). Future research should also continue to examine how different goal-domain frameworks align with patients' own conceptualizations of change and outcome. Third, future studies could extend this work by examining associations between goal attainment and negative outcomes, such as deterioration or dissatisfaction with treatment (Steinbrenner et al., 2025). Finally, therapy goals were treated here as fixed at baseline, although goal priorities and meanings may shift over the course of treatment. Capturing the dynamic development of goals and their attainment over time may provide a more process-sensitive understanding of how individualized goals relate to therapeutic change.

Limitations

Several limitations should be noted when interpreting the present findings. First, as mentioned above, the number of patients selecting certain goal categories, particularly existential and well-being-related goals, was relatively low, which may have reduced statistical power to detect meaningful effects in these domains. However, this lower frequency may also reflect other possibilities, such as lower perceived relevance, greater difficulty in articulating these goals, or differences in how they are framed during goal setting. Second, because only patients who reported at least three therapy goals were included a priori, the findings may generalize primarily to patients who are able and willing to engage in collaborative goal setting. As goal articulation is likely associated with motivation and therapeutic alliance (e.g., DeFife & Hilsenroth, 2011; Ryan et al.,

2011), a higher number of reported goals (independent of their content or diversity) may itself reflect more favorable psychotherapy process characteristics. Third, all therapists in the study were CBT trainees, potentially limiting the generalizability of the results to other therapeutic orientations or more experienced clinicians. Fourth, goal types were categorized based on their position within the BIT structure. However, some goals may have conceptually overlapped across categories, introducing ambiguity in the classification. Finally, all outcome measures were based on patient self-report, where therapist-rated improvements could also be important indicators.

Conclusion

This study highlights the clinical value of collaboratively defined therapy goals and their attainment as meaningful indicators of therapeutic outcome. Progress toward individualized goals was systematically reflected in standardized outcome measures, although not uniformly across goal content domains. Substantial therapist-level effects in goal selection indicate that goal formulation is shaped not only by patient characteristics but also by therapeutic context. Taken together, these results support the integration of individualized therapy goals alongside standardized measures to enable a more nuanced and patient-centered assessment of psychotherapy outcomes.

Funding: The authors have no funding to report.

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

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

Ethics Statement: The procedures of this study were approved by the ethical review board of Osnabrueck University (Ethik-28/2025). All subjects gave written informed consent in accordance with the Declaration of Helsinki (2013).

Related Versions: We note that the data stem from the same outpatient clinic as in Schöttke et al. (2014), but the samples do not overlap, as the earlier study included patients treated between 2002 and 2010, whereas the present study draws on patients treated from 2014 onwards.

Reporting Guidelines: This article is written according to the JARS-QUANT guidelines.

Preregistration: The analyses for this study were not preregistered.

Data Availability: The datasets generated and analyzed in the current study contain sensitive patient information and therefore cannot be shared publicly. Data may be made available upon reasonable request to the corresponding author, in accordance with institutional and ethical guidelines.

References

Ballespí, S., Vives, J., Alonso, N., Sharp, C., Ramírez, M. S., Fonagy, P., & Barrantes-Vidal, N. (2019). To know or not to know? Mentalization as protection from somatic complaints. *PLoS One, 14*(5), Article e0215308. https://doi.org/10.1371/journal.pone.0215308

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baur, A., Trösken, A., & Renneberg, B. (2024). Content and attainment of individual treatment goals in CBT. *Psychotherapy Research, 34*(1), 111–123. https://doi.org/10.1080/10503307.2023.2173681

Berking, M., Grosse Holtforth, M., Jacobi, C., & Kröner-Herwig, B. (2005). Empirically based guidelines for goal-finding procedures in psychotherapy: Are some goals easier to attain than others? *Psychotherapy Research, 15*(3), 316–324. https://doi.org/10.1080/10503300500091801

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy, 16*(3), 252–260. https://doi.org/10.1037/h0085885

Boswell, J. F., Sauer-Zavala, S. E., Gallagher, M. W., Delgado, N. K., & Barlow, D. H. (2012). Readiness to change as a moderator of outcome in transdiagnostic treatment. *Psychotherapy Research, 22*(5), 570–578. https://doi.org/10.1080/10503307.2012.688884

Brunstein, J. C., Schultheiss, O. C., & Grässmann, R. (1998). Personal goals and emotional well-being: The moderating role of motive dispositions. *Journal of Personality and Social Psychology, 75*(2), 494–508. https://doi.org/10.1037/0022-3514.75.2.494

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101*(1), 147–158. https://doi.org/10.1037/0033-2909.101.1.147

Bürkner, P.-C. (2017). brms: An *R* package for Bayesian multilevel models using *Stan. Journal of Statistical Software, 80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Chevance, A., Ravaud, P., Tomlinson, A., Le Berre, C., Teufer, B., Touboul, S., Fried, E. I., Gartlehner, G., Cipriani, A., & Tran, V. T. (2020). Identifying outcomes for depression that matter to patients, informal caregivers, and health-care professionals: Qualitative content analysis of a large international online survey. *The Lancet Psychiatry, 7*(8), 692–702. https://doi.org/10.1016/S2215-0366(20)30191-7

Choi, N. G., Caamano, J., Vences, K., Marti, C. N., & Kunik, M. E. (2021). Acceptability and effects of tele-delivered behavioral activation for depression in low-income homebound older adults: In their own words. *Aging & Mental Health, 25*(10), 1803–1810. https://doi.org/10.1080/13607863.2020.1783516

Christensen, R. (2023). *Ordinal-regression models for ordinal data* [R package version 2023.12-4.1; computer software]. https://doi.org/10.32614/CRAN.package.ordinal

Cooper, M., & Xu, D. (2023). Goals Form: Reliability, validity, and clinical utility of an idiographic goal-focused measure for routine outcome monitoring in psychotherapy. *Journal of Clinical Psychology, 79*(3), 641–666. https://doi.org/10.1002/jclp.23344

Cronbach, L. J., & Furby, L. (1970). How we should measure „change“: Or should we? *Psychological Bulletin, 74*(1), 68–80. https://doi.org/10.1037/h0029382

DeFife, J. A., & Hilsenroth, M. J. (2011). Starting off on the right foot: Common factor elements in early psychotherapy process. *Journal of Psychotherapy Integration, 21*(2), 172–191. https://doi.org/10.1037/a0023889

Del Re, A. C., Flückiger, C., Horvath, A. O., Symonds, D., & Wampold, B. E. (2012). Therapist effects in the therapeutic alliance–outcome relationship: A restricted-maximum likelihood meta-analysis. *Clinical Psychology Review, 32*(7), 642–649. https://doi.org/10.1016/j.cpr.2012.07.002

Derogatis, L. R. (1977). *SCL-90-R, administration, scoring & procedures manual-I for the R(evised) version.* Johns Hopkins University School of Medicine.

De Smet, M. M., Meganck, R., De Geest, R., Norman, U. A., Truijens, F., & Desmet, M. (2020). What "good outcome" means to patients: Understanding recovery and improvement in psychotherapy for major depression from a mixed-methods perspective. *Journal of Counseling Psychology, 67*(1), 25–39. https://doi.org/10.1037/cou0000362

Devine, S., Uanhoro, J. O., Otto, A. R., & Flake, J. K. (2024). Approaches for quantifying the ICC in multilevel logistic models: A didactic demonstration. *Collabra: Psychology, 10*(1), Article 94263. https://doi.org/10.1525/collabra.94263

Dirmaier, J., Harfst, T., Koch, U., & Schulz, H. (2006). Therapy goals in inpatient psychotherapy: Differences between diagnostic groups and psychotherapeutic orientations. *Clinical Psychology & Psychotherapy, 13*(1), 34–46. https://doi.org/10.1002/cpp.470

Franke, G. H. (2002). *SCL-90-R: Symptom-Checkliste von L. R. Derogatis – Deutsche Version* (*2. Aufl.*) [SCL-90-R: Symptom Checklist of L. R. Derogatis – German Version (2nd ed.)]. Beltz Test GmbH.

Gleacher, A. A., Olin, S. S., Nadeem, E., Pollock, M., Ringle, V., Bickman, L., Douglas, S., & Hoagwood, K. (2016). Implementing a measurement feedback system in community mental health clinics: A case study of multilevel barriers and facilitators. *Administration and Policy in Mental Health, 43*(3), 426–440. https://doi.org/10.1007/s10488-015-0642-0

Grawe, K. (2004). *Psychological therapy.* Hogrefe & Huber.

Grosse Holtforth, M. (2001). Was möchten Patienten in ihrer Therapie erreichen? Die Erfassung von Therapiezielen mit dem Berner Inventar für Therapieziele (BIT) [What do patients want to achieve in their therapy? The assessment of therapy goals with the Bern Inventory of Treatment Goals (BIT)]. *Verhaltenstherapie & Psychosoziale Praxis, 34*(2), 241–258.

Grosse Holtforth, M., & Grawe, K. (2000). Fragebogen zur Analyse Motivationaler Schemata (FAMOS) [Questionnaire for the analysis of motivational schemas]. *Zeitschrift für Klinische Psychologie und Psychotherapie: Forschung und Praxis, 29*(3), 170–179. https://doi.org/10.1026//0084-5345.29.3.170

Grosse Holtforth, M., & Grawe, K. (2002). Bern Inventory of Treatment Goals: Part 1. Development and first application of a taxonomy of treatment goal themes. *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 12*(1), 79–99. https://doi.org/10.1080/713869618

Grosse Holtforth, M., & Michalak, J. (2012). Motivation in psychotherapy. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 441–462). Oxford University Press.

Grosse Holtforth, M., Reubi, I., Ruckstuhl, L., Berking, M., & Grawe, K. (2004). The value of treatment-goal themes for treatment planning and outcome evaluation of psychiatric inpatients. *The International Journal of Social Psychiatry, 50*(1), 80–91. https://doi.org/10.1177/0020764004040955

Grosse Holtforth, M., Wyss, T., Schulte, D., Trachsel, M., & Michalak, J. (2009). Some like it specific: The difference between treatment goals of anxious and depressed patients. *Psychology and Psychotherapy: Theory, Research and Practice, 82*(3), 279–290. https://doi.org/10.1348/147608308X397040

Hatfield, D. R., & Ogles, B. M. (2007). Why some clinicians use outcome measures and others do not. *Administration and Policy in Mental Health, 34*(3), 283–291. https://doi.org/10.1007/s10488-006-0110-y

Ionita, G., & Fitzpatrick, M. (2014). Bringing science to clinical practice: A Canadian survey of psychological practice and usage of progress monitoring measures. *Canadian Psychology / Psychologie Canadienne, 55*(3), 187–196. https://doi.org/10.1037/a0037355

Jensen-Doss, A., Smith, A. M., Becker-Haimes, E. M., Mora Ringle, V., Walsh, L. M., Nanda, M., Walsh, S. L., Maxwell, C. A., & Lyon, A. R. (2018). Individualized progress measures are more acceptable to clinicians than standardized measures: Results of a national survey. *Administration and Policy in Mental Health, 45*(3), 392–403. https://doi.org/10.1007/s10488-017-0833-y

Kiresuk, T. J., & Sherman, R. E. (1968). Goal attainment scaling: A general method for evaluating comprehensive community mental health programs. *Community Mental Health Journal, 4*(6), 443–453. https://doi.org/10.1007/BF01530764

Kiresuk, T. J., Smith, A., & Cardillo, J. E. (1994). *Goal attainment scaling*. Psychology Press. https://doi.org/10.4324/9781315801933

Krasny-Pacini, A., Evans, J., Sohlberg, M. M., & Chevignard, M. (2016). Proposed criteria for appraising goal attainment scales used as outcome measures in rehabilitation research. *Archives of Physical Medicine and Rehabilitation, 97*(1), 157–170. https://doi.org/10.1016/j.apmr.2015.08.424

Lambert, M. J., Hannöver, W., Nisslmüller, K., Richard, M., & Kordy, H. (2002). Fragebogen zum Ergebnis von Psychotherapie [Questionnaire on the results of psychotherapy]. *Zeitschrift für Klinische Psychologie und Psychotherapie, 31*(1), 40–46. https://doi.org/10.1026/0084-5345.31.1.40

Lambert, M. J., Harmon, C., Slade, K., Whipple, J. L., & Hawkins, E. J. (2005). Providing feedback to psychotherapists on their patients' progress: Clinical results and practice suggestions. *Journal of Clinical Psychology, 61*(2), 165–174. https://doi.org/10.1002/jclp.20113

Lindhiem, O., Bennett, C. B., Orimoto, T. E., & Kolko, D. J. (2016). A meta-analysis of personalized treatment goals in psychotherapy: A preliminary report and call for more studies. *Clinical Psychology, 23*(2), 165–176. https://doi.org/10.1111/cpsp.12153

Lloyd, C. E. M., Duncan, C., & Cooper, M. (2019). Goal measures for psychotherapy: A systematic review of self-report, idiographic instruments. *Clinical Psychology, 26*(3), Article e12281. https://doi.org/10.1111/cpsp.12281

Margraf, J., & Schneider, S. (2009). *Lehrbuch der Verhaltenstherapie* [Textbook to cognitive-behavioral therapy]. Springer Medizin Verlag.

Mattila, A. K., Kronholm, E., Jula, A., Salminen, J. K., Koivisto, A.-M., Mielonen, R.-L., & Joukamaa, M. (2008). Alexithymia and somatization in general population. *Psychosomatic Medicine, 70*(6), 716–722. https://doi.org/10.1097/PSY.0b013e31816ffc39

Michalak, J., & Grosse Holtforth, M. (2006). Where do we go from here? The goal perspective in psychotherapy. *Clinical Psychology, 13*(4), 346–365. https://doi.org/10.1111/j.1468-2850.2006.00048.x

Michalak, J., Grosse Holtforth, M., & Berking, M. (2007). Patientenziele in der Psychotherapie [The goals of clients in psychotherapy]. *Psychotherapeut, 52*(1), 6–15. https://doi.org/10.1007/s00278-006-0518-6

Michalak, J., Klappheck, M. A., & Kosfelder, J. (2004). Personal goals of psychotherapy patients: The intensity and the "why" of goal-motivated behavior and their implications for the therapeutic process. *Psychotherapy Research, 14*(2), 193–209. https://doi.org/10.1093/ptr/kph017

Michalak, J., & Schulte, D. (2002). Zielkonflikte und Therapiemotivation [Goal conflicts and therapy motivation]. *Zeitschrift für Klinische Psychologie und Psychotherapie, 31*(3), 213–219. https://doi.org/10.1026/0084-5345.31.3.213

Miller, R. C. (2008). The somatically preoccupied patient in primary care: Use of attachment theory to strengthen physician-patient relationships. *Osteopathic Medicine and Primary Care, 2*(1), Article 6. https://doi.org/10.1186/1750-4732-2-6

Puschner, B., Becker, T., & Bauer, S. (2015). Routine outcome measures in Germany. *International Review of Psychiatry, 27*(4), 329–337. https://doi.org/10.3109/09540261.2015.1014025

Quilty, L. C., Mainland, B. J., McBride, C., & Bagby, R. M. (2013). Interpersonal problems and impacts: Further evidence for the role of interpersonal functioning in treatment outcome in major depressive disorder. *Journal of Affective Disorders, 150*(2), 393–400. https://doi.org/10.1016/j.jad.2013.04.030

Ramnerö, J., & Jansson, B. (2016a). The stability of treatment goals, as assessed by a Swedish version of the Bern Inventory of Treatment Goals. *Nordic Psychology, 68*(1), 30–39. https://doi.org/10.1080/19012276.2015.1045544

Ramnerö, J., & Jansson, B. (2016b). Treatment goals and their attainment: A structured approach to assessment and evaluation. *Cognitive Behaviour Therapist, 9*, Article e2. https://doi.org/10.1017/S1754470X15000756

Raudenbush, S., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). SAGE.

Raue, P. J., McGovern, A. R., Kiosses, D. N., & Sirey, J. A. (2017). Advances in psychotherapy for depressed older adults. *Current Psychiatry Reports, 19*(9), Article 57. https://doi.org/10.1007/s11920-017-0812-8

R Core Team. (2018). *R: A language and environment for statistical computing* [Computer software]. https://www.r-project.org/

Ryan, R. M., Lynch, M. F., Vansteenkiste, M., & Deci, E. L. (2011). Motivation and autonomy in counseling, psychotherapy, and behavior change: A look at theory and practice. *The Counseling Psychologist, 39*(2), 193–260. https://doi.org/10.1177/0011000009359313

Sales, C. M. D., Ashworth, M., Ayis, S., Barkham, M., Edbrooke-Childs, J., Faísca, L., Jacob, J., Xu, D., & Cooper, M. (2023). Idiographic patient reported outcome measures (I-PROMs) for routine outcome monitoring in psychological therapies: Position paper. *Journal of Clinical Psychology, 79*(3), 596–621. https://doi.org/10.1002/jclp.23319

Sales, C., Faísca, L., Ashworth, M., & Ayis, S. (2023). The psychometric properties of PSYCHLOPS, an individualized patient-reported outcome measure of personal distress. *Journal of Clinical Psychology, 79*(3), 622–640. https://doi.org/10.1002/jclp.23278

Sawrikar, V., Newton, C., & Nicklas, L. (2023). What do patients want from therapy? Understanding treatment goals of patients with long-term conditions referred for cognitive behavioural therapy in primary care. *Cognitive Behaviour Therapist, 16*, Article e32. https://doi.org/10.1017/S1754470X23000272

Saxon, D., Barkham, M., Foster, A., & Parry, G. (2017). The contribution of therapist effects to patient dropout and deterioration in the psychological therapies. *Clinical Psychology & Psychotherapy, 24*(3), 575–588. https://doi.org/10.1002/cpp.2028

Schmidt, J., Lamprecht, F., & Wittmann, W. W. (1989). Zufriedenheit mit der stationären Versorgung. Entwicklung eines Fragebogens und erste Validitätsuntersuchungen [Satisfaction with inpatient care: Development of a questionnaire and initial validity studies]. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 39*, 248–255.

Schöttke, H., Trame, L., & Sembill, A. (2014). Relevance of therapy goals in outpatient cognitive-behavioral and psychodynamic psychotherapy. *Psychotherapy Research, 24*(6), 711–723. https://doi.org/10.1080/10503307.2014.881578

Silberschatz, G. (2017). Improving the yield of psychotherapy research. *Psychotherapy Research: Journal of the Society for Psychotherapy Research, 27*(1), 1–13. https://doi.org/10.1080/10503307.2015.1076202

Sittler, M. C., Lechner-Meichsner, F., Wilz, G., & Kessler, E. (2022). Does age matter? Initial treatment goals of older adults with major depression in outpatient cognitive behavioural therapy. *Clinical Psychology & Psychotherapy, 29*(2), 554–566. https://doi.org/10.1002/cpp.2646

Steinbrenner, T., Lalk, C., Kabjesz, A., Ferizaj, D., Pena Loray, J. S., Iovoli, F., & Rubel, J. (2025). Exploring negative experiences in psychotherapy using an NLP approach on online forum data. *Npj Mental Health Research, 4*(1), Article 59. https://doi.org/10.1038/s44184-025-00172-4

Trachsel, M., Itten, S., Stauffer, B., Grosse Holtforth, M., & Hofer, D. (2008). Erreichung störungsübergreifender Therapieziele in einer störungsspezifischen ambulanten Gruppenpsychotherapie phobischer Patienten [Attainment of generic therapy goals in a specialized group psychotherapy for phobic outpatients]. *PPmP - Psychotherapie · Psychosomatik · Medizinische Psychologie, 58*(09–10), 359–365. https://doi.org/10.1055/s-2007-986292

Tryon, G. S. (2018). Goals and psychotherapy research. In M. Cooper & D. Law (Eds.), *Working with goals in psychotherapy and counselling* (pp. 87–110). Oxford University Press. https://doi.org/10.1093/med-psych/9780198793687.003.0005

Tryon, G. S., & Winograd, G. (2011). Goal consensus and collaboration. *Psychotherapy, 48*(1), 50–57. https://doi.org/10.1037/a0022061

Vathke, C., Kessler, E.-M., Gellert, P., O'Sullivan, J. L., & Lech, S. (2023). Motivation and psychotherapy in older adults with care needs and depression. *GeroPsych, 36*(2), 49–62. https://doi.org/10.1024/1662-9647/a000305

Wampold, B. E., & Owen, J. (2021). Therapist effects: History, methods, magnitude, and characteristics of effective therapists. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change: 50th anniversary edition* (pp. 297–326). John Wiley & Sons.

Wong, G. Y., & Mason, W. M. (1985). The hierarchical logistic regression model for multilevel analysis. *Journal of the American Statistical Association, 80*(391), 513–524. https://doi.org/10.1080/01621459.1985.10478148

World Health Organization. (2004). *International statistical classification of diseases and related health problems: Alphabetical index* (Vol. 3). https://iris.who.int/handle/10665/42980

Zavlis, O., Fonagy, P., & Luyten, P. (2025). The most important aims of psychotherapy: To love, to work, and to find meaning. *The Lancet Psychiatry, 12*(3), 173–174. https://doi.org/10.1016/S2215-0366(25)00006-9

Zimmermann, D., Rubel, J., Page, A. C., & Lutz, W. (2017). Therapist effects on and predictors of non-consensual dropout in psychotherapy. *Clinical Psychology & Psychotherapy, 24*(2), 312–321. https://doi.org/10.1002/cpp.2022

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

Online Application of Imagery Rescripting for Pathological Affective Dependence (ImRs-PAD) in Survivors of Intimate Partner Violence: Results From a Multiple Case Series

Erica Pugliese[1,2], Allison Uvelli[3], Arnold A. P. van Emmerik[1], Arnoud Arntz[1]

[1] *Clinical Psychology Group, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands.* **[2]** *School of Cognitive Psychology, Rome, Italy.* **[3]** *Department of Medical Science, Surgery and Neurosciences, University of Siena, Siena, Italy.*

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e22027, https://doi.org/10.32872/cpe.22027

Received: 2025-09-10 • **Accepted:** 2026-05-14 • **Published (VoR):** 2026-08-31

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

Corresponding Author: Erica Pugliese, Department of Psychology, Clinical Psychology Group, University of Amsterdam, Nieuwe Achtergracht 129, 1001 NK Amsterdam, The Netherlands. +31 (0)20 525 6810. E-mail: e.pugliese@uva.nl

Supplementary Materials: Materials [see Index of Supplementary Materials]

Abstract

Background: This study examined the feasibility and preliminary clinical indications of an imagery rescripting intervention integrating the Pathological Affective Dependence framework (ImRs-PAD) in survivors of Intimate Partner Violence (IPV). Pathological Affective Dependence (PAD) refers to a trauma-related relational vulnerability associated with difficulties in disengaging from abusive intimate relationships.

Method: A multiple case series design was used with three female IPV survivors referred by a mental health center in Italy (mean age = 31.3 years, *SD* = 2.08). All participants reported histories of abusive relationships, marked difficulties with separation, and clinically significant levels of PAD. Participants received an online ImRs-PAD intervention. Repeated self-report measures assessed changes in PAD, trauma-related symptoms, depression, anxiety, somatic symptoms, resilience, self-compassion, and psychological well-being.

Results: Across cases, reductions were observed in PAD, depressive and anxiety symptoms, and trauma-related and somatic complaints, alongside increases in resilience, self-compassion, and overall psychological well-being. The online delivery of the intervention was feasible, with no technical or safety issues reported.

Conclusion: The findings provide preliminary support for the potential clinical relevance of ImRs-PAD for IPV survivors. Addressing PAD may facilitate disengagement from violent partners and reduce the risk of re-exposure to severe forms of violence.

Keywords

intimate partner violence (IPV), pathological affective dependence (PAD), relational trauma, imagery rescripting (ImRs), mechanisms of change in ImRs, psychotherapy

Highlights

- Online imagery rescripting was feasible and safe for IPV survivors with pathological affective dependence.
- The intervention was associated with reductions in pathological affective dependence and trauma-related symptoms.
- Improvements were observed in self-compassion, resilience, and psychological well-being.
- Addressing unmet relational needs may support disengagement from abusive relationships.

Intimate Partner Violence (IPV) represents a pervasive public health issue with profound psychological, physical, and social consequences, particularly for women (Jewkes & Abrahams, 2026). The World Health Organization (WHO, 2024) defines IPV as harmful behaviors by a current or former intimate partner, including physical aggression, sexual coercion, psychological abuse, and controlling behaviors. Remaining in abusive relationships is influenced by structural and individual obstacles, including economic, social, and cultural pressures, as well as the perpetrator's psychological characteristics (i.e., anger problems, anxiety, depression, suicidal behaviour, personality disorders, alcoholism, or problem gambling; Cuesta-García & Crespo, 2022; Neal & Edwards, 2017). Over time, increasing attention has also been directed toward the relational dynamics between victims and perpetrators of IPV (Pugliese et al., 2024). Within this context, Pathological Affective Dependence (PAD) has been proposed as an emerging clinical construct describing a vulnerability that may increase the likelihood of remaining in or returning to abusive relationships (Castillo-Gonzáles et al., 2024; Pugliese et al., 2023). PAD is a relational condition characterized by a persistent and intense need to maintain a specific intimate relationship despite its dysfunctional nature, including the presence of psychological distress or interpersonal violence (Pugliese et al., 2023). Individuals with PAD remain trapped in these dangerous relationships as they experience an internal conflict: the compulsion to save the relationship, even at the expense of their own well-being, and the desire to protect themselves from further harm. Two conditions typically fuel this internal struggle: partner abuse and perceived inability to leave the relationship. PAD

encompasses both conditions and the resulting internal struggle (Pugliese, Uvelli, et al., 2025).

PAD can be conceptualized as both a trait—reflecting enduring dysfunctional relational patterns—and a state, referring to the situational activation of these patterns within a specific relationship, particularly in abusive or violent contexts (Pugliese, Uvelli, et al., 2025). In addition, PAD appears to be associated with reduced cognitive and metacognitive functioning in IPV survivors, which may partly explain difficulties in help-seeking and relational disengagement (Pugliese, Papa, et al., 2025). According to the PAD theory (Pugliese et al., 2023), the etiology of PAD lies in the chronic frustration of at least one of three core relational needs—love (acceptance and care), dignity (being respected and valued), and safety (physical and emotional security)—within early caregiving relationships (Papa & Pugliese, 2025; Papa et al., 2026; Smith-Marek et al., 2015; Speranza et al., 2022). In this framework, relational trauma is conceptualized as the cumulative impact of persistent unmet relational needs. In adulthood, these experiences may consolidate into trait PAD, understood as a stable pattern of dysfunctional relational functioning. Individuals with such histories often engage in intimate relationships with an implicit reparative aim, seeking to fulfill unmet developmental needs (Pezzoli et al., 2025; Phillips et al., 2025). However, these efforts are frequently directed toward relational contexts that resemble early caregiving environments, thereby reactivating and reinforcing maladaptive interpersonal patterns (Pugliese et al., 2023). Within this process, current relationships may elicit state PAD, defined as the situational manifestation of these dynamics within a specific relational context. This condition is typically characterized by ongoing abuse, a perceived inability to disengage, and a persistent conflict between separation and relationship maintenance. Such dynamics contribute to both the maintenance of abusive relationships and the perpetuation of maladaptive relational cycles. From this perspective, PAD may be understood as the enduring relational imprint of repeated relational trauma, predisposing individuals to recurrent patterns of dysfunction and abuse (Silvestri et al., 2025; Speranza et al., 2022). Despite its clinical significance, there remains a notable lack of interventions specifically designed to address psychological mechanisms rooted in early trauma that increase vulnerability to revictimization—mechanisms that are central to PAD.

Study Rationale

Imagery Rescripting (ImRs) is one promising evidence-based intervention to address the unmet needs central to PAD. This experiential technique, originally developed within Schema Therapy (Arntz & Weertman, 1999), allows clients to revisit early aversive memories, fantasies, or nightmares and transform their emotional meaning (Müller-Engelmann & Steil, 2017; Nilsson et al., 2019). One of the hypothesized mechanisms of change in ImRs is the fulfillment of unmet needs (Koetsier et al., 2024). To our knowledge, however, no study has examined whether satisfying unmet needs from childhood

and adulthood using guided imagery may be beneficial for survivors of IPV presenting with PAD (for a review, see Kip et al., 2023; Kroener et al., 2023). Specifically, it remains unclear whether, after these needs are addressed in imagination, survivors report reductions in PAD. Additionally, previous research on treatments for women IPV survivors has largely focused on PTSD, depression, anxiety, and self-esteem (Emirza & Uzun, 2024), often overlooking complex PTSD (ICD-11, 2018) symptoms—namely disturbances in self-organization, including affective dysregulation, negative self-concept, and relational difficulties—linked to ongoing trauma (Cloitre et al., 2018), as well as positive outcomes such as resilience and self-compassion. ImRs-PAD is expected to foster a more supportive and protective stance toward the vulnerable self, potentially increasing self-compassion. This is particularly relevant in survivors of IPV, who often experience high levels of self-criticism for not leaving the relationship earlier, despite recognizing its dysfunctional or abusive nature (Crapolicchio et al., 2021). Such self-criticism may be further reinforced by self-blame processes, as perpetrators frequently attribute responsibility for the violence to the victim (Naismith et al., 2024). This is consistent with the PAD construct, particularly the internal conflict dimension, which reflects the struggle between recognizing that the relationship is harmful and feeling unable to separate from the partner (Pugliese, Uvelli, et al., 2025). Moreover, despite substantial therapeutic and support efforts, many IPV survivors remain resistant to change within existing care systems (Paphitis et al., 2022). Existing evidence-based interventions for IPV survivors primarily target symptom reduction (e.g., PTSD, depression, anxiety) and may not directly address the underlying relational mechanisms that maintain dependence on abusive partners. These approaches may be less effective in modifying unmet relational needs and the associated internal conflict that characterizes PAD. The present study aims to address this gap by framing PAD as a relational pathway through which traumatic family and partner experiences may contribute to persistent difficulties in disengaging from abusive relationships. Addressing PAD may therefore represent a trauma-informed pathway to reduce re-exposure to violence, complementing symptom-focused PTSD treatments.

The Current Study

To address this gap, the present study explores the feasibility and preliminary effectiveness of an online imagery rescripting intervention (ImRs-PAD) tailored for female survivors of IPV with clinically meaningful PAD symptoms. Using an exploratory multiple case series design, the intervention specifically targets three core unmet needs—love, dignity, and safety—identified from participants' autobiographical narratives and relational histories. In addition to evaluating the feasibility of online delivery, the study examines whether ImRs-PAD reduces PAD, complex post-traumatic stress, depression, anxiety, and somatic complaints while enhancing psychological well-being, self-compassion, and resilience. Online access is particularly relevant for IPV survivors who often face barriers to in-person care due to ongoing partner control or surveillance (van Gelder et al.,

2023). The present study is an independent online multiple case series conducted in Italy, designed to provide a preliminary evaluation of the protocol prior to the implementation of a preregistered in-person multiple baseline study trial (ClinicalTrials.gov Identifier: NCT06670326). The preregistered trial is conducted in collaboration with several anti-violence centers and is currently ongoing in the Netherlands.

Hypotheses

We formulated a set of hypotheses organized into primary and secondary outcomes, designed to capture the direct effects of the intervention at different stages of the protocol.

- *Primary Outcomes:* Participants will show significant reductions in both state PAD and trait PAD from pre-treatment to post-treatment, with effects maintained at follow-up.
- *Secondary Outcomes:* Participants will show significant increases in resilience, self-compassion, and psychological well-being, as well as reductions in depression, anxiety, and complex post-traumatic stress symptoms from pre-treatment to post-treatment, with improvements sustained or enhanced at follow-up.

Method

Participants

Three female participants were referred by a psychiatrist from a mental health center in Italy to participate in the current study. All presented with a history of abusive relationships with current or former partners and reported marked difficulties in achieving or maintaining separation despite recognizing the need to do so. The mean age was 31.3 years (SD = 2.08). All participants were unmarried at the time of treatment and reported recurrent difficulties in maintaining stable and autonomous romantic relationships. They also showed a marked inability to leave dysfunctional relationships and experienced exacerbations of depressive and anxiety symptoms during separation phases. All patients referred during the study period who met the inclusion criteria and provided informed consent were included. Inclusion criteria were sufficient fluency in Italian to complete the study procedures, age between 18 and 60 years, and clinically meaningful levels of PAD, as indicated by a score of 47 or higher on the Pathological Affective Dependence Scale – State and Trait versions (PADS-S, PADS-T; Pugliese, Uvelli, et al., 2025). Exclusion criteria were assessed through a comprehensive clinical interview by a licensed psychologist and psychotherapist with 10 years of clinical experience. These included the evaluation of current and past psychiatric conditions, substance use, medical history, and ongoing treatments. The assessment aimed to rule out the presence of psychotic disorders, severe substance use disorder, intellectual disability, debilitating chronic medical conditions, prior receipt of imagery rescripting–based therapy, and current psychophar-

macological treatment. Detailed descriptions of individual cases included in this multiple case series are reported in Appendix A.

Procedures

The intervention was conducted by two licensed clinical psychologists, both formally trained in cognitive-behavioral therapy (CBT) for PAD, and engaged in peer supervision throughout the treatment process. One therapist, with over 4 years of experience in both clinical practice and research, delivered the treatment to all three participants. The second therapist, with over 10 years of experience in both clinical practice and research, provided clinical supervision and administered the assessment measures. Three clients participated in a 5-month program consisting of weekly individual sessions (approximately 20 sessions each, including 4–6 assessment sessions, 13 treatment sessions, and the remaining sessions focused on follow-up).

All treatment sessions were conducted online via Google Meet using secure, individualized access links. For each session, a unique private link was generated and sent to the participant's personal email address the day before the appointment. Access to the session was restricted to the invited participant, ensuring confidentiality and controlled access. A different link was used for each session and for each participant.

Before the intervention, all clients underwent a comprehensive psychological assessment. All self-report measures were completed independently by participants outside the therapy sessions. Therapists did not use questionnaire scores to guide the intervention and had no access to interim results during treatment. The treatment program started in March 2024 and concluded in August 2025. All participants provided informed consent, and the study was approved by the Ethics Committee of the Scuola di Psicoterapia Cognitiva (protocol number N. Pr 2/24).

The treatment component, ImRs-PAD, is a trauma-informed intervention based on Imagery Rescripting (ImRs), consisting of 13 weekly individual sessions (50 minutes each), including one psychoeducation session followed by 12 ImRs sessions. Grounded in PAD theory (Pugliese et al., 2023), ImRs-PAD aims to address chronic frustration of core relational needs for love, dignity, and safety across early family experiences, abusive intimate relationships, and anticipated future relational challenges. Across stages, responsibility for rescripting progressively shifts from the therapist to the client, fostering autonomy, affect regulation, and relational agency. ImRs-PAD is organized into three stages, each specifically targeting the reduction of PAD traits, PAD state, and relapse.

Measures

The Pathological Affective Dependence Scale (PADS; Pugliese, Uvelli, et al., 2025) is a 17-item self-report questionnaire available in two parallel forms that measure state (PADS-S) and trait (PADS-T) affective dependence, respectively. It assesses three dimensions of

pathological dependence, including internal conflict (IC), inability to separate from the partner (IS), and partner abuse (PA). Items are rated on a 5-point Likert scale (1 = *not at all*, 5 = *always*). The total score ranges from 17 to 85, with higher scores indicating greater levels of affective dependence. In the original validation study (Pugliese, Uvelli, et al., 2025), Cronbach's α was .89 for both the state and trait versions. Example items include *"I am aware that I am suffering in this relationship, but at the same time, I cannot leave it"* (IC); *"I would do anything in order not to lose my importance in my partner's sight"* (IS); and *"My partner devalues me"* (PA). In this study, the PADS was administered at baseline (T0), after every session (T1-T12), and at follow-up.

The Self-Compassion Scale (SCS; Neff, 2003) is a 26-item questionnaire designed to measure six dimensions of self-compassion, including self-kindness, self-judgment, common humanity, isolation, mindfulness, and over-identification. Each item is rated on a 5-point Likert scale (1 = *almost never*, 5 = *almost always*). The total score ranges from 26 to 130, with higher scores reflecting greater self-compassion. To obtain the scores for the subscales, the mean of the items belonging to each subscale should be calculated. Cronbach's α in the original study was .92 for the total score and between .75 and .81 for the subscales. Example items include *"I try to be loving toward myself when I'm feeling emotional pain"* and *"I'm disapproving and judgmental about my own flaws and inadequacies."* In this study, the SCS was administered at baseline (T0), after 6 sessions (T6), after 12 sessions (T12), and at follow-up.

The Patient Health Questionnaire (PHQ; Kroenke et al., 2001, 2002; Spitzer et al., 2006) is a widely used self-report instrument that screens for somatic, anxiety, and depressive symptoms. It consists of 31 items across different modules, each rated on a 4-point Likert scale (0 = *not at all*, 3 = *nearly every day*). The score range differs for each module (e.g., PHQ-9 depression: 0–27; PHQ-15 somatic symptoms: 0–45; GAD-7 anxiety: 0–21). Higher scores indicate greater symptom severity. Cronbach's α values in validation studies ranged from .86 to .89 (Kroenke et al., 2001, 2002; Spitzer et al., 2006). Example items include *"Over the last 2 weeks, how often have you had little interest or pleasure in doing things?"* and *"Over the last 2 weeks, how often have you felt nervous, anxious, or on edge?"* In this study, the PHQ was administered at T0, T6, T12, and follow-up.

The Ego-Resiliency Scale – Revised (ER89-R; Alessandri et al., 2007) is a 10-item self-report questionnaire designed to assess psychological resilience, particularly the ability to adapt to changing circumstances flexibly. Items are rated on a 4-point Likert scale (1 = *does not apply at all*, 4 = *applies very strongly*). The total score ranges from 10 to 40, with higher scores indicating greater ego-resiliency. Cronbach's α in the Italian validation was .76 (Alessandri et al., 2007). Example items include: *"I quickly get over and recover from being startled"* and *"I enjoy dealing with new and unusual situations."* The ER89-R was administered at T0, T6, T12, and follow-up.

The International Trauma Questionnaire (ITQ; Cloitre et al., 2018) is an 18-item measure assessing symptoms of post-traumatic stress disorder (PTSD) and complex PTSD (CPTSD) according to ICD-11 criteria. Each item is scored on a 5-point Likert scale (0 = *not at all*, 4 = *extremely*). The total score ranges from 0 to 72, with higher scores indicating greater symptom severity. Cronbach's α in the original validation study (Cloitre et al., 2018) ranged from .88 to .94 across the different subscales. Example items include: *"When I am reminded of the trauma, I get intense physical reactions (for example, sweating, heart beating fast, nausea)"* and *"I feel worthless because of the trauma."* In this study, the ITQ was administered at T0, T6, T12, and follow-up.

The Psychological General Well-Being Index – Short (PGWB-S; Grossi et al., 2006) is a 6-item self-report scale measuring overall subjective well-being. Each item is rated on a 6-point Likert scale (0 = *none of the time*, 5 = *all of the time*), with total scores ranging from 0 to 30. Higher scores indicate better psychological well-being. Cronbach's α in validation studies (Grossi et al., 2006) ranged from .85 to .90. Example items include *"Have you been bothered by nervousness or your 'nerves' during the past month?"* and *"How often have you felt cheerful and light-hearted?"* The PGWB-SF was administered at T0, T6, T12, and follow-up.

The Young Schema Questionnaire – Short Form (YSQ-SF; Young & Brown, 2005) is a 90-item self-report questionnaire designed to assess 18 early maladaptive schemas. Items are rated on a 6-point Likert scale (1 = *completely untrue of me*, 6 = *describes me perfectly*). The total score ranges from 90 to 540, with higher scores reflecting stronger schema endorsement. Cronbach's α for the subscales in the original validation study (Young & Brown, 2005) ranged from .83 to .96. Example items include: *"I worry that people I feel close to will leave me or abandon me"* and *"I feel that people will take advantage of me."* In this study, the YSQ-SF was administered only at baseline (T0).

All self-report measures were administered using standardized procedures to minimize response bias and social desirability effects. Questionnaires were completed independently by participants online via a secure survey platform. Measures were administered and collected by a research team member not involved in therapy, and therapists had no access to questionnaire responses at any assessment point, ensuring a clear separation between clinical and research roles.

Statistical Analysis

Data analysis was conducted in two steps. First, pre- (T0) and post-treatment (T12) scores were compared for each case using the Reliable Change Index (RCI; Jacobson & Truax, 1991), which expresses the ratio between the observed change and the standard error of Italian measurement. Values greater than 1.96 were considered statistically significant, indicating a reliable change. The RCI was selected because it allows the evaluation of clinically meaningful changes at the individual level, which is particularly appropriate for case-series designs in small samples. Second, differences were compared using the

RCI across intermediate and follow-up phases (T0–T6, T6–T12, T12–FU, and T0–FU), to capture both short-term and long-term changes. All analyses were performed at the individual level, in line with the descriptive and exploratory nature of case-series designs.

Results

Across all three cases, results showed consistent improvements on every measure from baseline (T0) to follow-up (FU). Specifically, both the state and trait versions of the PADS revealed marked reductions, with final scores approaching the minimum possible value for each participant. In parallel, self-compassion (SCS) increased steadily, with Case 1 reaching the highest overall levels at follow-up. Depressive, somatic, and anxiety symptoms measured with the PHQ decreased substantially in all participants. At the same time, psychological well-being (PGWB) showed a progressive rise, particularly in Case 1 and Case 3, where scores nearly doubled or tripled relative to baseline. Ego-resiliency (ER-89-R) also increased across cases, reflecting greater adaptive functioning over time. Finally, trauma-related symptoms (ITQ) declined sharply, with Case 1 and Case 2 showing the fastest early reductions, while Case 3 improved more gradually but reached substantially lower symptom levels at follow-up. Overall, the data highlight a consistent pattern of symptom reduction and functional improvement across all domains, with convergence of outcomes among the three clients despite some differences in the pace of change (see Table 1).

Table 1

Pre-Treatment (T0), Half-Treatment (T6), Post-Treatment (T12), and Follow-Up (FU) Scores of Participants on All Outcome Measures

Variable	Case 1 T0-T6-T12-FU	Case 2 T0-T6-T12-FU	Case 3 T0-T6-T12-FU
PADS-S	69-30-18-17	71-64-39-20	68-33-22-17
PADS-T	75-38-30-18	72-66-39-22	61-32-22-17
SCS	1.53-2.86-4.08-4.50	2.30-2.50-3.05-3.80	1.66-2.23-2.46-2.86
PHQ	49-25-22-10	25-19-13-8	31-23-20-10
PGWB	47.58-91.5-106.14-109.8	43.92-47.58-62.22-75.13	14.64-25.62-30.22-41.14
ER-89-R	38-46-48-50	37-43-44-52	30-40-42-45
ITQ	59-16-11-6	51-34-18-10	59-41-33-20

Note. T0 = pre-treatment; T6 = half-treatment; T12 = post-treatment; FU = follow-up; PADS-S = Pathological Affective Dependence Scale-State version; PADS-T = Pathological Affective Dependence Scale-Trait version; SCS = Self-Compassion Scale; PHQ = Patient Health Questionnaire; PGWB = Psychological General Well-Being Index; ER-89-R = Ego-Resiliency Scale-Revised; ITQ = International Trauma Questionnaire.

All three participants demonstrated statistically reliable improvements across the outcome measures, with differences in timing and magnitude (Table 2). A concise overview of the temporal pattern of reliable changes across outcome measures is reported in Table 3.

Table 2

Reliable Change Index (RCI) of the Outcome Measures Over Time

Variable	Case 1 T0-T12 / T0-FU / T0-T6 / T6-12 / T12-FU	Case 2 T0-T12 / T0-FU / T0-T6 / T6-12 / T12-FU	Case 3 T0-T12 / T0-FU / T0-T6 / T6-12 / T12-FU
PADS-S	8.79* / 8.96* / 6.72* / 2.07* / 0.17	5.52* / 8.79* / 1.21 / 4.31* / 3.27*	7.93* / 8.79* / 0.69 / 1.90 / 0.86
PADS-T	7.66* / 9.70* / 6.30* / 1.36 / 2.04*	5.62* / 8.51* / 1.02 / 4.59* / 2.89*	6.64* / 7.49* / 4.93* / 1.70 / 0.85
SCS	-7.08* / -9.00* / -3.96* / -3.70* / -1.27	-2.08* / -4.55* / -0.56 / -1.67 / -2.27*	-2.22* / -3.64* / -1.58 / -0.70 / -1.21
PHQ	6.53* / 9.43* / 5.80* / 0.73 / 2.90*	2.90* / 4.11* / 1.45 / 1.45 / 1.21	2.66* / 5.08* / 1.93 / 0.73 / 2.42*
PGWB	-19.91* / -21.16* / -14.93* / -4.98* / -1.24	-6.22* / -10.61* / -1.24 / -4.98* / -4.39*	-5.30* / -9.01* / -3.73* / -1.56 / -3.71*
ER-89-R	-8.37* / -10.04* / -6.69* / -1.67 / -1.67	-5.86* / -12.55* / -5.02* / -0.84 / -6.69*	-10.04* / -12.55* / -8.37* / -1.67 / -2.51*
ITQ	11.87* / 13.11* / 10.63* / 1.24 / 1.24	8.16* / 10.14* / 4.20* / 3.96* / 1.98*	6.43* / 9.64* / 4.45* / 1.98* / 3.21*

Note. PADS-S = pathological affective dependence scale-state version; PADS-T = pathological affective dependence-trait version; SCS = self-compassion scale; PHQ = patient health questionnaire; PGWB = psychological general well-being index; ER-89-R = ego-resiliency-89-revised; ITQ = international trauma questionnaire. T0–T12 = comparison between pre-treatment and the last treatment session; T0–FU = comparison between pre-treatment and follow-up; T0–T6 = comparison between pre-treatment and mid-treatment (after the sixth imagery with rescripting session); T6–T12 = comparison between mid-treatment and the last treatment session; T12–FU = comparison between the last treatment session and follow-up.
*Significant results at |RCI| > 1.96.

Table 3

Summary of Temporal RCI Changes Across Assessment Phases

Variable	Case 1 T0-T12 / T0-T6 / T6-T12 / T12-FU / T0-FU	Case 2 T0-T12 / T0-T6 / T6-T12 / T12-FU / T0-FU	Case 3 T0-T12 / T0-T6 / T6-T12 / T12-FU / T0-FU
PADS-S	8.79* / 6.72* / 2.07* / 0.17 / 8.96*	5.52* / 1.21 / 4.31* / 3.27* / 8.79*	7.93* / 0.69 / 1.90 / 0.86 / 8.79*
PADS-T	7.66* / 6.30* / 1.36 / 2.04* / 9.70*	5.62* / 1.02 / 4.59* / 2.89* / 8.51*	6.64* / 4.93* / 1.70 / 0.85 / 7.49*
SCS	-7.08* / -3.96* / -3.70* / -1.27 / -9.00*	-2.08* / -0.56 / -1.67 / -2.27* / -4.55*	-2.22* / -1.58 / -0.70 / -1.21 / -3.64*
PHQ	6.53* / 5.80* / 0.73 / 2.90* / 9.43*	2.90* / 1.45 / 1.45 / 1.21 / 4.11*	2.66* / 1.93 / 0.73 / 2.42* / 5.08*
PGWB	-19.91* / -14.93* / -4.98* / -1.24 / -21.16*	-6.22* / -1.24 / -4.98* / -4.39* / -10.61*	-5.30* / -3.73* / -1.56 / -3.71* / -9.01*
ER-89-R	-8.37* / -6.69* / -1.67 / -1.67 / -10.04*	-5.86* / -5.02* / -0.84 / -6.69* / -12.55*	-10.04* / -8.37* / -1.67 / -2.51* / -12.55*
ITQ	11.87* / 10.63* / 1.24 / 1.24 / 13.11*	8.16* / 4.20* / 3.96* / 1.98* / 10.14*	6.43* / 4.45* / 1.98* / 3.21* / 9.64*

Note. RCI = reliable change index; PADS-S = Pathological Affective Dependence Scale-State version; PADS-T = Pathological Affective Dependence Scale-Trait version; SCS = Self-Compassion Scale; PHQ = Patient Health Questionnaire; PGWB = Psychological General Well-Being Index; ER-89-R = Ego-Resiliency Scale-Revised; ITQ = International Trauma Questionnaire.
*Significant results at |RCI| > 1.96.

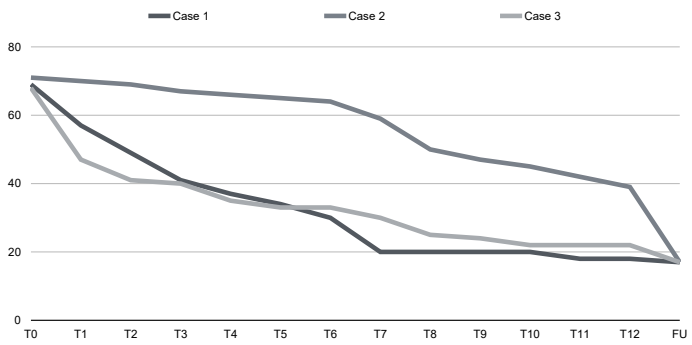
Case 1 showed the most rapid and extensive response. Substantial improvements were evident already at mid-treatment, with reliable reductions in pathological affective dependence, general psychopathology, and trauma-related symptoms, as well as increases in well-being, self-compassion, and resilience. These gains were largely maintained and consolidated through post-treatment and follow-up. Case 2 exhibited a more gradual trajectory. Early changes were limited, but significant progress emerged between mid- and post-treatment, particularly in affective dependence, trauma symptoms, and psychological well-being. Additional gains were observed at follow-up, including further

improvements in resilience. Case 3 also showed a slower but consistent pattern of change. Improvements were selective at mid-treatment, mainly in resilience and affective dependence, and became more evident by post-treatment. Further progress occurred at follow-up, especially in well-being and trauma-related symptoms. Taken together, the results indicate that while the three participants followed different trajectories, they all achieved clinically meaningful improvements by the end of treatment and maintained or enhanced these gains at follow-up.

Figures 1, 2, 3, 4, 5, 6, 7 illustrate the longitudinal trajectories of the main outcome measures from baseline to follow-up. Overall, visual inspection showed a consistent reduction in PADS-S and PADS-T scores across the three cases, with particularly marked improvements between baseline and post-treatment that were generally maintained or further improved at follow-up. A similar pattern emerged for ITQ and PHQ scores, which decreased over time in all participants. In parallel, SCS, PGWB, and ER-89-R scores increased across the intervention and follow-up phases. Although the timing and magnitude of change varied across cases and outcome domains, the overall pattern indicated improvement across both symptom-related and resilience-related outcomes. These trajectories are presented descriptively and should be interpreted in relation to the individual case patterns.

Figure 1

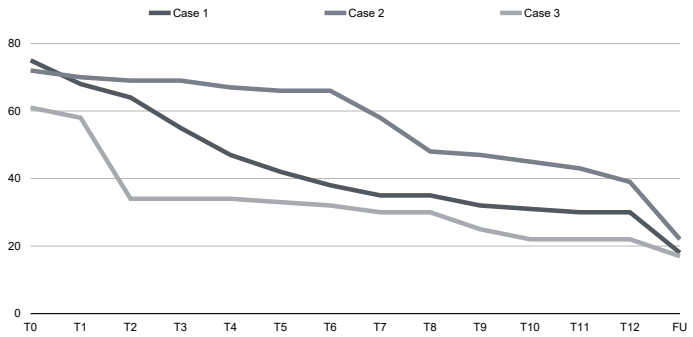
Time Series of State Pathological Affective Dependence: Changes in the Total Score of PADS-S



Note. PADS-S = Pathological Affective Dependence Scale-State version; T0 = pre-treatment; T6 = half-treatment; T12 = post-treatment; FU = follow-up. Higher scores indicate higher state pathological affective dependence.

Figure 2

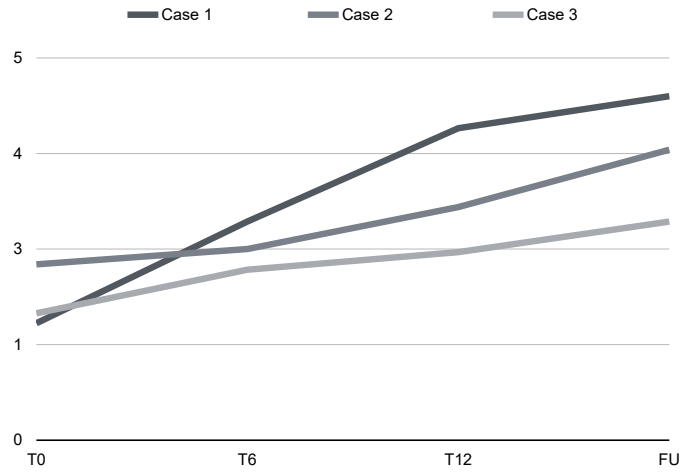
Time Series of Trait Pathological Affective Dependence: Changes in the Total Score of PADS-T



Note. PADS-T = Pathological Affective Dependence Scale-Trait version; T0 = pre-treatment; T6 = half-treatment; T12 = post-treatment; FU = follow-up. Higher scores indicate higher trait pathological affective dependence.

Figure 3

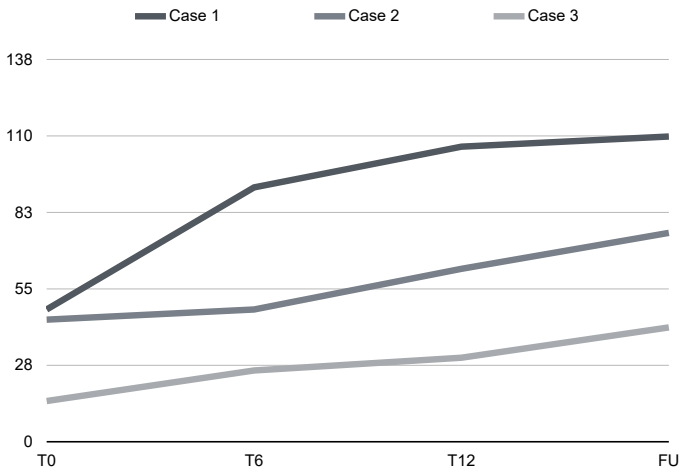
Self-Compassion Score Over Time: Changes in the Total Score of SCS



Note. SCS = Self-Compassion Scale; T0 = pre-treatment; T6 = half-treatment; T12 = post-treatment; FU = follow-up. Higher scores indicate higher self-compassion.

Figure 4

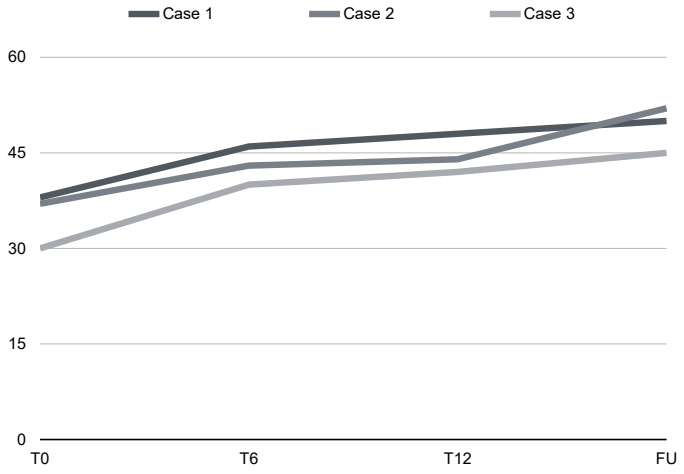
Psychological Well-Being Over Time: Changes in the Total Score of PGWB



Note. PGWB = Psychological General Well-Being Index; T0 = pre-treatment; T6 = half-treatment; T12 = post-treatment; FU = follow-up. Higher scores indicate higher psychological well-being.

Figure 5

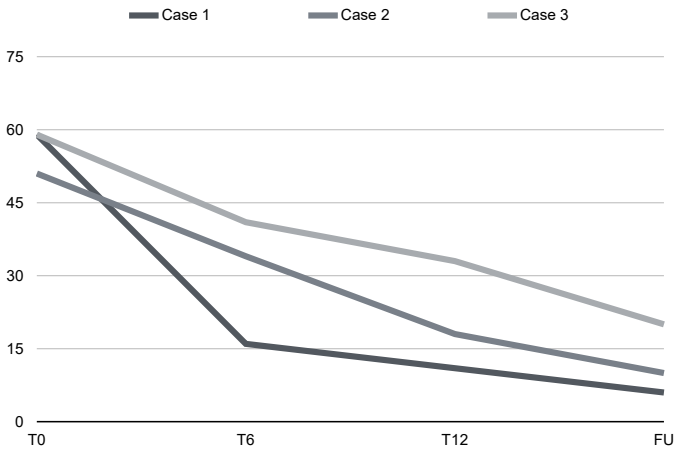
Resilience Over Time: Changes in the Total Score of ER89-R



Note. ER-89-R = Ego-Resiliency Scale-Revised; T0 = pre-treatment; T6 = half-treatment; T12 = post-treatment; FU = follow-up. Higher scores indicate higher ego-resiliency.

Figure 6

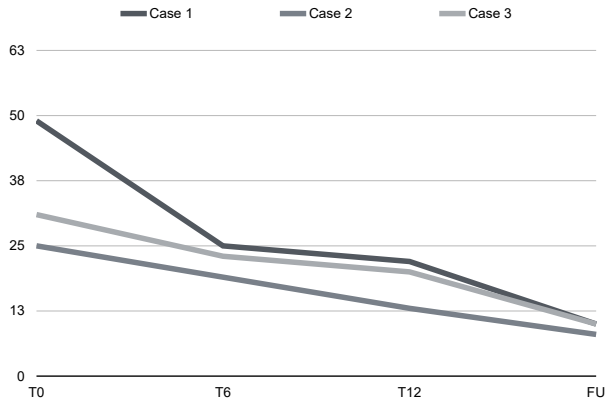
Trauma Symptoms Over Time: Changes in the Total Score of ITQ



Note. ITQ = International Trauma Questionnaire; T0 = pre-treatment; T6 = half-treatment; T12 = post-treatment; FU = follow-up. Higher scores indicate higher trauma-related symptoms.

Figure 7

Depression, Anxiety, and Somatic Symptoms Over Time: Changes in the Total Score of PHQ



Note. PHQ = Patient Health Questionnaire; T0 = pre-treatment; T6 = half-treatment; T12 = post-treatment; FU = follow-up. Higher scores indicate higher depression, anxiety, and somatic symptom severity.

Discussion

To our knowledge, this is the first study to apply the online imagery rescripting (ImRs) to women with pathological affective dependence (PAD) who have survived intimate partner violence (IPV). Across cases, ImRs-PAD was associated with consistent reductions in PAD and other negative symptomatology (complex post-traumatic stress, depression, anxiety), alongside parallel improvements in resilience, self-compassion, and well-being from baseline to follow-up. Taken together, these findings provide preliminary support for the potential efficacy of ImRs in this population. Moreover, they offer initial indications regarding possible mechanisms of change in ImRs (see also Visco-Comandini et al., 2025). Specifically, our results tentatively suggest that addressing unmet relational needs in imagination may contribute to reductions in PAD and related psychological distress, a condition theorized to originate from the chronic frustration of core needs within both the family of origin and abusive partner relationships (Pugliese, Uvelli, et al., 2025).

Evidence from the present study suggests that early maladaptive schemas, as assessed by the YSQ (Young & Brown, 2005), may play a role in shaping individual responses to the intervention (Uvelli et al., 2025). Participants with a predominant self-sacrifice schema appeared to show slower improvement. In contrast, those with more trauma-related schemas (e.g., abandonment, vulnerability) showed earlier reductions in distress following imagery rescripting. Participants with prominent shame/defectiveness schemas showed an intermediate pattern of change, suggesting that different schema profiles may affect both the speed and course of treatment response. Importantly, despite these differences, all participants eventually showed disengagement from abusive relationships. Although these findings are preliminary and based on a small sample, they highlight the potential role of schema-level processes in explaining differences in treatment outcomes.

Feasibility of Online Delivery

The online delivery of ImRs-PAD appeared acceptable and manageable for all participants. No major technical difficulties, safety concerns, or treatment interruptions were reported during the intervention. The online format may have facilitated engagement by reducing practical and relational barriers to care (i.e., mobility constraints, partner surveillance, and privacy concerns), common among IPV survivors (van Gelder et al., 2023). Although these observations are preliminary and based on a small number of cases, they suggest that online delivery of ImRs-PAD may represent a viable option for trauma-informed care in this population.

Timing of Therapeutic Change in ImRs-PAD

Improvements in primary and secondary outcomes did not emerge uniformly across participants. This variability may be explained by a combination of relational factors

(e.g., type of violence experienced: psychological and/or physical) and contextual factors (e.g., time since separation, ongoing contact with the abusive partner).

Relational Risk Factors: The Type of Violence

Differences in the type of violence (physical and/or psychological) experienced may influence the pace of therapeutic change (Jordan et al., 2010). Case 1 primarily reported psychological violence, such as devaluation, control, and verbal abuse, with no indication of severe physical aggression. This pattern may explain her relatively early improvements in PAD, depression, and trauma-related symptoms, as her therapeutic work was not compounded by the more pervasive bodily threat associated with physical violence. Also, Case 3 presented a history of severe psychological and verbal violence, but her therapeutic trajectory was additionally shaped by repeated re-entries into the abusive relationship. These returns prolonged distress and slowed down the stabilization process. However, once the imagery worked effectively, addressing both her fear of abandonment and need for safety, her progress became more consistent and sustainable. By contrast, Case 2 had experienced both psychological and physical violence, with recurrent episodes of intimidation and physical harm. This dual exposure may contribute to her slower therapeutic response compared to Case 1 and Case 3, as reflected in the delayed improvements in PAD and trauma-related symptoms. This pattern is consistent with evidence that combined forms of abuse are linked to more severe psychological consequences and greater treatment resistance (Iverson et al., 2011). Together, these findings suggest that the type of violence—psychological versus combined psychological and physical—may contribute to differential timing of therapeutic gains, with more severe and multifaceted abuse associated with slower change.

Contextual Risk Factors: The Separation From the Abusive Partner

Two contextual factors may influence treatment outcomes: (1) the time since the separation from the abusive partner occurred and (2) whether there is an ongoing or intermittent contact versus no current contact with the (ex) partner. Shorter time since last contact and ongoing or intermittent contact are typically linked to less favorable outcomes, including a higher risk of returning to the abusive relationship—a dynamic associated with escalating violence and even femicide (McFarlane et al., 2004). Similarly, Iverson et al. (2011) found that women with current partner contact were less likely to initiate PTSD treatment and showed poorer outcomes compared to those with past but not ongoing abuse. These findings suggest that continued exposure to IPV may undermine both treatment engagement and effectiveness. Our findings are consistent with this evidence. Case 1, who had fully disengaged from the abusive partner for several years, demonstrated the fastest and most stable improvements across outcomes. Case 2, separated for about a year but maintaining intermittent contact, showed the slower progression, with gains consolidating later in treatment. Case 3, who was separated more

recently and continued to have active contact with her abusive partner—including repeated returns to the relationship—displayed a mixed trajectory, with delayed stabilization of outcomes. These patterns indicate that the degree of separation from the abusive partner may be another factor that critically shapes the timing and stability of therapeutic change. Nevertheless, despite these contextual challenges, all three cases benefited from ImRs-PAD. This suggests that while achieving a no-contact condition maximizes treatment effectiveness, the intervention can still yield meaningful improvements even when ongoing or intermittent contact with the abusive partner persists.

Clinical Implications

This study provides preliminary clinical indications that ImRs-PAD may be adapted to address IPV-related difficulties in survivors presenting with PAD. By targeting unmet relational needs for love, safety, and dignity, ImRs offers a therapeutic approach that goes beyond symptom reduction, aiming to modify dysfunctional relational schemas that maintain dependence on abusive partners. Clinically, ImRs-PAD may support disengagement from violent relationships and potentially reducing the likelihood of returning to abusive partners—a situation associated with adverse outcomes, including heightened risk of femicide (McFarlane et al., 2004). These implications must be interpreted cautiously, given the preliminary nature of the evidence, and clinical application should proceed with care until findings are replicated in larger, controlled studies.

Limitations and Future Directions

This study has several limitations. First, it relied on a multiple-case series design, which limits the ability to draw causal inferences and generalize findings to broader populations. Second, the sample was relatively small and homogeneous, limiting applicability across different cultural, relational, or clinical contexts. Third, outcomes were primarily based on self-report measures, which may be subject to bias. Finally, online delivery may also present specific limitations in this population. For IPV survivors who are still living with a controlling or abusive partner, the home environment may not constitute a safe therapeutic space. Sessions may be subject to monitoring, interruption, or recording. These conditions may reduce the survivor's perceived safety and interfere with therapeutic openness. For this reason, when delivering the intervention online, particular attention should be paid to safety conditions. Therapists should ensure that patients have access to a private and secure space for the session and can interrupt it at any time if they feel at risk. Moreover, given the potentially reparative role of the therapeutic relationship in trauma-related relational difficulties, the absence of physical co-presence may affect engagement and treatment outcomes. In this regard, the possibility of subsequent in-person treatment should be considered, when feasible, to further support the

therapeutic process and enhance the reparative potential of the therapeutic relationship. Future research should directly compare online and in-person delivery modalities.

Despite these limitations, the findings are encouraging and highlight new directions for future research. Controlled studies with larger and more diverse samples are needed to further evaluate the effectiveness of ImRs-PAD and to identify predictors of early versus late treatment response. Once preliminary efficacy has been established, future studies may explore relational, contextual, and personality factors as potential moderators. For example, ongoing contact with the abusive partner, the type of violence (physical, sexual, psychological, or verbal), and/or specific schema may critically shape the pace and stability of improvements. Furthermore, dismantling studies could help clarify the relative contributions of early-life rescripting versus interventions targeting current and future partner-related dynamics. Longitudinal follow-up at three, six, and twelve months will be essential to determine whether the benefits of ImRs-PAD are sustained over time and translate into safer relational choices. In addition, incorporating behavioral outcomes—such as the intention to return to the abusive partner—and biological indicators, including cortisol, micro-RNAs, and inflammatory markers, could provide deeper insight into how ImRs-PAD supports survivors in breaking free from PAD and IPV. Finally, dedicated studies are needed to clarify the mechanisms of change in ImRs-PAD, showing how addressing unmet needs during rescripting translates into therapeutic gains.

Conclusion

This multiple case series provides preliminary support for the effectiveness of ImRs-PAD, a specific three-stage ImRs protocol grounded in PAD theory (Pugliese et al., 2023). According to this theory, PAD originates from the frustration of core needs—being loved, feeling safe, and experiencing self-worth—within childhood caregiving and adult intimate relationships. PAD may represent a relevant psychological risk pathway for IPV (Pugliese, Papa, et al., 2025; Pugliese, Uvelli, et al., 2025). After addressing these memories in guided imaginations and meeting the unmet relational needs during rescripting, we observed reductions in PAD symptoms, along with improvements in well-being, resilience, and self-compassion, as well as decreases in depression, anxiety, trauma-related, and somatic symptoms. These findings broaden the use of ImRs, showing its relevance for PAD and IPV survivors. They also contribute to a better understanding of the mechanisms of change in ImRs.

Funding: This research was supported by the *Stichting Achmea Slachtoffer en Samenleving (SASS)* for the project titled *"Should I Stay or Should I Go? Afhankelijkheid in gewelddadige partnerrelaties"*.

Acknowledgments: We sincerely thank Stichting Achmea Slachtoffer en Samenleving for their support in facilitating this research. We are also deeply grateful to the survivors of intimate partner violence who voluntarily participated in the study, generously sharing their experiences and contributing to the advancement of knowledge in this field.

Competing Interests: Erica Pugliese is the sole developer and trainer of the ImRs-PAD intervention evaluated in this study. Her remuneration for training activities partially supports her ongoing research. Arnoud Arntz occasionally offers training in ImRs, the remuneration goes to the university to support research. The authors declare no further conflicts of interest.

Author Contributions: Conceptualization: EP; intervention protocol development: EP; writing—original draft preparation: EP, AU; writing—review and editing: EP, AU, AA, AvE; Methodology: EP; Formal Analysis: AU visualization: AU; supervision: AA, AvE; project administration: EP. All authors have read and agreed to the published version of the manuscript.

Ethics Statement: The study was conducted in accordance with the Declaration of Helsinki and relevant national ethical guidelines for research involving human participants. Ethical approval was obtained from the Ethics Committee of the Scuola di Psicoterapia Cognitiva (protocol number Pr 2/24). All participants provided written informed consent, and data were collected and stored in accordance with applicable data protection regulations.

Preregistration: This manuscript reports a preliminary evaluation of ImRs-PAD conducted within the framework of a broader preregistered clinical study. The full study protocol is registered at ClinicalTrials.gov under the identifier NCT06670326. The present article reports the preliminary findings from the intervention protocol before completion of the full study. Therefore, the findings should be interpreted as preliminary and exploratory in relation to the broader preregistered trial.

Reporting Guidelines: The manuscript was prepared in accordance with the CARE guidelines for case reports and case series.

Use of AI Tools: The authors used AI-based language editing tools to improve the quality of the English. All scientific content, interpretations, and conclusions are the sole responsibility of the authors.

Social Media Accounts: *Erica Pugliese:* LinkedIn

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

Supplementary Materials

The appendix associated with this article includes the following contents (for access, see Pugliese et al., 2026S):

- Appendix A. Detailed description of the ImRs-PAD sessions. The appendix provides an overview of the structure, aims, and clinical focus of each session of the ImRs-PAD intervention.

Index of Supplementary Materials

Pugliese, E., Uvelli, A., van Emmerik, A. A. P., & Arntz, A. (2026S). *Supplementary materials to "Online application of imagery rescripting for pathological affective dependence (ImRs-PAD) in survivors of intimate partner violence: Results from a multiple case series"* [Appendix]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.22357

References

Alessandri, G., Vecchio, G. M., Steca, P., Caprara, M. G., & Caprara, G. V. (2007). A revised version of Kremen and Block's Ego Resiliency Scale in an Italian sample. *TPM – Testing, Psychometrics, Methodology in Applied Psychology, 14*(3-4), 165–183. https://www.tpmap.org/wp-content/uploads/2014/11/14.3.4.pdf

Arntz, A., & Weertman, A. (1999). Treatment of childhood memories: Theory and practice. *Behaviour Research and Therapy, 37*(8), 715–740. https://doi.org/10.1016/S0005-7967(98)00173-9

Castillo-Gonzáles, M., Mendo-Lázaro, S., León-del-Barco, B., Terán-Andrade, E., & López-Ramos, V.-M. (2024). Dating violence and emotional dependence in university students. *Behavioral Sciences, 14*(3), Article 176. https://doi.org/10.3390/bs14030176

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

Cuesta-García, A., & Crespo, M. (2022). Barriers for help-seeking in female immigrant survivors of intimate partner violence: A systematic review. *Revista de Victimología / Journal of Victimology, 14*, 33–59.

Crapolicchio, E., Regalia, C., Di Bernardo, G. A., & Cinquegrana, V. (2021). The role of relational dependence, forgiveness and hope on the intention to return with an abusive partner. *Journal of Social and Personal Relationships, 38*(9), 2474–2493. https://doi.org/10.1177/02654075211011546

Emirza, E. G., & Uzun, S. (2024). The effect of psychotherapeutic interventions on posttraumatic stress disorder symptoms, depression and self-esteem in women exposed to violence: A systematic review and meta-analysis. *Current Psychology, 43*, 26301–26316. https://doi.org/10.1007/s12144-024-06285-x

Grossi, E., Groth, N., Mosconi, P., Cerutti, R., Pace, F., Compare, A., & Apolone, G. (2006). Development and validation of the short version of the Psychological General Well-Being Index (PGWB-S). *Health and Quality of Life Outcomes, 4*, Article 88. https://doi.org/10.1186/1477-7525-4-88

Iverson, K. M., Resick, P. A., Suvak, M. K., Walling, S., & Taft, C. T. (2011). Intimate partner violence exposure predicts PTSD treatment engagement and outcome in cognitive processing therapy. *Behavior Therapy, 42*(2), 236–248. https://doi.org/10.1016/j.beth.2010.06.003

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

Jewkes, R., & Abrahams, N. (2026). Global burden of disease from intimate partner violence against women and sexual violence against children: A call to action. *Lancet, 407*(10523), 2–3. https://doi.org/10.1016/S0140-6736(25)02598-X

Jordan, C. E., Campbell, R., & Follingstad, D. R. (2010). Violence and women's mental health: The impact of physical, sexual, and psychological aggression. *Annual Review of Clinical Psychology, 6*, 607–628. https://doi.org/10.1146/annurev-clinpsy-090209-151437

Kip, A., Schoppe, L., Arntz, A., & Morina, N. (2023). Efficacy of imagery rescripting in treating mental disorders associated with aversive memories – An updated meta-analysis. *Journal of Anxiety Disorders, 99*, Article 102772. https://doi.org/10.1016/j.janxdis.2023.102772

Koetsier, T., Apeldoorn, J., & Nugter, A. (2024). Working mechanisms of imagery rescripting (ImRs) in adult patients with childhood-related PTSD: A pilot study. *European Journal of Psychotraumatology, 15*(1), Article 2339702. https://doi.org/10.1080/20008066.2024.2339702

Kroener, J., Hack, L., Mayer, B., & Sosic-Vasic, Z. (2023). Imagery rescripting as a short intervention for symptoms associated with mental images in clinical disorders: A systematic review and meta-analysis. *Journal of Psychiatric Research, 166*, 49–60. https://doi.org/10.1016/j.jpsychires.2023.09.010

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

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2002). The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine, 64*(2), 258–266. https://doi.org/10.1097/00006842-200203000-00008

McFarlane, J., Malecha, A., Gist, J., Watson, K., Batten, E., Hall, I., & Smith, S. (2004). Increasing the safety-promoting behaviors of abused women. *American Journal of Nursing, 104*(3), 40–51. https://doi.org/10.1097/00000446-200203000-00019

Müller-Engelmann, M., & Steil, R. (2017). Cognitive restructuring and imagery modification for posttraumatic stress disorder (CRIM-PTSD): A pilot study. *Journal of Behavior Therapy and Experimental Psychiatry, 54*, 44–50. https://doi.org/10.1016/j.jbtep.2016.06.004

Naismith, I., Ripoll-Nuñez, K., & Henao, G. B. (2024). Depression, anxiety, and posttraumatic stress disorder following intimate partner violence: The role of self-criticism, guilt, and gender beliefs. *Violence Against Women, 30*(3-4), 791–811. https://doi.org/10.1177/10778012221142917

Neal, A. M., & Edwards, K. M. (2017). Perpetrators' and victims' attributions for IPV: A critical review of the literature. *Trauma, Violence & Abuse, 18*(3), 239–267. https://doi.org/10.1177/1524838015603551

Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity, 2*(3), 223–250. https://doi.org/10.1080/15298860309027

Nilsson, J. E., Knutsson, J., Jalamo, B. S., & Lundh, L. G. (2019). Imagery rescripting of early memories in health anxiety disorder: A feasibility and non-randomized pilot study. *Journal of Behavior Therapy and Experimental Psychiatry, 65*, Article 101491. https://doi.org/10.1016/j.jbtep.2019.101491

Papa, C., & Pugliese, E. (2025, September 4-6). *Pathological affective dependence and fear of intimacy: Opposite but intertwined risk factors for intimate partner violence* [Poster abstract]. 55th EABCT Congress, Glasgow, United Kingdom.

Papa, C., Uvelli, A., Matos, M., Floridi, M., Franquillo, A. C., Perdighe, C., Mancini, F., & Pugliese, E. (2026). The hidden roots of interpersonal aggressiveness: Dependency, jealousy, and revenge. *Victims & Offenders*. Advance online publication. https://doi.org/10.1080/15564886.2026.2652379

Paphitis, S. A., Bentley, A., Asher, L., Osrin, D., & Oram, S. (2022). Improving the mental health of women intimate partner violence survivors: Findings from a realist review of psychosocial interventions. *PLoS One, 17*(3), Article e0264845. https://doi.org/10.1371/journal.pone.0264845

Pezzoli, P., Pingault, J. B., Eley, T. C., McCrory, E., & Viding, E. (2025). Causal and common risk pathways linking childhood maltreatment to later intimate partner violence victimization. *Molecular Psychiatry, 30*(5), 2027–2037. https://doi.org/10.1038/s41380-024-02813-0

Phillips, A. R., Williams, R., Okuoimose, V., Hiller, R. M., McCrory, E., Viding, E., & Pezzoli, P. (2025). Lived-experience perspectives on the psychological factors linking childhood maltreatment to later intimate partner violence victimization. *European Journal of Psychotraumatology, 16*(1), Article 2524922. https://doi.org/10.1080/20008066.2025.2524922

Pugliese, E., Papa, C., Filippelli, M. C., Murabito, F., Riposo, S., & Carcione, A. (2025). Pathological affective dependence, intimate partner violence, and metacognition: The moderating role of different forms of violence in Italian women survivors. *Journal of Aggression, Maltreatment & Trauma, 34*(11), 1641–1660. https://doi.org/10.1080/10926771.2025.2585270

Pugliese, E., Saliani, A. M., Mosca, O., Maricchiolo, F., & Mancini, F. (2023). When the war is in your room: A cognitive model of pathological affective dependence (PAD) and intimate partner violence (IPV). *Sustainability, 15*(2), Article 1624. https://doi.org/10.3390/su15021624

Pugliese, E., Uvelli, A., van Emmerik, A., Ferretti, F., Saliani, A. M., Foschino-Barbaro, M. G., Vigilante, T., Celitti, E., Quintavalle, C., Mancini, F., & Arntz, A. (2025). Pathological affective dependence as a risk factor for intimate partner violence: Initial psychometric validation of the Italian version of the pathological affective dependence scale. *Clinical Psychology & Psychotherapy, 32*(4), Article e70140. https://doi.org/10.1002/cpp.70140

Pugliese, E., Visco-Comandini, F., Papa, C., Ciringione, L., Cornacchia, L., Gino, F., Cannito, L., Fadda, S., & Mancini, F. (2024). Understanding trauma in IPV: Distinguishing complex PTSD, PTSD, and BPD in victims and offenders. *Brain Sciences, 14*(9), Article 856. https://doi.org/10.3390/brainsci14090856

Silvestri, V., Gobbo, S., Pugliese, E., Mancini, F., & Visco-Comandini, F. (2025). The perception of trustworthiness and emotional identification in women experiencing intimate partner violence: A behavioral pilot study. *Brain Sciences, 15*(5), Article 429. https://doi.org/10.3390/brainsci15050429

Smith-Marek, E. N., Cafferky, B., Dharnidharka, P., Mallory, A. B., Dominguez, M., High, J., Stith, S. M., & Mendez, M. (2015). Effects of childhood experiences of family violence on adult partner violence: A meta-analytic review. *Journal of Family Theory & Review, 7*(4), 498–519. https://doi.org/10.1111/jftr.12113

Speranza, A. M., Farina, B., Bossa, C., Fortunato, A., Maggiora Vergano, C., Palmiero, L., Quintigliano, M., & Liotti, M. (2022). The role of complex trauma and attachment patterns in intimate partner violence. *Frontiers in Psychology, 12*, Article 769584. https://doi.org/10.3389/fpsyg.2021.769584

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

Uvelli, A., Floridi, M., Agrusti, G., Franquillo, A. C., Fiumalbi, L., Micheloni, T., Arcuri, A., Iazzetta, S., & Gragnani, A. (2025). When adverse experiences influence the interpretation of ourselves, others and the world: A systematic review and meta-analysis of maladaptive schemas in victims of violence. *Clinical Psychology & Psychotherapy, 32*(4), Article e70114. https://doi.org/10.1002/cpp.70114

van Gelder, N. E., Ligthart, S. A., van Rosmalen-Nooijens, K. A., Prins, J. B., & Oertelt-Prigione, S. (2023). Effectiveness of the SAFE eHealth intervention for women experiencing intimate partner violence and abuse: Randomized controlled trial, quantitative process evaluation, and open feasibility study. *Journal of Medical Internet Research, 25*, Article e42641. https://doi.org/10.2196/42641

Visco-Comandini, F., Papa, C., Uvelli, A., Mancini, F., & Pugliese, E. (2025). Rewriting trauma: A systematic review of treatment effects of imagery rescripting for PTSD and Complex PTSD. *European Journal of Trauma & Dissociation / Revue Européenne du Trauma et de la Dissociation, 9*(4), Article 100609. https://doi.org/10.1016/j.ejtd.2025.100609

World Health Organization. (2024, March 25). *Violence against women*. World Health Organization. https://www.who.int/news-room/fact-sheets/detail/violence-against-women

Young, J. E., & Brown, G. (2005). *Young Schema Questionnaire-Short Form*; Version 3 (YSQ-S3, YSQ) [Database record]. APA PsycTests. https://doi.org/10.1037/t67023-000

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

Sleep Quality, Relationship Functioning, and Health-Related Quality of Life in Partners of CPAP-Treated OSA Patients

Giulia Marafioti [1], Enrica S. Vinci [2], Laura Culicetto [1], Viviana Lo Buono [1], Giovanni L. Cipriano [1], Rosalia Silvestri [3], Amelia Rizzo [2]

[1] *IRCCS Bonino-Pulejo Neurolesi Center, Messina, Italy.* **[2]** *National Institute of Social Welfare, Medical Legal Center of Messina, Messina, Italy.* **[3]** *Sleep Medicine Center, UOSD of Neurophysiopathology and Movement Disorders, Department of Clinical and Experimental Medicine, University of Messina, Messina, Italy.*

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e20295, https://doi.org/10.32872/cpe.20295

Received: 2025-10-14 • **Accepted:** 2026-06-21 • **Published (VoR):** 2026-08-31

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

Corresponding Author: Giulia Marafioti, IRCCS Bonino-Pulejo Neurolesi Center, Via Palermo S.S. 113, Contrada Casazza, 98124 Messina, Italy. Telephone: [+39 090/60128102]. E-mail: giulia.marafioti@irccsme.it

Abstract

Background: Obstructive Sleep Apnea Syndrome (OSAS) is a chronic condition that negatively affects not only the patient but also the partner, influencing their sleep quality, psychological well-being, and couple relationship. However, scientific literature has given little attention to the partner's experience, especially regarding the impact of Continuous Positive Airway Pressure (CPAP) treatment on their partners' health. The aim of the present study is to investigate sleep quality, couple relationship quality, and quality of life (QoL) in Italian-speaking partners of patients with OSAS treated with CPAP, compared to partners of healthy individuals.

Method: The study involved 100 participants, with a final sample of 48 partners of OSA patients and 49 partners of healthy subjects as a control group, assessed using the Pittsburgh Sleep Quality Index (PSQI), the Dyadic Adjustment Scale (DAS), and the SF-36 Health Survey. Statistical analysis included independent samples *t*-tests and Pearson's correlations.

Results: Partners of patients with OSAS reported significantly poorer sleep quality (PSQI), lower levels of dyadic consensus and cohesion (DAS), and reduced perception of physical and mental health (SF-36) compared to the control group. Correlations revealed that sleep quality was negatively associated with relationship quality and perceived health, while years of relationship were positively associated with affection and cohesion, but negatively with sleep quality.

Conclusions: The findings confirm the interconnection between sleep quality and relationship functioning. A clinical approach that includes the partner in OSA treatment pathways may improve the overall well-being of the dyad.

Keywords

obstructive sleep apnea, CPAP, sleep quality, couple relationship, quality of life, partner

Highlights

- Partners of patients with Obstructive Sleep Apnea (OSA) report poorer sleep, health, and relationship quality.
- OSA is a "shared" disorder that negatively impacts couple cohesion and consensus.
- Poor sleep quality in partners is linked to worse relationship functioning and health.
- A dyadic approach to OSA treatment may improve overall well-being for the couple.

Obstructive Sleep Apnea Syndrome (OSAS) falls within the category of Sleep-Disordered Breathing (SDB). SDBs are frequently comorbid with various pathologies; for example, they are common in patients with heart failure (Khokhrina et al., 2022). Moreover, they have been epidemiologically and pathophysiologically linked to ischemic heart disease, atrial fibrillation, arterial hypertension (Fava et al., 2011; Gonzaga et al., 2015; Gopalakrishnan & Tak, 2011), diabetes mellitus, metabolic syndrome (Drager et al., 2013), sudden death (Gami et al., 2005), and stroke (Yaggi et al., 2005). The prevalence of SDB is progressively increasing due to the rising incidence of obesity (Peppard et al., 2013). In this context, it is essential to distinguish between two primary syndromes: Obstructive Sleep Apnea (OSA) and Central Sleep Apnea (CSA). OSAS is a widespread but often unrecognized respiratory disorder characterized by episodes of partial or complete upper airway closure during sleep (Locke et al., 2022).

The complications of this syndrome frequently affect the cardiovascular system (hypertension initially appearing at night and later during the day) or the nervous system, leading to cognitive and mood disorders (depression, anxiety, emotional instability; Harding, 2000). These complications strongly impact not only the life of the affected individual but also that of their partner, often leading to imbalances in the relationship. Several studies in the literature have highlighted how sleep disorders or sleep deprivation can negatively affect cognitive, relational, and emotional processes, consequently influencing interpersonal interpretation abilities (Durmer & Dinges, 2005; Hasler & Troxel, 2010; Troxel et al., 2007). Individuals affected by OSAS are often considered ‘heavy snorers’. Various studies emphasize the great difficulty that partners experience in managing this condition (Cartwright, 2008; Luyster et al., 2016). Many partners report feeling deeply distressed while witnessing these nocturnal apneas (Troxel et al., 2009). On one hand, continuous positive airway pressure treatment often reduces snoring and apneic events, which are among the main causes of sleep disruption for

partners of individuals with OSAS. In this sense, CPAP can significantly improve the partner's sleep continuity and perceived restfulness. On the other hand, some partners may initially experience sleep disturbance due to device-related factors such as machine noise, airflow sounds, mask leaks, or the physical presence of the equipment in bed. For certain couples, especially in the early adaptation phase, these elements may temporarily disrupt sleep, with possible repercussions on the couple's relationship quality.

However, in the literature, few and relatively outdated studies have explored the association between couple quality and sleep quality. Additionally, the results are conflicting: some studies reveal a clear deterioration in relationship quality when a sleep disorder is present (McArdle et al., 2001; Virkkula et al., 2005), whereas others do not confirm such an association (Broström et al., 2010; Parish & Lyng, 2003). Further research has investigated how OSAS is a shared problem affecting both spouses, leading to fragmented sleep, daytime sleepiness, and reduced quality of life (QoL), often resulting in tense marital relationships (Doherty et al., 2003; Henry, 2016) found a significant association between chronic insomnia and the deterioration of romantic relationships. Pankhurst and Home (1994) were the first to analyze the effects of nocturnal movements of OSAS patients on their partners' sleep in couples without other sleep disorders. A gender-based study showed that men report more nighttime movements than women, and women are more likely to experience sleep disturbances caused by their partner compared to men (Baron et al., 2009). Some qualitative studies have highlighted the importance of partners' perspectives in understanding the experience of OSAS. Improvements in sleep and increased alertness in both individuals involved contribute to a better understanding of the health consequences of OSAS. The nighttime and daytime consequences of OSAS often lead individuals to sleep in separate beds, reducing intimacy and contributing to relationship tensions. This is supported by various qualitative and quantitative studies that report poor sleep quality and impaired QoL, affecting both OSA patients and their partners (Baron et al., 2011; Hoy et al., 1999). Conversely, some cross-sectional studies have demonstrated adverse associations between OSAS and relationship quality as reported by the patient or partner (McArdle et al., 2001; Virkkula et al., 2005). It is well known that QoL improves after CPAP treatment; however, the effects on the patient's bed partner have received little attention. The role of relationship factors in CPAP adherence remains poorly understood, but some studies suggest that partners may have both positive and negative influences on CPAP adoption and use. In this regard, some studies have identified that partners of OSAS patients also benefit from treatment, showing improved perception levels, QoL (Billmann & Ware, 2002), sleep quality (Luyster et al., 2016), and couple relationship quality (Beninati et al., 1999). Specifically, a study conducted at Rush University Medical Center in Chicago examined the sleep of married couples, recording it in a laboratory before and after the husband's CPAP treatment to assess adherence levels. The study found that adherence was strongly correlated with the wives' involvement, emphasizing that bed-sharing is fundamental for improving

treatment adherence (Cartwright, 2008). Based on the literature and previous research, this study aims to address the Research Objectives (RO):

- **RO 1** – Compare Sleep Quality, Relationship Functioning, and Health-Related Quality of Life in Partners of Patients with and without CPAP-Treated OSAS;
- **RO 2** – Compare differences in correlational patterns of Sleep Quality, Relationship Functioning, and Health-Related Quality of Life in Partners of patients with and without CPAP-Treated OSAS.

The impact of OSAS extends beyond the individual patient, significantly affecting the partner's sleep quality, psychological well-being, and relationship dynamics. The literature highlights the distress experienced by partners due to sleep disruptions, emotional strain, and reduced intimacy, often leading to tensions within the couple. However, findings remain inconsistent, with some studies confirming a decline in relationship satisfaction, while others do not observe such an association. CPAP therapy has been shown to improve sleep quality and overall well-being for both patients and their partners, yet its effects on relationship satisfaction and psychological health are still underexplored. The present study aims to bridge this gap by comparing sleep quality, couple satisfaction, and QoL in partners of people with and without CPAP treatment (RO1) and investigating differences in correlational patterns when compared to a healthy control group (RO2). By addressing these aspects, this research seeks to provide deeper insights into the broader implications of OSAS and highlight the importance of partner involvement in treatment adherence and overall well-being.

Materials and Method

The study protocol initially included the administration of an informed consent form, a document that thoroughly informed participants about the study and collected their authorization to participate. Subsequently, each participant completed a demographic form to collect information on demographics, education level, any pathologies, and comorbidities. The questionnaires were administered individually once a week, under the supervision of a psychologist who provided clarifications and support to the participants, if needed. This protocol ensured the accurate collection of data and ensured that participants had a clear understanding of the questions and the process of participating in the study.

Participants

A total of 100 subjects were recruited in the research, divided into two groups: 50 partners of OSA patients (see Table 1 for the sample characteristics), attending the Sleep Medicine Center of the University Hospital of Messina "G. Martino," between October 2019 and March 2020, and 50 partners of healthy individuals as a control group.

Participants were classified into two groups based on clinical status: individuals diagnosed with OSA and healthy controls (HC). OSA patients were routinely monitored for CPAP treatment adherence through regular clinical follow-up and device-recorded data.

Inclusion in the OSA group required a confirmed clinical diagnosis established through standard polysomnographic assessment and medical evaluation conducted by a sleep specialist. Only adults were included, and participants had to be in a stable romantic relationship, given the relational variables investigated in the study. Inclusion in the HC group required the absence of a diagnosis of sleep disorders and the absence of self-reported symptoms suggestive of sleep apnea, as well as the same relational status criteria applied to the OSA group.

Exclusion criteria for both groups included the presence of severe psychiatric disorders, neurodegenerative conditions, major uncontrolled medical illnesses, or cognitive impairments that could interfere with comprehension of the questionnaires. Individuals with current substance abuse or other diagnosed sleep disorders were also excluded. For the HC group specifically, any previous diagnosis of OSA or other clinically relevant sleep disturbances represented an exclusion criterion.

Two participants from the OSA partners group were excluded due to dropout prior to protocol completion. One participant from the HC group was excluded because they were not co-sleeping with their partner. In accordance with the predefined inclusion and exclusion criteria, the final sample consisted of 48 partners of individuals diagnosed with OSA and 49 HC group.

Participants were recruited using a non-probability sampling method. The OSA group was likely selected through consecutive sampling from patients attending a sleep clinic or undergoing diagnostic evaluation, while the HC group was recruited through convenience sampling from the community, ensuring comparability in basic sociodemographic characteristics. This sampling strategy aimed at obtaining two clinically distinct but demographically comparable groups rather than a randomly selected population sample.

Instruments

Pittsburgh Sleep Quality Index (PSQI)

The PSQI is a self-report questionnaire developed by researchers at the University of Pittsburgh (Buysse et al., 1989). It is intended to be a standardized sleep questionnaire that researchers and doctors can easily use, and it is utilized across various populations and in different contexts, including research activities and clinical settings. It consists of 19 items that assess the subjects' perceived sleep quality and disturbances over a 1-month period, and completion takes between 5 and 10 minutes (Buysse et al., 1989). The 19 items are grouped into 7 composite items, rated on a scale from 0 to 3, which, when summed, give the overall PSQI score. The overall score ranges from 0 to 21, where "≤ 5" corresponds to good sleep quality while "> 5" corresponds to poor sleep

quality. These 7 composite items explore sleep quality through a wide range of domains that include: The usual methods of awakening, the duration of sleep, sleep latency, the frequency and severity of specific problems that arise during sleep (such as the presence of pain, frequent urination, breathing difficulties, the presence of snoring, dream activity, temperature), the use of hypnotic medications, daytime disturbances, and usual effectiveness. All of this is measured retrospectively.

The PQSI exhibits good psychometric characteristics to the extent that it can be said that the authors have effectively developed a reliable, valid, and standardized measure of sleep quality. In addition to the promising reliability and validity of the measure, its brevity and accessibility as a free measure give it great potential for clinical practice (Buysse et al., 1989). In the present study, the reliability of the Italian version of the PSQI was assessed using Cronbach's alpha, which yielded values between .80 and .90, indicating good internal consistency of the scale.

Dyadic Adjustment Scale (DAS)

The DAS is a multidimensional tool designed by Spanier (1976). The DAS consists of 32 items divided into four subscales: Dyadic Consensus (DC); Dyadic Satisfaction (DS); Dyadic Cohesion (DH); Affective Expression (AE). The sum of the four scales provides a total score that expresses the overall level of agreement within the couple. The dyadic consensus scale (DC) consists of 13 items and assesses the degree of agreement and disagreement between partners on topics such as leisure time management and finances, or on religion, friendships, and household organization. The dyadic satisfaction scale (DS) consists of 10 items and evaluates the happiness or unhappiness that couples perceive regarding their relationship; this scale observes the frequency of arguments, the pleasure or lack thereof in being together, and the consideration of separation or divorce. The diadic cohesion scale (DH) consists of 5 items and evaluates the amount of time partners share pleasant activities such as social interests, dialogue, or having common goals. Finally, the affective expression scale (AE) consists of 4 items and evaluates how the couple expresses their feelings, love, and sexuality. It is a simple tool that takes about 6-7 minutes to complete; a quiet place, a pen, and the questionnaire are required. The scale is aimed at both married and unmarried couples, and it is filled out independently. The instruction given to the subjects is to carefully read the instructions and understand their meaning. It is considered the most widely used measurement tool for evaluating the "quality" of the relationship (Spanier, 1976). In the present study, the reliability of the Italian versions of the DAS was assessed using Cronbach's alpha, yielding values between .80 and .90 for both scales, indicating good internal consistency.

Total scores on the scale range from 0 to 150. Higher scores reflect greater relationship quality, satisfaction, and adjustment, whereas lower scores indicate relational dissatisfaction or distress. In his original validation study, Spanier (1976) proposed a clinical cutoff whereby scores below 101 indicate *relational distress*, and scores of 102 or higher

indicate a *non-distressed* relationship. For reference, Spanier reported mean total scores of 70.7 in a divorced sample and 114.8 in a married sample, supporting the discriminative measures' validity (Spanier, 1976).

Health Survey (SF-36)

The Short Form-36 Health Survey (SF-36; Ware & Sherbourne, 1992) was used to assess participants' perceived health status. The instrument evaluates eight domains: Physical Functioning, Role Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role Emotional, and Mental Health. Two summary measures can also be derived: the Physical Component Summary (PCS) and the Mental Component Summary (MCS). The Italian validation by Apolone and Mosconi (1998) demonstrated good psychometric properties, with Cronbach's alpha coefficients for the subscales ranging from 0.73 to 0.93, indicating satisfactory to excellent internal consistency.

Statistical Analysis

All statistical analyses were conducted with SPSS (Version 27.0). Independent-samples t tests were used to compare Sleep Quality, Relationship Functioning, and Health-Related QoL between partners of patients with CPAP-treated OSA and partners of individuals without OSA. Pearson's correlation coefficients (r) were computed separately within each group to examine associations among the three constructs. Differences in correlational patterns between the two groups were evaluated using Fisher's r to z transformation. For each analysis, unadjusted p-values were reported, and results were interpreted according to the conventional criterion of $p < .05$, without applying additional corrections for multiple comparisons. Effect sizes (Cohen's d for group differences and r for correlations) and 95% confidence intervals were calculated to aid interpretation of the findings.

Results

Sample Characteristics

The data analysis began with the qualitative analysis of various constructs. Table 1 presents the sociodemographic and relationship characteristics of the OSA (n = 48) and HC (n = 49) partner groups. All participants included in the dyadic analyses reported being currently involved in a stable romantic relationship. The category "in a stable relationship / partnered" includes not only individuals who are unmarried or cohabiting, but also those with other civil statuses (e.g., separated or widowed) who currently have a partner.

Table 1

Sociodemographic and Relationship Characteristics of the HC and OSA Groups

Variable / Category	HC *n* (%)	OSA *n* (%)
Number of valid protocols		
Valid	48 (96.0%)	50 (100.0%)
Invalid	2 (4.0%)	0 (0.0%)
Civil/marital status		
In a stable relationship / partnered	47 (97.9%)	49 (98.0%)
Other civil status (e.g., separated/widowed), with current partner	1 (2.1%)	1 (2.0%)
Relationship duration (years)		
< 1 year	0 (0.0%)	1 (2.0%)
1–3 years	3 (6.3%)	1 (2.0%)
4–6 years	2 (4.2%)	3 (6.0%)
7–10 years	2 (4.2%)	2 (4.0%)
11–15 years	9 (18.8%)	5 (10.0%)
16–20 years	15 (31.3%)	4 (8.0%)
21–25 years	7 (14.6%)	16 (32.0%)
> 25 years	10 (20.8%)	18 (36.0%)
Educational level		
Primary school	2 (4.2%)	7 (14.0%)
Secondary school	11 (22.9%)	21 (42.0%)
High school diploma	26 (54.2%)	12 (24.0%)
University degree	9 (18.8%)	10 (20.0%)
Sleeping with Partner		
Yes	48 (100.0%)	49 (98.0%)
No	0 (0.0%)	1 (2.0%)

Note. Statistics on the whole sample (N = 100), valid cases for OSA = Obstructive Sleep Apnea partners (n = 49), and for Healthy Controls (n = 48).

Given that the DAS assesses current couple functioning, DAS scores, DAS-based classifications, and correlations involving DAS variables were computed only for participants who reported having a current romantic partner. Participants not sharing the bed with their partner were excluded from dyadic analyses, despite reporting an ongoing romantic relationship; to this extent, co-sleeping status was described separately and was not used as a proxy for relationship status.

The OSA group shows a higher concentration in the longest relationship categories (> 20 years), while the HC group appears more in the mid-range durations (11 to 20 years). Educationally, the HC group is more represented in the high school category, whereas the OSAS group shows a greater proportion with lower secondary education. All participants in the HC group reported sleeping with their partner, whereas 49 out of 50 participants in the OSA group reported co-sleeping. Therefore, the participant who did not meet the inclusion criteria was excluded from the analysis.

RO 1 – Sleep Quality, Relationship Functioning, and Health-Related Quality of Life in Partners of Patients With and Without CPAP-Treated OSA

To address Research Objective 1 (RO1), Student's *t*-test for independent samples was performed (see Table 2) to explore whether the presence of OSA, even when treated with continuous positive airway pressure (CPAP), is associated with differences in partners' subjective sleep experience, relational dynamics, and perceived well-being.

Table 2

Mean Scores and Standard Deviations Obtained by the Participants

Measure / Group	*M*	*SD*
SF-36 PCS		
OSA Partners	63.23*	19.73
HC Partners	75.57*	10.85
SF-36 MCS		
OSA Partners	65.37*	16.34
HC Partners	74.77*	10.44
PSQI		
OSA Partners	11.06*	3.69
HC Partners	6.00*	3.37
Dyadic Consensus (DAS)		
OSA Partners	3.06*	0.74
HC Partners	4.07*	0.59
Dyadic Satisfaction (DAS)		
OSA Partners	3.67	0.62
HC Partners	3.64	0.68
Affective Expression (DAS)		
OSA Partners	2.27	0.53
HC Partners	2.15	0.46
Dyadic Cohesion (DAS)		
OSA Partners	1.04*	1.03
HC Partners	2.97*	0.98

Note. DAS = Dyadic Adjustment Scale; PCS = Physical Component Score; MCS = Mental Component Score; PSQI = Pittsburgh Sleep Quality Index; HC = Healthy Controls; OSA Partners = Obstructive Sleep Apnea partners.
*$p < .05$.

The data analysis conducted highlighted some statistically significant differences. Specifically, in the control group, there is a greater perception of physical health status compared to the group of partners with OSA patients, $t(95) = -3.80$, $p < .001$, and this result is also confirmed regarding the perception of mental health status, $t(95) = -3.80$, $p =$

.001. Again, regarding sleep quality, the group of partners of OSA patients reports poorer sleep quality compared to the control group, $t(95) = 0.85$, $p = .032$. Subsequently, we considered the construct of couple functioning. Specifically, in the dimension of dyadic consensus, the control group records higher levels of agreement within the couple, $t(95) = -0.48$, $p = .042$, compared to the group of partners of patients with OSA. No significant differences were found in the dimensions of dyadic satisfaction, emotional expression, and cohesion.

In other words, partners of patients with OSA seem to experience a tangible impact on their personal well-being, both physically and psychologically, and they also report poorer sleep. However, when the focus shifts to the quality of the relationship itself, the differences are more nuanced.

Specifically, the lower dyadic consensus suggests that these couples may experience more difficulties in agreement or coordination in everyday matters. Yet, the absence of differences in satisfaction, emotional expression, and cohesion indicates that the emotional bond and overall relationship satisfaction are not necessarily compromised. In essence, OSA appears to affect well-being and certain functional aspects of the relationship more than its emotional core.

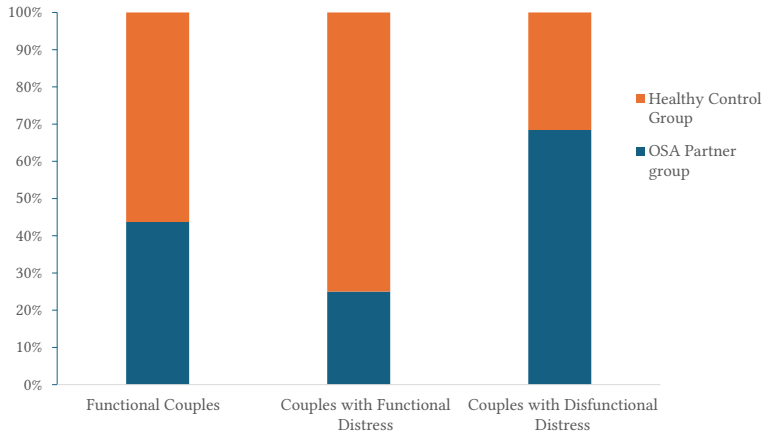
We first examined the overall level of couple functioning in the total sample. Results showed that 37% of couples fell within the functional range (DAS total score range 99-113), indicating high harmony and good relational adjustment. The majority, 57%, were classified in the functional distress range (71–98), suggesting the presence of relational strain but with the capacity to manage conflicts constructively. The remaining 6% (0–70) were characterized by marked relational dysfunction, reflecting persistent disagreement and poor conflict resolution.

We then compared the two groups. Among partners of patients with OSA, 59.2% of couples were classified as functional, 38.8% as experiencing dysfunctional distress, and 2% as functional distress. In contrast, the control group showed a higher proportion of functional relationships (76.3%), a lower percentage of dysfunctional distress (17.4%), and 6.3% in the functional distress range. These differences are illustrated in Figure 1.

Sleep quality was categorized according to the standard PSQI global cut-off (> 5) commonly used to distinguish poor from good sleepers (Buysse et al., 1989). Importantly, good sleep quality was not defined as an a priori inclusion criterion for the control group; rather, controls were recruited based on the general eligibility criteria described in the Method section. The fact that the control group consisted exclusively of "good sleepers" therefore reflects the observed distribution in this sample, and not a selection rule. Accordingly, 0% of controls scored > 5 (poor sleepers), whereas 55% of partners of patients with OSA exceeded the cut-off. Given this unexpectedly homogeneous profile in controls, group differences in sleep quality should be interpreted cautiously, as the absence of poor sleepers may contribute to an overestimation of between-group differences. To contextualize these findings, we also compared the observed sleep quality to

Figure 1

Comparison of Quality of Couple Relationship



Note. OSA Partners = Obstructive Sleep Apnea partners.

population-based normative evidence. In community samples, PSQI mean values are typically around ~ 4–5, and the proportion of individuals above the > 5 cut-off is commonly non-negligible (e.g., around one quarter in large population-based surveys; Hinz et al., 2017). In this respect, the control group in the present study appears to exhibit unusually good sleep quality relative to general population benchmarks, which further supports a cautious interpretation of group comparisons involving PSQI.

RO 2 – Differences in Correlational Patterns of Sleep Quality, Relationship Functioning, and Health-Related Quality of Life in Partners of Patients With and Without CPAP-Treated OSA

In the data analysis, the next examination proceeded with the differences in correlational patterns. The table below reports the Pearson correlations with the *p*-value obtained by the participants (Table 3 and 4).

The correlation analysis revealed several meaningful relationships in partners of patients undergoing OSA treatment. Dyadic consensus was positively associated with satisfaction, cohesion, affective expression, and the length of the relationship, indicating that longer-lasting couples tend to show greater harmony, emotional expression, and overall relationship satisfaction. Dyadic satisfaction also increased with cohesion and relationship duration, while affective expression grew with the years together. Cohesion was positively related to relationship length but negatively correlated with sleep quality

(PSQI), suggesting that poorer sleep reduces couple cohesion. Sleep quality itself was negatively associated with physical and mental health and, unexpectedly, with relationship duration, indicating that longer relationships may experience poorer sleep. Physical health (PCS) declined with longer relationship length, while mental health (MCS) decreased as sleep quality worsened, highlighting the interconnection between sleep, health perception, and relationship functioning.

Table 3

Correlations Among Partners of Patients With CPAP-Treated OSA (n = 48)

Variable	1	2	3	4	5	6	7
1. DAS DC	–	.37**	.41**	.41**	.04	.04	.12
		.009	*.004*	*.004*	*.805*	*.765*	*.410*
2. DAS DS		–	.21	.32*	−.02	.09	.23
			.144	*.027*	*.892*	*.525*	*.112*
3. DAS AE			–	.18	.16	−.14	−.12
				.217	*.271*	*.322*	*.413*
4. DAS DH				–	−.01	.15	.09
					.043	*.309*	*.551*
5. PSQI					–	−.51**	−.49**
						.000	*.000*
6. SF-36 PCS						–	.81**
							.000
7. SF-36 MCS							–

Note. Shows Pearson Correlations and *p*-values. DAS DC = Dyadic Adjustment Scale Dyadic Consensus; DAS DS = Dyadic Adjustment Scale Dyadic Satisfaction; DAS AE = Dyadic Adjustment Scale Affective Expression; DAS DH = Dyadic Adjustment Scale Dyadic Cohesion; PSQI = Pittsburgh Sleep Quality Index; SF-36 PCS = Physical Component Summary; SF-36 MCS = Mental Component Summary.
Italics indicate *p*-values. *$p < .05$. **$p < .01$.

In conclusion, in the OSA partner group, sleep quality was significantly associated with both physical and mental health but showed no correlation with relationship quality, suggesting that poor sleep primarily affects individual well-being rather than relational functioning. Conversely, in the healthy partner group, relationship quality was positively correlated with sleep quality and physical health, but not with mental health, indicating that for these partners, better relational functioning is linked to more restorative sleep and better perceived physical well-being, while mental health appears independent of relationship quality.

Table 4

Correlations Among Partners of Healthy Controls (n = 49)

Neuropsychological Assessment	1	2	3	4	5	6	7
1. DAS DC	–	.30*	.24	.31*	−.26	.15	.22
		.036	*.096*	*.031*	*.076*	*.303*	*.131*
2. DAS DS		–	.36*	.44**	−.39**	.33*	−.01
			.013	*.002*	*.006*	*.023*	*.964*
3. DAS AE			–	.52**	−.24	.31*	.25
				.000	*.104*	*.030*	*.085*
4. DAS DH				–	−.34*	.33*	.18
					.017	*.022*	*.217*
5. PSQI					–	−.31*	−.09
						.035	*.526*
6. SF-36 PCS						–	.42**
							.003
7. SF-36 MCS							–

Note. Shows Pearson Correlations and *p*-values. DAS DC = Dyadic Adjustment Scale Dyadic Consensus; DAS DS = Dyadic Adjustment Scale Dyadic Satisfaction; DAS AE = Dyadic Adjustment Scale Affective Expression; DAS DH = Dyadic Adjustment Scale Dyadic Cohesion; PSQI = Pittsburgh Sleep Quality Index; SF-36 PCS = Physical Component Summary; SF-36 MCS = Mental Component Summary.
Italics indicate *p*-values. $^{*}p < .05$. $^{**}p < .01$.

Discussion

The findings of the present study suggest that partners of patients diagnosed with OSA face multiple challenges, including poorer sleep quality, lower perceived physical and mental health, and reduced levels of couple cohesion and dyadic consensus. While not all subscales of the DAS revealed statistically significant differences, the overall pattern aligns with previous literature highlighting that OSA is not solely a patient-centered condition but a "shared" disorder with significant repercussions for partners and the couple as a relational unit (Bercea et al., 2013; Schröder & O'Hara, 2005).

To minimize the possibility that these differences reflect preexisting characteristics rather than the effects of OSA and CPAP use, both groups were selected using similar inclusion criteria (adulthood, stable relationships, absence of severe psychiatric or neurological conditions), and statistical analyses controlled for potential confounders such as age, relationship duration, and educational level. Although this is an observational study, these precautions support the interpretation that differences in sleep quality, health, and

couple functioning are consistent with the impact of OSA and CPAP adherence rather than innate couple characteristics.

Sleep quality emerged as a critical factor. Partners of OSA patients reported disrupted sleep, which was significantly associated with poorer physical and mental health but not directly with relationship quality. This is consistent with evidence that even mild or intermittent nocturnal disturbances, such as snoring or apneic events, can impair partner sleep and increase stress and interpersonal conflict (Luyster et al., 2016). In contrast, in the control group, relationship quality correlated with sleep quality and physical health, but not mental health, suggesting that for healthy partners, relational functioning is closely tied to restorative sleep and physical well-being, while mental health remains relatively independent.

Despite CPAP therapy being well established for improving outcomes in patients with OSA (McArdle et al., 1999), its effects on partners' well-being and couple functioning are less understood. The current findings tentatively support the idea that partner involvement in CPAP adherence may have broader benefits beyond patient health, potentially enhancing relational functioning, although longitudinal or intervention studies are needed to confirm this (Luyster et al., 2016).

Using preliminary classifications, couples were categorized as functional, experiencing functional distress, or dysfunctional. The majority (57%) fell into functional distress, characterized by conflict alongside the ability to resolve disagreements. A smaller proportion demonstrated persistent conflict (dysfunctional distress), while 37% were fully functional. Partners of OSA patients were more likely to experience dysfunctional distress (38%) than controls (17%), while functional relationships were more common in the control group (76%), suggesting greater cohesion and relational stability in healthy couples. These findings echo prior research linking chronic illness to increased relational strain (Luyster et al., 2016), though sampling characteristics may partially influence these differences. It is important to consider that the patient couples (PC) included in this study may represent a subset of partnerships that have endured despite the challenges posed by sleep apnea. Prior to treatment, these couples could have experienced more severe disruptions in relational functioning, including increased conflict, emotional distancing, or diminished cohesion.

Correlational analyses provided further insights. In partners of OSA patients, dyadic consensus correlated positively with satisfaction, cohesion, affective expression, and relationship duration, suggesting that longer relationships may sustain higher relational quality and emotional expression. However, cohesion was negatively correlated with sleep quality (PSQI), and sleep quality itself correlated negatively with perceived physical and mental health, highlighting how chronic sleep disruption can erode both individual well-being and aspects of relational functioning (Hasler & Troxel, 2010; Troxel et al., 2009). In the control group, these patterns were less complex: relationship quality corre-

lated mainly with satisfaction and cohesion, and sleep quality had weaker associations, reinforcing the specific burden of OSA on partners.

Across both groups, affective expression, satisfaction, and cohesion were positively associated with perceived health, underscoring the potential buffering role of emotional intimacy in mitigating health-related stressors (Revenson & DeLongis, 2010; Røsand et al., 2012).

In conclusion, partners of OSA patients exhibit lower sleep quality, poorer health perception, and reduced couple functioning compared to controls. While these findings should be interpreted cautiously due to the cross-sectional design, the absence of poor sleepers in the control group, and exploratory relational classifications, they emphasize the systemic nature of OSA, affecting both patients and partners. Emotionally responsive and cohesive relationships may serve as important resources, potentially buffering stress and supporting resilience. Future longitudinal research, more heterogeneous controls, and clearer operational definitions of relational categories are needed to disentangle condition-specific effects from sampling influences. Overall, these results reinforce the importance of conceptualizing OSA as a relational phenomenon, warranting assessment and intervention at the level of the couple, in addition to the patient.

Limitations

Several limitations should be acknowledged when interpreting the present findings. First, the cross-sectional design precludes any causal inference regarding the directionality of the associations observed among sleep quality, relationship functioning, and perceived health. Although significant correlations emerged, it remains unclear whether poorer sleep leads to relational strain and reduced well-being, whether relational difficulties exacerbate sleep disturbances, or whether these processes are reciprocally reinforced over time.

Second, the sample size was relatively modest and predominantly female, which may limit the generalizability of the findings. In addition, the control group unexpectedly included only participants classified as "good sleepers," despite sleep quality not being an inclusion criterion. Community-based samples generally include a non-negligible proportion of individuals with poor sleep quality thus, the present control group may not fully represent the variability of sleep quality observed in the general population. The observed differences in sleep quality should not be interpreted exclusively as reflecting the impact of OSA or CPAP treatment on partners, but also as potentially influenced by the unusually favorable sleep profile of the control group. Future studies should recruit more heterogeneous and representative comparison groups, including controls with different types of sleep quality, to better disentangle condition-related effects from sampling characteristics.

Moreover, the absence of correction for multiple statistical tests represents an additional limitation. This analytic choice increases the risk of type I error and consequently

weakens the generalizability of the significant findings. Accordingly, the results should be interpreted with caution and considered as important preliminary evidence that may guide future studies using larger samples, confirmatory designs, and appropriate correction procedures for multiple comparisons.

Third, sleep quality was assessed exclusively through self-report measures (PSQI), without objective sleep recordings. Although the PSQI is widely validated and has demonstrated good internal consistency in the present study, subjective sleep perception may not fully capture the complexity of sleep disturbances in partners of patients with OSA. Similarly, relationship functioning was assessed using self-report questionnaires, which may be influenced by response bias or social desirability.

Furthermore, the categorization of couple functioning (e.g., "functional distress") was exploratory and not based on clinically validated cut-offs specifically established for relational typologies. While this approach allowed for a descriptive differentiation of relational profiles, these classifications should be interpreted with caution and considered preliminary.

Despite these limitations, the study offers several important contributions. It addresses an underexplored population of partners of CPAP-treated OSA patients within a dyadic framework, integrating sleep quality, relationship functioning, and health-related QoL in a single model. The use of validated instruments (PSQI, DAS, SF-36) and the comparison with a demographically similar control group strengthen the methodological rigor. Most importantly, the findings contribute to a growing body of literature conceptualizing OSA as a "shared" condition embedded within a relational system. By highlighting the interconnected nature of sleep, health perception, and couple functioning, this study provides clinically relevant insights that support the development of couple-centered assessment and intervention strategies in sleep medicine.

Conclusions and Future Directions

In conclusion, our comparative analysis highlights a general decline across multiple domains, including sleep quality, health perception, and relationship functioning among partners of patients with OSAS relative to the control group. Although these findings must be interpreted cautiously in light of methodological limitations, the exploratory nature of relational classifications, and the cross-sectional design, they contribute to a growing body of evidence suggesting that obstructive sleep apnea should not be conceptualized solely as an individual clinical condition. Rather, OSA appears to extend its impact to the relational sphere, shaping partners' well-being and the overall quality of couple functioning.

The observed pattern underscores the interconnected nature of sleep, relational quality, and perceived health. While the present design does not allow causal conclusions, the findings suggest that relational functioning and individual psychological well-being are closely intertwined. In this context, emotionally responsive and cohesive relationships

may represent an important relational resource, potentially buffering stress and fostering resilience within the dyad.

From a clinical perspective, these results support the importance of adopting a dyadic approach in the management of OSAS. Including partners in assessment procedures, psychoeducation, and CPAP adherence strategies may enhance not only patient outcomes but also relational stability and mutual well-being.

Future research should employ longitudinal designs to clarify the directionality of associations between sleep disturbances and couple functioning, recruit more heterogeneous and representative control samples, and further refine the operationalization of relational profiles. The integration of objective sleep measures and dyadic-level analyses would also strengthen the understanding of how chronic sleep disruption affects couple dynamics over time. Advancing this line of inquiry may ultimately inform more comprehensive, couple-centered models of care in sleep medicine.

Funding: The authors have no funding to report.

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

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

Author Contributions: Conceptualization: GM, RS. Methodology: ESV. Software: GM, ESV, GL. Validation: LC, VLB, GLC. Formal analysis: GM, AR. Investigation: GM. Resources: LC, VLB, GLC. Data curation: GM, AR. Writing—original draft preparation: GM, AR. Writing—review and editing: VLB. GLC. Visualization: all authors. Supervision: RS. Project administration: AR. Funding acquisition: none. All authors have read and agreed to the published version of the manuscript.

Ethics Statement: The study was conducted in accordance with the Declaration of Helsinki and followed the 'Ethical Guidelines of the Italian Association of Psychology' and the Deontological Psychologist Ethics Code. Informed consent was obtained from all subjects involved in the study.

Preregistration: No preregistration exists for this study.

Reporting Guidelines: APA Guidelines for Psychological Assessment and Evaluation were followed [available link: APA Guidelines for Psychological Assessment and Evaluation]

Data Availability: The data that support the findings of this study are available from the corresponding author upon reasonable request. Materials: The materials used in this study are described in the Method section.

References

Apolone, G., & Mosconi, P. (1998). The Italian SF-36 health survey. *Journal of Clinical Epidemiology, 51*(11), 1025–1036. https://doi.org/10.1016/S0895-4356(98)00094-8

Baron, K. G., Smith, T. W., Berg, C. A., Czajkowski, L. A., Gunn, H., & Jones, C. R. (2011). Spousal involvement in CPAP adherence among patients with obstructive sleep apnea. *Sleep & Breathing / Schlaf & Atmung, 15*(3), 525–534. https://doi.org/10.1007/s11325-010-0374-z

Baron, K. G., Smith, T. W., Czajkowski, L. A., Gunn, H. E., & Jones, C. R. (2009). Relationship quality and CPAP adherence in patients with obstructive sleep apnea. *Behavioral Sleep Medicine, 7*(1), 22–36. https://doi.org/10.1080/15402000802577751

Beninati, W., Harris, C. D., Herold, D. L., & Shepard, J. W., Jr. (1999). The effect of snoring and obstructive sleep apnea on the sleep quality of bed partners. *Mayo Clinic Proceedings, 74*(10), 955–958. https://doi.org/10.1016/S0025-6196(11)63991-8

Bercea, R. M., Patacchioli, F. R., Ghiciuc, C. M., Cojocaru, E., & Mihaescu, T. (2013). Serum testosterone and depressive symptoms in severe OSA patients. *Andrologia, 45*(5), 345–350. https://doi.org/10.1111/and.12022

Billmann, S. J., & Ware, J. C. (2002). Marital satisfaction of wives of untreated sleep apneic men. *Sleep Medicine, 3*(1), 55–59. https://doi.org/10.1016/S1389-9457(01)00118-6

Broström, A., Nilsen, P., Johansson, P., Ulander, M., Strömberg, A., Svanborg, E., & Fridlund, B. (2010). Putative facilitators and barriers for adherence to CPAP treatment in patients with obstructive sleep apnea syndrome: A qualitative content analysis. *Sleep Medicine, 11*(2), 126–130. https://doi.org/10.1016/j.sleep.2009.04.010

Buysse, D. J., Reynolds, C. F., III, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4

Cartwright, R. (2008). Sleeping together: A pilot study of the effects of shared sleeping on adherence to CPAP treatment in obstructive sleep apnea. *Journal of Clinical Sleep Medicine, 4*(2), 123–127. https://doi.org/10.5664/jcsm.27128

Doherty, L. S., Kiely, J. L., Lawless, G., & McNicholas, W. T. (2003). Impact of nasal continuous positive airway pressure therapy on the quality of life of bed partners of patients with obstructive sleep apnea syndrome. *Chest, 124*(6), 2209–2214. https://doi.org/10.1378/chest.124.6.2209

Drager, L. F., Togeiro, S. M., Polotsky, V. Y., & Lorenzi-Filho, G. (2013). Obstructive sleep apnea. *Journal of the American College of Cardiology, 62*(7), 569–576. https://doi.org/10.1016/j.jacc.2013.05.045

Durmer, J. S., & Dinges, D. F. (2005). Neurocognitive consequences of sleep deprivation. *Seminars in Neurology, 25*(1), 117–129. https://doi.org/10.1055/s-2005-867080

Fava, C., Montagnana, M., Favaloro, E., Guidi, G., & Lippi, G. (2011). Obstructive sleep apnea syndrome and cardiovascular diseases. *Seminars in Thrombosis and Hemostasis, 37*(3), 280–297. https://doi.org/10.1055/s-0031-1273092

Gami, A. S., Howard, D. E., Olson, E. J., & Somers, V. K. (2005). Day–night pattern of sudden death in obstructive sleep apnea. *The New England Journal of Medicine, 352*(12), 1206–1214. https://doi.org/10.1056/NEJMoa041832

Gonzaga, C., Bertolami, A., Bertolami, M., Amodeo, C., & Calhoun, D. (2015). Obstructive sleep apnea, hypertension and cardiovascular diseases. *Journal of Human Hypertension, 29*(12), 705–712. https://doi.org/10.1038/jhh.2015.15

Gopalakrishnan, P., & Tak, T. (2011). Obstructive sleep apnea and cardiovascular disease. *Cardiology in Review, 19*(6), 279–290. https://doi.org/10.1097/CRD.0b013e318223bd08

Harding, S. M. (2000). Complications and consequences of obstructive sleep apnea. *Current Opinion in Pulmonary Medicine, 6*(6), 485–489. https://doi.org/10.1097/00063198-200011000-00004

Hasler, B. P., & Troxel, W. M. (2010). Couples' nighttime sleep efficiency and concordance: Evidence for bidirectional associations with daytime relationship functioning. *Psychosomatic Medicine, 72*(8), 794–801. https://doi.org/10.1097/PSY.0b013e3181ecd08a

Henry, J. D. (2016). Clinical assessment of social cognitive function in neurological disorders. *Nature Reviews Neurology, 12*(1), 28–39. https://doi.org/10.1038/nrneurol.2015.229

Hinz, A., Glaesmer, H., Brähler, E., Löffler, M., Engel, C., Enzenbach, C., Hegerl, U., & Sander, C. (2017). Sleep quality in the general population: Psychometric properties of the Pittsburgh Sleep Quality Index, derived from a German community sample of 9284 people. *Sleep Medicine, 30*, 57–63. https://doi.org/10.1016/j.sleep.2016.03.008

Hoy, C. J., Vennelle, M., Kingshott, R. N., Engleman, H. M., & Douglas, N. J. (1999). Can intensive support improve continuous positive airway pressure use in patients with the Sleep Apnea/Hypopnea Syndrome? *American Journal of Respiratory and Critical Care Medicine, 159*(4), 1096–1100. https://doi.org/10.1164/ajrccm.159.4.9808008

Khokhrina, A., Andreeva, E., & Degryse, J.-M. (2022). A systematic review on the association of sleep-disordered breathing with cardiovascular pathology in adults. *NPJ Primary Care Respiratory Medicine, 32*(1), Article 41. https://doi.org/10.1038/s41533-022-00307-6

Locke, B. W., Lee, J. J., & Sundar, K. M. (2022). OSA and Chronic Respiratory Disease: Mechanisms and epidemiology. *International Journal of Environmental Research and Public Health, 19*(9), Article 5473. https://doi.org/10.3390/ijerph19095473

Luyster, F. S., Dunbar-Jacob, J., Aloia, M. S., Martire, L. M., Buysse, D. J., & Strollo, P. J. (2016). Patient and partner experiences with Obstructive Sleep Apnea and CPAP Treatment: A qualitative analysis. *Behavioral Sleep Medicine, 14*(1), 67–84. https://doi.org/10.1080/15402002.2014.946597

McArdle, N., Devereux, G., Heidarnejad, H., Engleman, H. M., Mackay, T. W., & Douglas, N. J. (1999). Long-term use of CPAP therapy for sleep apnea/hypopnea syndrome. *American Journal of Respiratory and Critical Care Medicine, 159*(4), 1108–1114. https://doi.org/10.1164/ajrccm.159.4.9807111

McArdle, N., Kingshott, R. N., Engleman, H. M., Mackay, T. W., & Douglas, N. J. (2001). Partners of patients with sleep apnoea/hypopnoea syndrome: Effect of CPAP treatment on sleep quality and quality of life. *Thorax, 56*(7), 513–518. https://doi.org/10.1136/thx.56.7.513

Pankhurst, F. P., & Home, J. A. (1994). The influence of bed partners on movement during sleep. *Sleep, 17*(4), 308–315. https://doi.org/10.1093/sleep/17.4.308

Parish, J. M., & Lyng, P. J. (2003). Quality of life in bed partners of patients with obstructive sleep apnea or hypopnea after treatment with continuous positive airway pressure. *Chest, 124*(3), 942–947. https://doi.org/10.1378/chest.124.3.942

Peppard, P. E., Young, T., Barnet, J. H., Palta, M., Hagen, E. W., & Hla, K. M. (2013). Increased prevalence of sleep-disordered breathing in adults. *American Journal of Epidemiology, 177*(9), 1006–1014. https://doi.org/10.1093/aje/kws342

Revenson, T. A., & DeLongis, A. (2010). Couples coping with chronic illness. In S. Folkman (Ed.), *The Oxford handbook of stress, health and coping* (pp. 101-123). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195375343.013.0006

Røsand, G.-M. B., Slinning, K., Eberhard-Gran, M., Røysamb, E., & Tambs, K. (2012). The buffering effect of relationship satisfaction on emotional distress in couples. *BMC Public Health, 12*(1), Article 66. https://doi.org/10.1186/1471-2458-12-66

Schröder, C. M., & O'Hara, R. (2005). Depression and obstructive sleep apnea (OSA). *Annals of General Psychiatry, 4*, Article 13. https://doi.org/10.1186/1744-859X-4-13

Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 38*(1), 15–28. https://doi.org/10.2307/350547

Troxel, W. M., Buysse, D. J., Hall, M., & Matthews, K. A. (2009). Marital happiness and sleep disturbances in a multi-ethnic sample of middle-aged women. *Behavioral Sleep Medicine, 7*(1), 2–19. https://doi.org/10.1080/15402000802577736

Troxel, W. M., Robles, T. F., Hall, M., & Buysse, D. J. (2007). Marital quality and the marital bed: Examining the covariation between relationship quality and sleep. *Sleep Medicine Reviews, 11*(5), 389–404. https://doi.org/10.1016/j.smrv.2007.05.002

Virkkula, P., Bachour, A., Hytönen, M., Malmberg, H., Salmi, T., & Maasilta, P. (2005). Patient- and bed partner-reported symptoms, smoking, and nasal resistance in sleep-disordered breathing. *Chest, 128*(4), 2176–2182. https://doi.org/10.1378/chest.128.4.2176

Ware, J. E., & Sherbourne, C. D. (1992). The MOS 36-Item short-form health survey (SF-36). *Medical Care, 30*(6), 473–483. https://doi.org/10.1097/00005650-199206000-00002

Yaggi, H. K., Concato, J., Kernan, W. N., Lichtman, J. H., Brass, L. M., & Mohsenin, V. (2005). Obstructive sleep apnea as a risk factor for stroke and death. *The New England Journal of Medicine, 353*(19), 2034–2041. https://doi.org/10.1056/NEJMoa043104

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

Trusting the Psychotherapist, Trusting the Governmental System and Perhaps Trusting the Neighbor Next Door: Exploring the Institutional Nature of Trust in Psychotherapists and Examining Different Groups of Trusters in a Large German Patient Sample

Marie-Theresa Kaufmann[1], Christoph Kasinger[2], Ayline Heller[2,3], Elmar Brähler[2,4], Bernhard Strauß[1]

[1] *Institute for Psychosocial Medicine, Psychotherapy and Psychooncology, Friedrich Schiller University Jena, Jena, Germany.* **[2]** *Department of Psychosomatic Medicine and Psychotherapy, University Medical Center, Johannes Gutenberg University Mainz, Mainz, Germany.* **[3]** *Survey Design and Methodology, GESIS – Leibniz-Institute for the Social Sciences, Mannheim, Germany.* **[4]** *Department of Medical Psychology and Medical Sociology, University of Leipzig, Leipzig, Germany.*

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e19731, https://doi.org/10.32872/cpe.19731

Received: 2025-09-10 • **Accepted:** 2026-05-14 • **Published (VoR):** 2026-08-31

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

Corresponding Author: Bernhard Strauß, Institut für Psychosoziale Medizin, Psychotherapie und Psychoonkologie, Stoystr. 3, 07740 Jena, Germany. Tel.: + 49 3641 9 398020. E-mail: bernhard.strauss@med.uni-jena.de

Supplementary Materials: Materials [see Index of Supplementary Materials]

Abstract

Background: Trust in psychotherapists is considered primarily interpersonal. However, psychotherapists could be viewed as representatives of the (healthcare) system. In this study, we investigated whether trust in psychotherapists is more closely linked to interpersonal or institutional trust. Furthermore, we explored whether people can be distinguished into different trust groups.

Method: In a German cross-sectional study sample, we identified 1,300 persons (61.7% female; *M* = 48.46 years; *SD* = 14.45 years) indicating experiences with psychotherapy (current and/or in the past). Correlation and linear regression analysis were applied to explore connections between interpersonal (SOEP Interpersonal Trust Scale), institutional (ALLBUS Institutional Trust Scale), and trust in psychotherapists (IIV Psychotherapists). Then, hierarchical cluster analysis was used to identify profiles of interpersonal, institutional, and trust in psychotherapists. Group differences

of the cluster solution were analyzed for dimensions that are important for trust and psychotherapy such as having experienced trauma.

Results: Trust in psychotherapists is not solely associated with interpersonal factors but is also modestly associated with institutional trust. Furthermore, a three-cluster solution emerged with different trust profiles: *low trusters* (30%) with low scores for all trust measures, *high trusters* (49%) with the highest values, and *system trusters* (21%) with the highest trust scores in psychotherapists and institutions, but lower interpersonal trust values. The three groups showed different manifestations of e.g., trauma (LEC-5 Personal Experienced Subscale).

Conclusion: There is an institutional nature of trust in psychotherapists, which should be reflected and included in research and practice. Implications could include the common measurement of trust in psychotherapists in therapy and broadening psychotherapeutic interventions according to trust needs. To this end, new ways of building trust both in psychotherapy and institutions in general should be explored.

Keywords

trust in psychotherapists, psychotherapy, systematic change, institutional trust, interpersonal trust

Highlights

- Trust in psychotherapists should be considered partly institutional in nature.
- Trust in psychotherapists is associated with both interpersonal and institutional trust.
- Three distinct trust profiles emerged: low trusters, high trusters, and system trusters.

"[T]rust (or, symmetrically, distrust) is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action, both before he can monitor such action" (Gambetta, 2000, p. 216). Understanding trust in psychotherapists is crucial for assessing impact factors of psychotherapy (e.g., Gaebel et al., 2014; Strupp et al., 1964). Although trust in psychotherapists is often discussed as interpersonal (e.g., Kassebaum, 2004; Ozawa & Sripad, 2013), there is limited evidence to support this assumption. As psychotherapists operate within the public institution 'healthcare system', connections with institutional trust are conceivable. A comprehensive understanding of the interrelationships between trust in psychotherapists, institutions, and interpersonal trust is essential for developing effective approaches to fostering trust in psychotherapy.

Trust in Psychotherapists

Trust in psychotherapists plays a key role in help-seeking behavior (Cavanagh et al., 2022). It is also strongly linked to better treatment outcomes, particularly when patients regularly communicate with their psychotherapist about whether they feel respected and trust in their relationship (Connolly Gibbons et al., 2023). Trust ratings of psychotherapists in their patients and vice versa do not necessarily have to correlate, but

therapists' evaluations of being congruent and empathic correlate with patients' trust ratings (Peschken & Johnson, 1997). It remains questionable whether one can really speak of interpersonal trust between patient and practitioner. It may be a question of trust between two representatives of two groups who meet (Sousa-Duarte et al., 2020).

Trust measures can either refer to psychotherapists in general or to one's own psychotherapist with whom one is undergoing treatment. While there is a larger number of survey instruments available for measuring trust in one's own treating psychotherapist, which is a time critical event that must be completed during psychotherapy (e.g., Crits-Christoph et al., 2019), there is a lack of questionnaires that focus on general trust in psychotherapists (Kassebaum, 2004), which this study investigated.

Interpersonal Trust

The concept of interpersonal trust is understood as trust in strangers or trust in people in general (Beierlein et al., 2012; Naef & Schupp, 2009; Rotter, 1967; Yamagishi & Yamagishi, 1994). There are different definitions, which contain (a) a subject, (b) an action, and (c) an expectation concerning the future. "Trust, however, involves present decisions, often based on another person's past behavior, that require anticipating some action that hasn't yet happened" (Borum, 2010, p. 8). Lower levels of interpersonal trust are associated with reduced economic performance as lower income or higher rates of unemployment (Lichter et al., 2021; Nikolova et al., 2022). Individuals with higher interpersonal trust values are happier and more socially integrated (Rotter, 1980). Furthermore, greater interpersonal trust is linked to higher quality of life (Tokuda et al., 2008) and better mental health (Folkman et al., 1986).

Institutional Trust

Institutional trust is typically measured using scales that assess the level of trust in specific institutions (Hudson, 2006; Schupp & Wagner, 2004), which are evaluated by residents of a country (González & Smith, 2017). Institutional trust can be separated into three categories: political trust (e.g., in the parliament), trust in impartial institutions (e.g., the healthcare system), and trust in control institutions (e.g., the media). Among these, the healthcare system generally achieves the highest scores in comparison with other institutions, such as the media (Rothstein & Stolle, 2002). Trust is also a crucial factor in both private and social contexts, which are highly relevant for psychotherapy: Higher levels of institutional trust have been linked to enhanced social cohesion and mental health outcomes (Defeyter et al., 2021; van Prooijen et al., 2022). Moreover, institutional trust can act as a protective factor against the negative impact of adverse life events on well-being (Glatz & Eder, 2020; Lee, 2022). It also appears to play a significant role in the economic management of a country and for coping with crises (Chen et al., 2024; Eitze et al., 2021; Lichter et al., 2021; Reinemann et al., 2022).

Trust in Psychotherapists, Interpersonal Trust, and Institutional Trust – Considerations About Connections

Studies on the measurement of trust in (healthcare) systems, healthcare practitioners and the trust between practitioners, and patients primarily rely on questionnaires focused on the patient-practitioner relationship (Aboueid et al., 2023; Ozawa & Sripad, 2013). Thereby, these instruments neglect to address system trust (Ozawa & Sripad, 2013).

Notably, interpersonal and institutional trust appear to be distinct concepts as they differ according to their relationship to other constructs and attitudes (Krastev et al., 2023) and can be distinguished in factor analysis (Naef & Schupp, 2009). A causal influence of institutional trust on interpersonal trust has been proposed (Sønderskov & Dinesen, 2016). On the other hand, interactions with practitioners are still considered an interpersonal factor in the development of institutional trust (Campos-Castillo et al., 2016). It is important to clarify these connections as they affect the effectiveness of psychotherapy.

Several factors identified in the literature as influencing the utilization of psychotherapy overlap with those associated with higher levels of institutional trust: The sociodemographic factors such as biological sex (female), higher education, not living together with a partner, income and age appear to be relevant in understanding the likelihood of psychotherapy utilization (e.g., Hahm et al., 2020; Petrowski et al., 2014; Vessey & Howard, 1993). Similarly, the sociodemographic factors gender, high school diploma, age and income have been identified as significant factors for institutional trust (Baroudi et al., 2022; Campbell, 2004; Campbell, 2023). This may indicate that there are the same influencing factors and possibly conceptual similarities of trust in psychotherapists and institutions. Since the present cross-sectional study design is exploratory, further relationships will be recorded here. Psychotherapy is embedded within the healthcare system and is regulated by institutional frameworks in Germany. As such, trust in psychotherapists may not only reflect interpersonal evaluations but also institutional trust that governs and provides psychotherapy.

Linking trust in psychotherapeutic settings to institutional trust could offer valuable insights into the factors influencing therapy effectiveness and accessibility. Understanding this connection can help identify barriers to treatment, improve patient engagement, and highlight the broader societal impact of institutional trust on health behaviors.

In this study, we aimed to investigate the relationships between the three different forms of trust: in psychotherapists, others, and institutions. Based on the literature, we assumed a stronger connection of trust in psychotherapists to institutional trust than to interpersonal trust.

Patients come to therapy with varying backgrounds, including experiences of trauma, resilience, personality traits, and beliefs. For this reason, the results were further contextualized in terms of factors relevant to psychotherapy: Associations between trust patterns and traumatic life events (Goenjian et al., 1997) were also investigated, as they

contribute to psychological burden and psychotherapy usage. Differences in psychological resilience factors, such as general resilience (Bartholomew et al., 2022), belief in a just world (Furnham, 2003), and willingness to forgive others – a resilience factor for traumatic life events (Wade et al., 2005) – were examined. Personality traits were also analyzed as introversion and neuroticism may be risk factors for psychological disorders (Zinbarg et al., 2008).

Method

Study Design, Setting, and Participants

An online cross-sectional study was conducted in Germany in December 2022 by two projects[1], which were funded by the German 'Federal Ministry of Research, Technology and Space' ('Bundesministerium für Forschung, Technologie und Raumfahrt'; BMFTR). The study was approved by the Ethics Committee of the University Hospital Jena (Reg.-Nr. 2022-2817-Bef). The analyzed questions were part of an extensive survey in which a series of health-related and psychological data were collected. Around 2000 East Germans, 2000 West Germans, and 500 internal migrants from East to West Germany, over the age of 18 up to 74, were contacted via the panel provider bilendi & respondi. The provider sent out three separate invitations for the three groups (East Germans, West Germans, and internal migrants). For the last group, a low frequency in the population (Heller et al., 2020) was expected, and a correspondingly high number of invitations was necessary. The participating panelists were unaware that the aforementioned groups were being targeted. They began the survey and were excluded based on their feedback regarding the relevant sociodemographic factors if they did not meet these criteria. They received financial compensation. Consequently, the response rates varied between 0.1% and 9%. Quotas on gender (female/male) and age (uncrossed) were only applied to East and West Germans, not to internal migrants, due to the low frequency of internal migrants in the population. There were no missing values, as the participants could only continue and complete the questionnaire if they filled out all fields. A total number of n = 170 people were excluded due to unrealistic values (number of people in the household > 10; frequency of one's own unemployment > 10). This resulted in a sample of N = 4,619 people (Table S1 Appendix). In our analysis, only n = 1,300 people that experienced psychotherapy after the fall of the Berlin wall and the end of the Cold War in the year 1990 were included (Table 1). We did not consider those with psychotherapy experience before the end of the Cold War 1990, as Germany was divided into two completely

1) The two projects 'SiSaP' (University Hospital Jena) and 'DDR-Psych' (University Hospital Mainz) investigated the history of psychotherapy in the GDR (German Democratic Republic) and the psychological consequences of the former GDR.

different governmental systems with different healthcare systems in these decades (see section Psychotherapy Experience).

Table 1

Subsample Description of the Cross-Sectional Study (Online)

Variable	Participants (n = 1,300)
Female gender	n = 799 (61.7%)
Age	M = 48.46 (SD = 14.45)
Home ownership	n = 449 (34.5%)
Interpersonel trust*	M = 6.54 (SD = 1.63)
Trust in institutions**	M = 3.77 (SD = 1.27)
Trust in psychotherapists***	M = 3.86 (SD = 0.73)

Note. The relative proportions refer to the total number of the sub-sample. Reflects M and SD of participants showing trust from a scale from 3 to 12*, 1 to 7**, 1 to 5***.

Instruments

Psychotherapy Experience

The psychotherapy variable (binary: yes/no) was generated based on the response to the following question: '*Have you ever received psychotherapy as an inpatient in a clinic or as an outpatient, without a stay in a hospital?*' The response options were '*no*', '*yes, started and already completed*', *'yes, started and discontinued*' and *'yes, currently*'. A distinction was made between '*no*' and '*yes*' answers. For people who were born before the 1st of January 1980[2] the question was asked twice – once related to the Cold War period when Germany was still divided and once to the period after 1990. In conducting the calculations, only those therapeutic experiences were used that occurred subsequent to the fall of the Berlin wall and the end of the Cold War in 1990 because of system comparability. In total n = 1,300 received psychotherapy after the end of the Cold War.

Trust Questionnaires

Institutional trust was measured by the questionnaire of the German General Social Survey (ALLBUS; GESIS, 2018). It contains 13 categories such as the healthcare system, the German parliament or television and is a validated instrument (Naef & Schupp, 2009). For each category, the participant rates on a 7-point scale how much he or she trusts in it (*1 not at all*; *7 very much*) – an averaged total value between 1 and 7 was calculated. Higher values indicate higher levels of trust. The scale showed excellent internal consistency (α = .95), indicating a highly reliable measure.

2) This was a distinction measure for people that experienced the Cold War (similar to Ulke et al., 2021).

Interpersonal trust was measured with the validated three item scale of the German Socio-Economic Panel Study with questions such as *"In general, you can trust people"* (SOEP; Naef & Schupp, 2009; *1 not at all; 4 very much*). Although brief, the SOEP measure is widely used in large-scale studies and provides an efficient assessment of interpersonal trust. A sum score was calculated with values ranging between 3 and 12. Higher values indicate higher levels of trust. The scale showed moderate internal consistency (α = .60), which may reflect the conceptual heterogeneity of the construct. Given the exploratory nature of the study, this level of reliability was considered acceptable.

Trust in psychotherapists was measured with the validated scale for trust in psychotherapists in Kassebaum's (2004) questionnaire including further trust categories such as trust in neighbours which were not used. The questionnaire asks e.g., *"Psychotherapists can be a great help in serious crises"* (*1 NO - - I completely disagree with this statement.; 5 YES + + I completely agree with this statement.*). A mean value was calculated with results ranging between 1 and 5. Higher values indicate higher levels of trust. The scale demonstrated good internal consistency (α = .79), indicating a reliable measurement of the construct.

Classification Questionnaires for the Trust-Profiles

Potentially traumatic life events were assessed with the Life Event Checklist (LEC-5; Gray et al., 2004; Weathers et al., 2013). The questionnaire comprises 16 events, which may be experienced by the individual, witnessed or heart about by the individual or with the individual experiencing it through the job. In addition, there is the possibility of assessing uncertainty regarding whether an event has been experienced or not and also reporting being not affected. A sum score of events which happened to oneself personally was calculated.

Belief in a just world was measured with the validated questionnaire of Dalbert and colleagues (1987; Schmitt et al., 2008) comprising six items on a scale from *1* (disagreeing) up to *6* (agreeing) – indicating more belief with a higher sum score.

The willingness to forgive was measured with the validated questionnaire of Allemand and colleagues (2008). Two subscales are utilized to ascertain an individual's willingness to forgive another person who has (not) expressed regret for their actions on a scale from *1* (not at all) up to *5* (very much) – indicating more willingness with a higher score. A total score was calculated.

Resilience was measured with the validated RS-5 scale comprising five items (Wagnild & Young, 1993; Schmalbach et al., 2016). Higher values indicate higher levels of resilience.

Personality factors were measured with the validated Big Five Inventory BFI-10 with two items for each category (openness, extraversion, agreeableness, conscientiousness and neuroticism). One item of each category is inverted. Each item had to be answered

on a five-point scale from *1* (disagreeing) up to *5* (agreeing; Rammstedt & John, 2007). A mean score was calculated.

Statistical Analysis

The analyses were carried out using the statistical software IBM SPSS (Version 21). Analyses concerning gender issues included only male and female respondents due to the small number of people who selected the category 'diverse' or 'other'. Results were considered significant at $\alpha = 0.05$.

Identification of Connections Between Interpersonal, Institutional and Trust in Psychotherapists

Correlational and linear regression analyses were applied to identify connections between the three trust scores. Correlational analysis included in the first analysis trust in psychotherapist and institutional trust and in a second analysis trust in psychotherapists and interpersonal trust. The linear regression analysis to predict trust in psychotherapists included institutional trust an interpersonal trust.

Identification of Pattern-Based Subcategories With Different Trust-Profiles and Validation of Trust-Profiles

Trust in psychotherapists, institutional trust, and interpersonal trust were used as input for an agglomerative hierarchical clustering algorithm (HCA). HCA used the Ward error sum-of-squares agglomeration method with Euclidean distance between individuals. Following these steps, individuals who have experienced psychotherapy were categorized into pattern-based subcategories (clusters). The number of clusters was determined by inspecting the dendrogram, with particular attention to increases in linkage distances between successive fusion steps. A marked increase in these distances suggested a three-cluster solution, indicating that further merging would substantially reduce within-cluster homogeneity. To assess the separation of the clusters, a discriminant analysis was subsequently performed. The differences between clusters were statistically significant ($p < .001$; Table S2 Appendix).

Classification of Trust-Profiles

Kruskal-Wallis-Tests were mainly used for the calculations, only for the Big Five personality traits openness and agreeableness were ANOVAs used. Pairwise comparisons were calculated with the Dunn test (Kruskal-Wallis-Test) and the Scheffé procedure (ANOVA).

Results

Identification of Connections Between Interpersonal, Institutional, and Trust in Psychotherapists

Trust in psychotherapists and institutional trust correlated on a significant level of $r = 0.259$ ($p < .001$). Trust in psychotherapists and interpersonal trust also correlated significantly $r = 0.226$ ($p < .001$). Multiple linear regression analysis ($R^2 = 0.089$ (adj. $R^2 = 0.088$); $F(2, 1297) = 63.64$, $p < .001$) confirmed the relevance of both, with institutional trust showing a slightly stronger connection as standardized β reached a higher value (Table 2).

Table 2

Multiple Linear Regression Model

	Multiple linear regression model (n = 1,300)		
Variable	***B* (95% CI)**	**β**	***p***
Constate term	2.985 (2.781; 3.188)		< .001***
Interpersonal trust	0.085 (0.055; 0.114)	0.158	< .001***
Trust in institutions	0.142 (0.104; 0.180)	0.207	< .001***

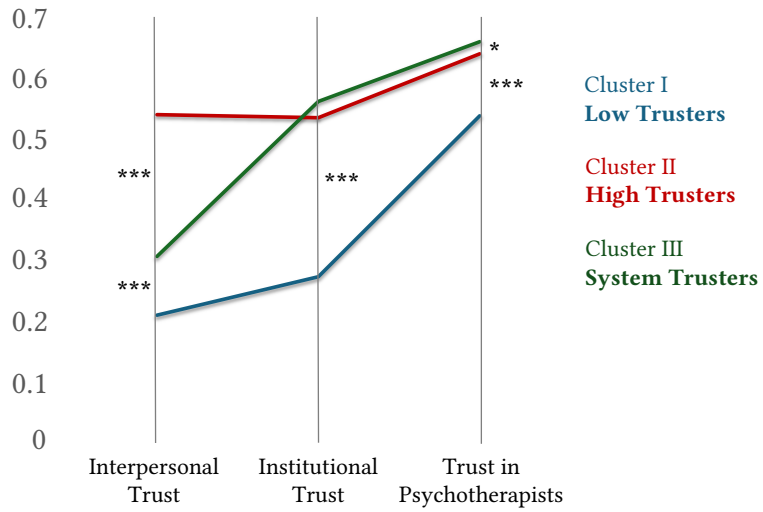
Note. B = unstandardized coefficient; β = standardized coefficient.
***$p < .001$.

Pattern-Based Subcategories With Different Trust-Profiles Among Participants That Have Experienced Psychotherapy

The hierarchical clustering resulted in a 3-cluster solution for those participants that have gone through psychotherapy. The trust-profiles for each cluster can be seen in Figure 1. Each cluster of the exploratory study design showed its own trust-profile. Based on the mean values of the three trust categories (trust in psychotherapists, institutional trust, interpersonal trust) the clusters 1-3 can be described the following way: *low trusters* (Cluster I; n = 391, 30%) with low scores for all trust measures, *high trusters* (n = 642, 49%) with highest interpersonal trust values and high trust values in psychotherapists and institutions and *system trusters* (n = 267, 21%) with highest trust scores in psychotherapists and high levels of institutional trust, but lower interpersonal trust values.

Figure 1

Trust Profiles per Cluster for Psychotherapy Patients



Note. The mean scores per cluster were standardized: $(m_{Cluster} - Min)/(Max - Min)$.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Validation of Trust-Profiles

In order to assess the separation between the clusters, a discriminative analysis was examined and confirmed the three clusters. A total of two discriminant functions was obtained for the three cluster categories. Their Wilk's λ values were 0.229 and 0.821, all reaching significant levels with $p < .001$. The eigenvalue of the first discriminant function was 2.582, which could explain 92.2% of the variance. The eigenvalue of the second discriminant function was 0.219 and explained 7.8% of the variance (Table 3). On the two discriminant functions, the structural load of institutional trust and interpersonal trust were highest (Table 4). The centroid of each cluster in the discriminant function is shown in Table 5. The reclassification results reached an accuracy rate of 87.2% (Table 6). In sum, the discriminant analysis indicated a clear separation between the three clusters but should be interpreted as a descriptive assessment rather than an independent validation of the cluster solution.

Table 3

Discriminant Function Significance Test for the Three Trust Clusters of the Hierarchical Cluster Analysis (HCA)

Function	Eigenvalues	Explained Variance	Cumulative Variation	ρ^2	Wilk's λ
1	2.582	92.2%	92.2%	0.849	0.229***
2	0.219	7.8%	100%	0.424	0.821***

Note. Function 1 represents the first discriminant function, function 2 represents the second discriminant function.
***$p < .001$.

Table 4

Structural Loadings in Trust Dimensions for the Two Cluster Discriminant Functions

	Cluster Discriminant Function	
Variable	**1**	**2**
Interpersonal trust	0.933	-0.389
Trust in psychotherapists	-0.024	0.389
Trust in institutions	0.443	0.788

Note. Function 1 represents the first discriminant function, function 2 represents the second discriminant function.

Table 5

Functions at Group Centroids

	Functions at Group Centroids	
Cluster	**1**	**2**
Cluster I: Low Trusters	-2.134	-0.349
Cluster II: High Trusters	1.524	-0.164
Cluster III: System Trusters	-0.539	0.905

Note. Function 1 represents the first discriminant function, function 2 represents the second discriminant function.

Table 6

Classification Accuracy of the Three Cluster Solution

	Forecast Cluster			
Cluster Category	**1**	**2**	**3**	**Total**
1. Low Trusters	347 (89%)	0 (0%)	44 (11%)	391
2. High Trusters	8 (1%)	610 (95%)	24 (4%)	642
3. System Trusters	36 (14%)	55 (21%)	176 (66%)	267
Classification accuracy		87.2%		

Note. Function 1 represents the first discriminant function, Function 2 represents the second discriminant function, Function 3 represents the third discriminant function. Rows represent the original cluster assignments derived from the cluster analysis, whereas columns represent the cluster membership predicted by the discriminant analysis. Cell entries show the number of participants and the corresponding row percentages. Values on the diagonal indicate correctly classified participants; off-diagonal values indicate misclassifications. Overall, 87.2% of participants were correctly classified.

Classification of Trust-Profiles

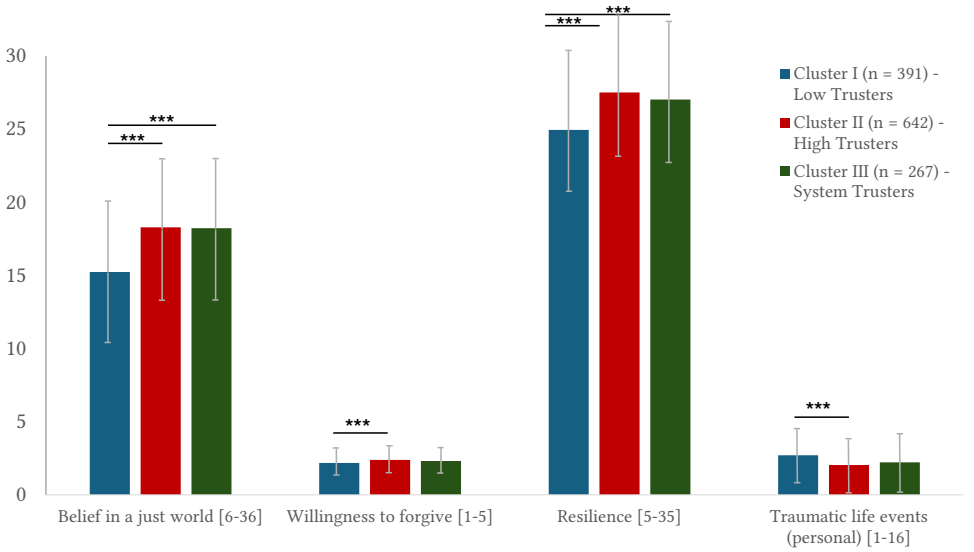
The three trust groups differed according to all of the classification variables – expect conscientiousness (Table S3 Appendix): System trusters were less willing to forgive others, were less agreeable and extroverted, than the high trusters, who reached highest scores in these categories, but showed higher values than the low trusters. Both system trusters and high trusters more often believed in a just world. High trusters showed higher resilience scores and experienced less critical and potentially traumatic life events in comparison with the low trusters. The high trusters reached the highest score for openness and low trusters showed higher values of neuroticism (Figure 2 & Figure 3).

Discussion

In this study, we investigated whether trust in psychotherapists is more closely linked to interpersonal or institutional trust. Our findings build on previous research, which predominantly considers trust in psychotherapists as an interpersonal construct (e.g., Ozawa & Sripad, 2013; Kassebaum, 2004). Our findings suggest that trust in psychotherapists is not solely associated with interpersonal factors but is also modestly connected with institutional trust, indicating that broader systemic attitudes may play a role in shaping patients' trust in psychotherapy. Based on HCA, we could identify three groups depending on the levels of institutional trust, trust in psychotherapists, and interpersonal trust: there are patients who show high levels in all three types of trust (high trusters), persons with low trust values in general (low trusters), and patients who trust the system and psychotherapists, but not others (system trusters).

Figure 2

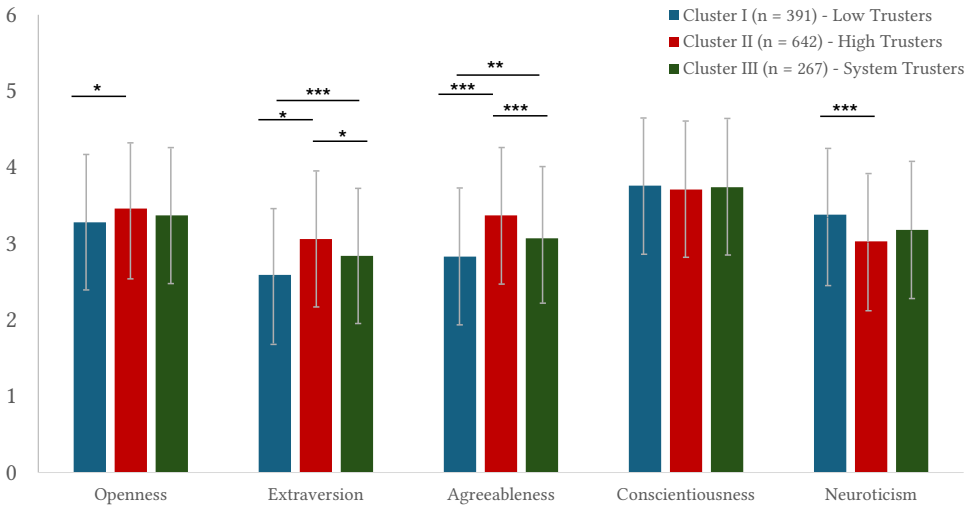
Belief in a Just World, Willingness to Forgive, Resilience and Trauma in the Three Trust Groups



Note. Pairwise comparisons. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Figure 3

Personality Traits in the Three Trust Groups



Note. Pairwise comparisons of the Big Five Personality Traits. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

While the clusters were derived exclusively from trust variables, subsequent analyses examining external variables revealed differences between clusters in terms of personality characteristics and life experiences. These findings provide preliminary support for the broader psychological relevance of the identified trust profiles. The three clusters of trust patterns exhibited partly different character traits and life experiences associated with psychotherapy usage and institutional trust. However, these findings should be interpreted with caution, as the cluster solution itself is based solely on trust variables. There may be several key starting points for psychotherapists in this context. The following considerations regarding therapeutic implications are speculative and should be interpreted with caution, as they are not directly tested in the present study. And it is important to distinguish between general trust in psychotherapists and trust in one's own psychotherapist. The present study focuses on general trust, which reflects attitudes toward the profession rather than the quality of a specific therapeutic relationship. Firstly, it is important to recognize that, from the patient's perspective, psychotherapists also represent institutions within which they work. Consequently, a trusting patient may also trust the system as a whole. For system trusters who trust psychotherapists but not others, additional support might be relevant to help them build new trusting relationships in private contexts, as observed in cases of psychosis patients (Ridenour et al., 2024) or young patients with suicide ideation (Hill et al., 2019). Interventions such as social skills training within the therapeutic context could be relevant for this group. In addition, there are patients who not only trust psychotherapists but also others and the political system – either as a result of psychotherapy or as a preexisting trait. For this group of high-trusters, potential risk factors for the emergence of psychiatric symptoms in the future may include impending critical life events or challenging circumstances, but at first, it is necessary to examine potential risk factors in longitudinal research. At the same time, practitioners may encounter patients who exhibit low trust overall (low-trusters). Low-trusters seem to be an especially vulnerable group: They experienced more traumatic life events, showed less resilience, were less open, extroverted, agreeable and reached higher levels of neuroticism. They were less willing to forgive others and believed less in a just world. For these patients, it could be relevant to foster trust within the therapeutic relationship. Additionally, the development of resilience factors, such as social participation, establishing effective coping strategies, and cultivating personal interests, may prove beneficial, too. These potential therapeutic implications should be interpreted with caution, as they are not directly tested in the present study and remain speculative.

Further investigations on psychotherapeutic relationships are warranted to prove distinct connections. Future research could examine whether cluster affiliation is associated with specific psychiatric outcomes and should take into account that trust in psychotherapists is associated with institutional trust. This is an important aspect regarding the conceptualization of trust in psychotherapists.

These findings align with the recommendations of Gaebel and colleagues (2014) to preserve trust in the healthcare system and psychiatrists. Our findings broaden the range of strategies for enhancing trust in psychotherapists and, in turn, may also improve the effectiveness of psychotherapy. We recommend the regular evaluation of trust in psychotherapy.

Limitations

The representativeness of this study is limited due to the non-random sampling and the inherent characteristics of panel surveys. The results of these correlational analyses suggest the potential for the formulation of causal hypotheses regarding trust in psychotherapists, which should be investigated further. The analysis can only be described as exploratory. Given the diverse definitions of trust in psychotherapists, it cannot be ruled out that different questionnaires might yield varying cluster solutions. A combined analysis incorporating multiple trust questionnaires could provide further information and strengthen our findings. Furthermore, a limitation of this study is that it assesses general trust in psychotherapists rather than trust in a specific therapist, which may differ in important ways. Also, the concept of epistemic trust – the capacity to understand others' and one's personal behavior according to mental states, which is relevant in the psychotherapeutic relationship – was not measured (Fonagy & Allison, 2014). In addition, the identified clusters represent statistical groupings based on the included trust variables and should not be interpreted as clearly distinct psychological types. Besides, the data used is only based on self-reports. Furthermore, the experience of psychotherapy was quite broad: It included participants who are currently in therapy, have completed therapy, or have discontinued therapy. These experiences can vary and may be related to trust in psychotherapists. Finally, the results refer to the German healthcare system, which is characterized by relatively high institutional regulation and broad access to psychotherapeutic services. Given cultural differences, no generalizations can be made. These structural features may contribute to comparatively higher levels of institutional trust, which in turn could influence trust in psychotherapists. Research suggests that the relative importance of interpersonal versus institutional trust may vary across cultural contexts (e.g., Sikorski & Albrecht, 2025). Therefore, the generalizability of the present results to other countries and healthcare systems may be limited and should be examined in future cross-cultural studies.

Conclusion

Trusting the psychotherapist may imply trusting the system: This perspective emphasizes the need for an adjustment of the conceptualization of psychotherapy as a relationship embedded within an institutional framework. In the therapeutic relationship, a third entity – the institution or system – plays a role alongside the patient and the

psychotherapist. This underscores the importance of evaluating not only the patients' trust in the psychotherapists but also their perception of the (healthcare) system as perceived by both the psychotherapist and the patient. This approach may help optimize the efficacy of psychotherapeutic treatments.

Funding: The study was funded as part of the BMBF-funded joint project 'Seelenarbeit im Sozialismus' (SiSaP), subproject Jena (FKZ01UJ1908AY) and 'DDR-Vergangenheit und psychische Gesundheit: Risiko- und Schutzfaktoren' (DDR-PSYCH; FKZ01UJ1911AY), Mainz.

Acknowledgments: We thank all participants for their contribution to the study.

Competing Interests: The authors declare that they have no conflict of interest.

Author Contributions: All authors designed the analyses. MTK wrote the report, developed and conducted the research question and all the analysis strategies. AH, CK, EB and BS contributed to reviewing and commenting on the report and approved the final content. EB and BS are the principal investigators of the entire research project.

Ethics Statement: The survey was approved by the data protection and ethics committee of Jena University Hospital (Reg.-Nr. 2022-2817-Bef)., in accordance with the Declaration of Helsinki. Participation was voluntarily and anonymous. Respondents gave written and informed consent before starting the questionnaire.

Preregistration: The study was not preregistered.

Reporting Guidelines: This study was conducted in accordance with the JARS–Quant Guidelines.

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

Supplementary Materials

The Supplementary Materials contain the following items (for access, see Kaufmann et al., 2026S):

- Table S1. Sample description of the cross-sectional study (online)
- Table S2. Profiles of the three clusters of the hierarchical cluster analysis (HCA)

Index of Supplementary Materials

Kaufmann, M., Kasinger, C., Heller, A., Brähler, E., & Strauß, B. (2026S). *Supplementary materials to "Trusting the psychotherapist, trusting the governmental system and perhaps trusting the neighbor next door: Exploring the institutional nature of trust in psychotherapists and examining different groups of trusters in a large German patient sample"* [Appendix]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.22342

References

Aboueid, S. E., Herati, H., Nascimento, M. H., Ward, P. R., Brown, P. R., Calnan, M., Perlman, C. M., & Meyer, S. B. (2023). How do you measure trust in social institutions and health professionals? A systematic review of the literature (2012–2021). *Sociology Compass, 17*(9), Article e13101. https://doi.org/10.1111/soc4.13101

Allemand, M., Sassing-Meng, A., Huber, S., & Schmitt, M. (2008). Entwicklung und Validierung einer Skala der Bereitschaft zu verzeihen [Development and validation of a scale of willingness to forgive]. *Diagnostica, 54*(2), 71–84. https://doi.org/10.1026/0012-1924.54.2.71

Baroudi, M., Goicolea, I., Hurtig, A.-K., & San-Sebastian, M. (2022). Social factors associated with trust in the health system in northern Sweden: A cross-sectional study. *BMC Public Health, 22*(1), Article 881. https://doi.org/10.1186/s12889-022-13332-4

Bartholomew, T. T., Robbins, K. A., Joy, E. E., Kang, E., & Maldonado Aguiñiga, S. (2022). Clients' resilience and distress in psychotherapy: A preliminary meta-analysis. *Counselling Psychology Quarterly, 35*(2), 344–362. https://doi.org/10.1080/09515070.2020.1777389

Beierlein, C., Kemper, C. J., Kovaleva, A., & Rammstedt, B. (2012). *Kurzskala zur Messung des zwischenmenschlichen Vertrauens: Die Kurzskala Interpersonales Vertrauen (KUSIV3)* [Short scale for measuring interpersonal trust: The Short Interpersonal Trust Scale (KUSIV3)] (GESIS Working Papers, 2012/22). GESIS. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-312126

Borum, R. (2010). *The science of interpersonal trust* (Mental Health Law & Policy Faculty Publications, 574). https://digitalcommons.usf.edu/mhlp_facpub/574

Campbell, R. (2023). Germany united? Trust in democratic institutions thirty years after Unification. *German Politics and Society, 41*(1), 1–24. https://doi.org/10.3167/gps.2023.410101

Campbell, W. R. (2004). The sources of institutional trust in East and West Germany: Civic culture or economic performance? *German Politics, 13*(3), 401–418. https://doi.org/10.1080/0964400042000287437

Campos-Castillo, C., Woodson, B. W., Theiss-Morse, E., Sacks, T., Fleig-Palmer, M. M., & Peek, M. E. (2016). Examining the relationship between interpersonal and institutional trust in political and health care contexts. In E. Shockley, T. M. S. Neal, L. M. PytlikZillig, & B. H. Bornstein (Eds.), *Interdisciplinary perspectives on trust* (pp. 99-115). Springer. https://doi.org/10.1007/978-3-319-22261-5_6

Cavanagh, D., Jurcik, T., & Charkhabi, M. (2022). How does trust affect help-seeking for depression in Russia and Australia? *International Journal of Social Psychiatry, 68*(8), 1561–1570. https://doi.org/10.1177/00207640211039253

Chen, G., Yao, Y., Zhang, Y., & Zhao, F. (2024). The impact of risk perception and institutional trust on COVID-19 vaccine hesitancy in China. *Human Vaccines & Immunotherapeutics, 20*(1), Article 2301793. https://doi.org/10.1080/21645515.2024.2301793

Connolly Gibbons, M. B., Duong, L., Zoupou, E., Kashden, J., Fisher, J., & Crits-Christoph, P. (2023). Randomized controlled trial evaluating feedback to community-based therapists based on patient reports of trust and respect. *Journal of Consulting and Clinical Psychology, 91*(6), 337–349. https://doi.org/10.1037/ccp0000807

Crits-Christoph, P., Rieger, A., Gaines, A., & Gibbons, M. B. C. (2019). Trust and respect in the patient-clinician relationship: Preliminary development of a new scale. *BMC Psychology, 7*, Article 91. https://doi.org/10.1186/s40359-019-0347-3

Dalbert, C., Montada, L., & Schmitt, M. (1987). Glaube an eine gerechte Welt als Motiv: Validierungskorrelate zweier Skalen [Belief in a just world as a motive: Validation correlates of two scales]. *Psychologische Beiträge, 29*, 596–615.

Defeyter, M. A., Stretesky, P. B., Long, M. A., Furey, S., Reynolds, C., Porteous, D., Dodd, A., Mann, E., Kemp, A., Fox, J., McAnallen, A., & Gonçalves, L. (2021). Mental well-being in UK higher education during COVID-19: Do students trust universities and the government? *Frontiers in Public Health, 9*, Article 646916. https://doi.org/10.3389/fpubh.2021.646916

Eitze, S., Felgendreff, L., Korn, L., Sprengholz, P., Allen, J., Jenny, M. A., Wieler, L. W., Thaiss, H., De Bock, F., & Betsch, C. (2021). Vertrauen der Bevölkerung in staatliche Institutionen im ersten Halbjahr der Coronapandemie: Erkenntnisse aus dem Projekt COVID-19 Snapshot Monitoring (COSMO) [Public trust in state institutions during the first half of the coronavirus pandemic: Findings from the COVID-19 Snapshot Monitoring (COSMO) project]. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz, 64*(3), 268–276. https://doi.org/10.1007/s00103-021-03279-z

Folkman, S., Lazarus, R. S., Gruen, R. J., & DeLongis, A. (1986). Appraisal, coping, health status, and psychological symptoms. *Journal of Personality and Social Psychology, 50*(3), 571–579. https://doi.org/10.1037/0022-3514.50.3.571

Fonagy, P., & Allison, E. (2014). The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy, 51*(3), 372–380. https://doi.org/10.1037/a0036505

Furnham, A. (2003). Belief in a just world: Research progress over the past decade. *Personality and Individual Differences, 34*(5), 795–817. https://doi.org/10.1016/S0191-8869(02)00072-7

Gaebel, W., Muijen, M., Baumann, A. E., Bhugra, D., Wasserman, D., Van der Gaag, R. J., Heun, R., & Zielasek, J. (2014). EPA guidance on building trust in mental health services. *European Psychiatry, 29*(2), 83–100. https://doi.org/10.1016/j.eurpsy.2014.01.001

Gambetta, D. (2000). Can we trust trust? In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations* (pp. 213-237). University of Oxford.

GESIS – Leibniz-Institut für Sozialwissenschaften. (2018). *ALLBUS – Allgemeine Bevölkerungsumfrage der Sozialwissenschaften. ALLBUS 2018. Fragebogendokumentation. Material zu den Datensätzen der Studiennummern ZA5270 und ZA5271* [ALLBUS – German General Social Survey. ALLBUS 2018. Questionnaire Documentation. Materials relating to the datasets for study numbers ZA5270 and ZA5271]. https://www.gesis.org/allbus/inhalte-suche/frageboegen

Glatz, C., & Eder, A. (2020). Patterns of trust and subjective well-being across Europe: New insights from repeated cross-sectional analyses based on the European social survey 2002–2016. *Social Indicators Research, 148*(2), 417–439. https://doi.org/10.1007/s11205-019-02212-x

Goenjian, A. K., Karayan, I., Pynoos, R. S., Minassian, D., Najarian, L. M., Steinberg, A. M., & Fairbanks, L. A. (1997). Outcome of psychotherapy among early adolescents after trauma. *American Journal of Psychiatry, 154*(4), 536–542. https://doi.org/10.1176/ajp.154.4.536

González, S., & Smith, C. (2017). *The accuracy of measures of institutional trust in household surveys: Evidence from the OECD trust database* (OECD Statistics Working Papers, No. 2017/11). OECD Publishing. https://doi.org/10.1787/18152031

Gray, M. J., Litz, B. T., Hsu, J. L., & Lombardo, T. W. (2004). Psychometric properties of the life events checklist. *Assessment, 11*(4), 330–341. https://doi.org/10.1177/1073191104269954

Hahm, S., Speerforck, S., Fleischer, T., Grabe, H. J., Beutel, M., & Schomerus, G. (2020). Die Effekte von Geschlecht, Bildung und Einkommen auf antizipierte Scham bei psychischen Erkrankungen – Ergebnisse einer deutschen Bevölkerungsstudie [The effects of gender, education, and income on anticipated shame regarding mental illness – Results of a German population study]. *Psychiatrische Praxis, 47*(3), 142–147. https://doi.org/10.1055/a-1081-7614

Heller, A., Tibubos, A. N., Beutel, M., & Brähler, E. (2020). *Mauer in den Köpfen? Einstellungen zur deutschen Einheit im Wandel* [A wall in the mind? Changing attitudes towards German Unity] (OBS-Arbeitspapier 42). Otto Brenner Stiftung. https://archiv.otto-brenner-stiftung.de/mauer-in-den-koepfen-deutsche-einheit/

Hill, R. M., Penner, F., Vanwoerden, S., Mellick, W., Kazimi, I., & Sharp, C. (2019). Interpersonal trust and suicide ideation among adolescent psychiatric inpatients: An indirect effect via perceived burdensomeness. *Suicide & Life-Threatening Behavior, 49*(1), 240–252. https://doi.org/10.1111/sltb.12433

Hudson, J. (2006). Institutional trust and subjective well-being across the EU. *Kyklos, 59*(1), 43–62. https://doi.org/10.1111/j.1467-6435.2006.00319.x

Kassebaum, U. B. (2004). *Interpersonelles Vertrauen: Entwicklung eines Inventars zur Erfassung spezifischer Aspekte des Konstrukts* [Interpersonal trust: Development of an inventory to assess specific aspects of the construct] (Doctoral dissertation). Universität Hamburg.

Krastev, S., Krajden, O., Vang, Z. M., Juárez, F. P. G., Solomonova, E., Goldenberg, M. J., Weinstock, D., Smith, M. J., Dervis, E., Pilat, D., & Gold, I. (2023). Institutional trust is a distinct construct related to vaccine hesitancy and refusal. *BMC Public Health, 23*(1), Article 2481. https://doi.org/10.1186/s12889-023-17345-5

Lee, S. (2022). Subjective well-being and mental health during the pandemic outbreak: Exploring the role of institutional trust. *Research on Aging, 44*(1), 10–21. https://doi.org/10.1177/0164027520975145

Lichter, A., Löffler, M., & Siegloch, S. (2021). The long-term costs of government surveillance: Insights from Stasi spying in East Germany. *Journal of the European Economic Association, 19*(2), 741–789. https://doi.org/10.1093/jeea/jvaa009

Naef, M., & Schupp, J. (2009). *Measuring trust: Experiments and surveys in contrast and combination* (SOEPpapers on Multidisciplinary Panel Data Research, No. 167). DIW Berlin. https://www.diw.de/de/diw_01.c.776382.de/s_5479.html

Nikolova, M., Popova, O., & Otrachshenko, V. (2022). Stalin and the origins of mistrust. *Journal of Public Economics, 208*, Article 104629. https://doi.org/10.1016/j.jpubeco.2022.104629

Ozawa, S., & Sripad, P. (2013). How do you measure trust in the health system? A systematic review of the literature. *Social Science & Medicine, 91*, 10–14. https://doi.org/10.1016/j.socscimed.2013.05.005

Peschken, W., & Johnson, M. (1997). Therapist and client trust in the therapeutic relationship. *Psychotherapy Research, 7*(4), 439–447. https://doi.org/10.1080/10503309712331332133

Petrowski, K., Hessel, A., Körner, A., Weidner, K., Brähler, E., & Hinz, A. (2014). Die Einstellung zur Psychotherapie in der Allgemeinbevölkerung [Attitudes towards psychotherapy in the general population]. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 64*(2), 82–85. https://doi.org/10.1055/s-0033-1361155

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Reinemann, C., Haas, A., & Rieger, D. (2022). "I don't care, 'cause I don't trust them!" The impact of information sources, institutional trust, and right-wing populist attitudes on the perception of the COVID-19 pandemic during the first lockdown in Germany. *SCM Studies in Communication and Media, 11*(1), 132–168. https://doi.org/10.5771/2192-4007-2022-1-132

Ridenour, J. M., Hamm, J. A., Wiesepape, C. N., & Buck, B. (2024). Restoring trust for people with psychosis through psychotherapy. *Journal of Nervous and Mental Disease, 212*(4), 228–234. https://doi.org/10.1097/NMD.0000000000001757

Rothstein, B., & Stolle, D. (2002, August 29-September 2). *How political institutions create and destroy social capital: An institutional theory of generalized trust* [Paper presentation]. Annual Meeting of the American Political Science Association. Boston, MA, United States.

Rotter, J. B. (1967). A new scale for the measurement of interpersonal trust. *Journal of Personality, 35*(4), 651–665. https://doi.org/10.1111/j.1467-6494.1967.tb01454.x

Rotter, J. B. (1980). Interpersonal trust, trustworthiness, and gullibility. *American Psychologist, 35*(1), 1–7. https://doi.org/10.1037/0003-066X.35.1.1

Schmalbach, B., Zenger, M., Strauß, B., Hinz, A., Steffens-Guerra, I., Decker, O., & Brähler, E. (2016). Validation and psychometric properties of the Resilience Scale-5 (RS-5) results of a representative survey of the German general population. *Health Science Journal, 10*(5), Article 13.

Schmitt, M., Dalbert, C., Montada, L., Gschwendner, T., Maes, J., Reichle, B., Radant, M., Schneider, A., & Brähler, E. (2008). Verteilung des Glaubens an eine gerechte Welt in der Allgemeinbevölkerung: Normwerte für die Skala Allgemeiner Gerechte-Welt-Glaube [Distribution of the belief in a just world in the general population: Normative values for the General Belief in a Just World scale]. *Diagnostica, 54*(3), 150–163. https://doi.org/10.1026/0012-1924.54.3.150

Schupp, J., & Wagner, G. G. (2004). Vertrauen in Deutschland: Großes Misstrauen gegenüber Institutionen [Trust in Germany: Widespread mistrust of institutions]. *DIW Wochenbericht, 71*(21), 311–313.

Sikorski, M. T., & Albrecht, A. (2025). Trust in the context of intercultural negotiations: A systematic review. *Negotiation and Conflict Management Research, 18*(1), 1–41. https://doi.org/10.34891/a0mc-jx98

Sønderskov, K. M., & Dinesen, P. T. (2016). Trusting the state, trusting each other? The effect of institutional trust on social trust. *Political Behavior, 38*(1), 179–202. https://doi.org/10.1007/s11109-015-9322-8

Sousa-Duarte, F., Brown, P., & Mendes, A. M. (2020). Healthcare professionals' trust in patients: A review of the empirical and theoretical literatures. *Sociology Compass, 14*(10), 1–15. https://doi.org/10.1111/soc4.12828

Strupp, H. H., Wallach, M. S., & Wogan, M. (1964). Psychotherapy experience in retrospect: Questionnaire survey of former patients and their therapists. *Psychological Monographs, 78*(11), 1–45. https://doi.org/10.1037/h0093869

Tokuda, Y., Jimba, M., Yanai, H., Fujii, S., & Inoguchi, T. (2008). Interpersonal trust and quality-of-life: A cross-sectional study in Japan. *PLoS ONE, 3*(12), Article e3985. https://doi.org/10.1371/journal.pone.0003985

Ulke, C., Fleischer, T., Muehlan, H., Altweck, L., Hahm, S., Glaesmer, H., Fegert, J. M., Zenger, M., Grabe, H. J., Schmidt, S., Beutel, M. E., Schomerus, G., Brähler, E., & Speerforck, S. (2021). Socio-political context as determinant of childhood maltreatment: A population-based study among women and men in East and West Germany. *Epidemiology and Psychiatric Sciences, 30*, Article e72. https://doi.org/10.1017/S2045796021000585

van Prooijen, J.-W., Spadaro, G., & Wang, H. (2022). Suspicion of institutions: How distrust and conspiracy theories deteriorate social relationships. *Current Opinion in Psychology, 43*, 65–69. https://doi.org/10.1016/j.copsyc.2021.06.013

Vessey, J. T., & Howard, K. I. (1993). Who seeks psychotherapy? *Psychotherapy, 30*(4), 546–553. https://doi.org/10.1037/0033-3204.30.4.546

Wade, N. G., Bailey, D. C., & Shaffer, P. (2005). Helping clients heal: Does forgiveness make a difference? *Professional Psychology, Research and Practice, 36*(6), 634–641. https://doi.org/10.1037/0735-7028.36.6.634

Wagnild, G. M., & Young, H. M. (1993). Development and psychometric evaluation of the Resilience Scale. *Journal of Nursing Measurement, 1*(2), 165–178.

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).* U.S. Department of Veterans Affairs. https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp#obtain

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion, 18*(2), 129–166. https://doi.org/10.1007/BF02249397

Zinbarg, R. E., Uliaszek, A. A., & Adler, J. M. (2008). The role of personality in psychotherapy for anxiety and depression. *Journal of Personality, 76*(6), 1649–1688. https://doi.org/10.1111/j.1467-6494.2008.00534.x

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

Announcement

Correction of Bettina Hufschmidt et al. (2026). Adaptive and Maladaptive Networks Using Ecological Momentary Assessment

Clinical Psychology in Europe, 2026, Vol. 8(3), Article e23516, https://doi.org/10.32872/cpe.23516

Published (VoR): 2026-08-31

In the published version of

> Hufschmidt, B., Ebert, M., Nemani, A., Kohl, V., Pahlen, L., Müller, S., Hofmann, S. G., & Stangier, U. (2026). Adaptive and maladaptive networks using ecological momentary assessment. *Clinical Psychology in Europe*, *8*(2), Article e17891. https://doi.org/10.32872/cpe.17891

the statement on Competing Interests was incomplete. Therefore, the authors would like to replace the published Competing Interests statement with the following:

> **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 by the Alexander von Humboldt Foundation (as part of the Alexander von Humboldt Professur); the Hessische Ministerium für Wissenschaft und Kunst (as part of the LOEWE Spitzenprofessur); 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); the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 521379614 – TRR 393/SFB Transregio 393; and the Excellence Cluster EXC3066 "The Adaptive Mind" (Project No. 533717223). He also receives compensation for his work as editor from the American Psychological Association and royalties and payments for his work from various publishers.

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.