

1

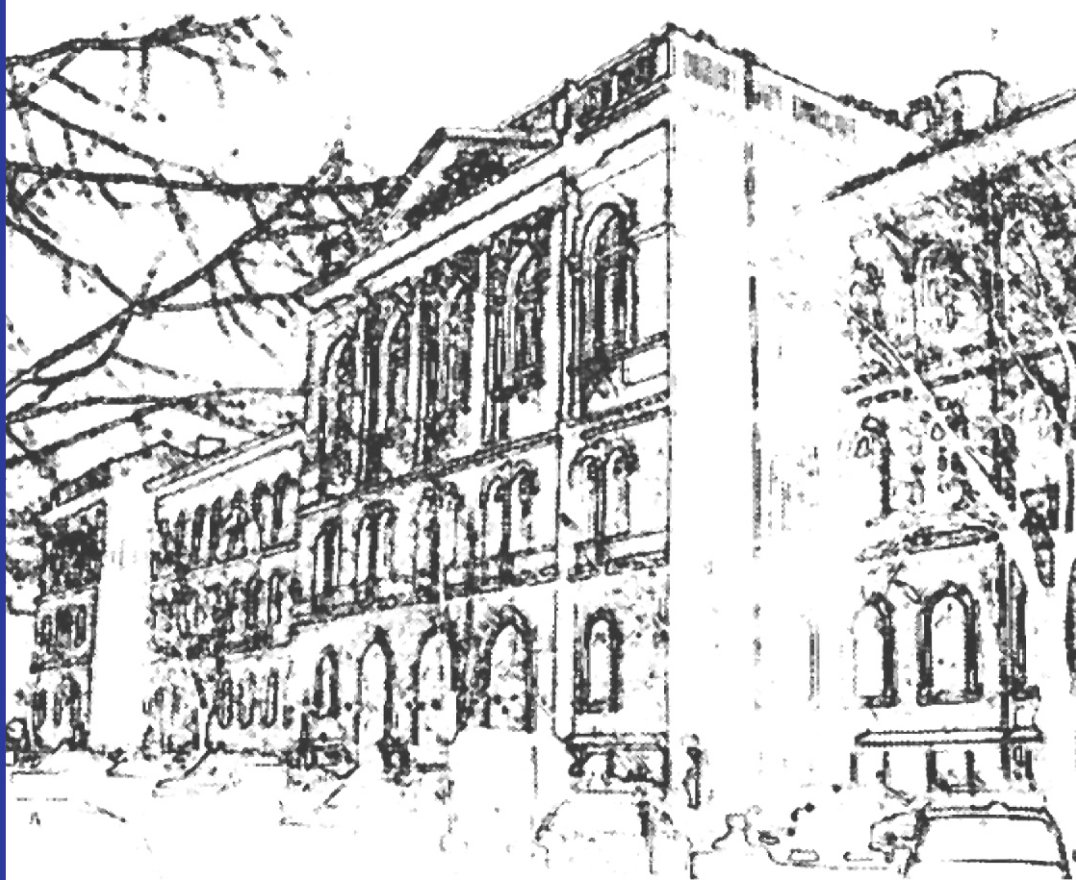
March 2026

Volume XXVI Number 1

The Official Journal of the
International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health
(IIAS of PAMH)



JOURNAL OF EVIDENCE-BASED PSYCHOTHERAPIES



Presa Universitară Clujeană

Journal of Evidence-Based Psychotherapies is the successor of the Journal of Cognitive and Behavioral Psychotherapies (2001-2013) and is published biannually (in March and September) in one volume per year by the International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health. The journal is devoted to the advancement of the clinical theory and practice of evidence-based psychotherapies (EBP) (e.g., evidence-based psychological assessments, evidence-based psychological treatments). The journal publishes original papers dealing with EBP and psychology, psychiatry, the medical and mental specialties, and allied areas of science. Submissions include (1) articles referring to clinical and experimental studies, discussions of theory, and systematic reviews for inclusion in Article Section, (2) articles referring to clinical discussions/developments for inclusion in the Clinical Forum Section, and (3) commentaries, letters to the editor, reviews and abstracts of books, tapes, and films, salient findings in EBP, and other information relevant to the journal's goal for inclusion in the Development and Resources Section. Finally, the journal seeks to publish special issues devoted to topics of particular interest, suggestions for which are welcomed by the Editor.

Indexed in: (1) **Abstracted:** Thomson ISI Web of Science/SSCI; SCOPUS; PsycINFO; and International Bibliography of Social Science; (2) **Full text:** EBSCO, ERIH+, and ProQuest

CHIEF EDITOR:

Oana A. David, Ph.D., Babeş-Bolyai University, Cluj-Napoca, Romania (2019-Present)

ASSOCIATE EDITOR:

Irving Kirsch, Ph.D., Harvard Medical School, USA/ Plymouth University, UK (2019-Present)

FOUNDING EDITOR:

Daniel David, Ph.D., Babeş-Bolyai University, Cluj-Napoca, Romania/Icahn School of Medicine at Mount Sinai, New York, USA

PAST EDITORS:

Daniel David, Ph.D., Babeş-Bolyai University, Cluj-Napoca, Romania

Aurora Szentagotai Tătar, Ph.D., Babeş-Bolyai University, Cluj-Napoca, Romania (2009-2013)

Ioana-Alina Cristea, Ph.D., Babeş-Bolyai University, Cluj-Napoca, Romania (2014-2019)

Dr. David Barlow,
Boston University, USA

Dr. Pim Cuijpers,
VU University Amsterdam, The Netherlands

Dr. Frank Dattilio,
Harvard Medical School, USA

Dr. Raymond DiGiuseppe,
St. John's University, USA

Dr. Windy Dryden,
London University, UK

Dr. Michael Eysenck,
London University, UK

Dr. Michael First,
Columbia University, USA

Dr. Joshua Fogel,
City University of New York, USA

Dr. David Haaga,
American University, USA

Dr. Steven Hayes,
University of Nevada, USA

Dr. Phillip Kendall,
Temple University, USA

Dr. Scott Lilienfeld,
Emory University, USA

Dr. Elisabeth Loftus,
University of California-Irvine, USA

Dr. Steven Jay Lynn,
SUNY at Binghamton, USA

Dr. Richard McNally,
Harvard University, USA

Dr. Donald Meichenbaum,
Waterloo University, Canada

Dr. Michael Nash,
Tennessee University, USA

Dr. Pietro Pietrini,
University of Pisa, Italy

Dr. Ann Vernon,
Albert Ellis Institute, USA

Dr. Jeffrey Young,
Columbia University, USA

ACTING EDITORIAL BOARD

Daniel David,
*Babeş-Bolyai University, Romania/
Icahn School of Medicine at Mount Sinai,
New York, USA*

Anca Dobrea, Ph.D.,
Babeş-Bolyai University, Romania

Dan Lucian Dumitraşcu, M.D., Ph.D.,
*Iuliu Hatieganu University of Medicine
and Pharmacy, Romania*

Claudio Gentili, M.D., Ph.D
University of Padua, Italy

Irina Holdevici, Ph.D.,
Titu Maiorescu University, Romania

Viorel Lupu, M.D., Ph.D.,
*Iuliu Hatieganu University of Medicine
and Pharmacy, Romania*

Simona Stefan, Ph.D.,
Babeş-Bolyai University, Romania

Aurora Szentagotai Tatar, Ph.D.,
Babeş-Bolyai University, Romania

Bogdan Voinescu, Ph.D.,
Babeş-Bolyai University, Romania

CONSULTING EDITORS

Richard Brown, Ph.D.,
Institute of Neurology, London, UK

László Tringer, Ph.D.,
Semmelweis University, Hungary

Guy Montgomery, Ph.D.,
*Icahn School of Medicine at Mount Sinai,
New York, USA*

Cristina Pojoga, M.D.,
Babeş-Bolyai University, Romania

Simona Popa, Ph.D.,
Oradea University, Romania

Agnieszka Popiel, M.D.,
Medical University of Warsaw, Poland

Jan Prasko, M.D., Ph.D.,
*3rd Medical Faculty, Charles
University, Czech Republic*

Adrian Rosan, Ph.D.,
Babeş-Bolyai University, Romania

Dan Octavian Rusu, Ph.D.,
Babeş-Bolyai University, Romania

Mihaela Stănculete, M.D.,
Adult Clinical Hospital, Romania

Mehmet Z. Sungur, M.D., Ph.D.,
*Medical School of
Marmara University, Turkey*

Jezdimir Zdravkovic, Ph.D.,
University of Nis, Serbia

Statistical Consultant

Silviu Matu, Ph.D.,
Babeş-Bolyai University, Romania

Scientific Secretary

Vasile Sîrbu, MA
Babeş-Bolyai University, Romania

Journal's Website Administrator

Răzvan Antonescu,
Cluj-Napoca, Romania

CONTENTS

ARTICLES SECTION

- An exploratory comparison of IEMT- versus EMDR-directed eye movements on changes in emotionality and distress during recall of negative memories 1**

DALENA VAN HEUGTEN – VAN DER KLOET, ANOUK BOONSTRA,
NAIMA TROUK, AND ANNE TEN BRINKE

- Feasibility and Usability of a VR-Based Training for Therapeutic Relationship Skills in Future Psychotherapists: A Pilot Study 19**

DANA-ANDREEA CRISTOLȚAN

- Maternal Depression and Children’s Socio-Emotional Development in the First Two Years of Life: A Meta-Analysis 39**

MĂDĂLINA RUXANDRA COSTIN, DIANA TĂUT,
ADRIANA SMARANDA BĂBAN, AND SEBASTIAN PINTEA

- Perceived Parental Blame and Emotion Regulation as Predictors of Adult Psychological Distress: Development and Validation of the Perceived Parental Blame Scale..... 69**

NESRULLAH OKAN

- Effect of aerobic exercise on relapse tendency in individuals with methamphetamine use disorder in a virtual social environment: Mediated by cue-induced craving and interoceptive sensibility 93**

XIAOXIA LIU AND ZEFAN FAN

- Do Values Change After Major Life Events? A Meta-analytic Review 117**

DORIS POP, OANA A. DAVID

- Effectiveness of Group-Based Cognitive Behavioral Therapy in Reducing Depression and Anxiety Among Adolescents in Residential Institutions: A Randomized Controlled Trial 137**

DING WAI ENG AND FIRDAUS MUKHTAR

- Improving students' self-esteem with an online group intervention based on the Fennell Model: an open-label, uncontrolled feasibility trial 163**

LUANA M. ALEXA, ANDREI RUSU

The International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health (IIAS of PAMH) is a research Institute of the Babes-Bolyai University (BBU), located in Cluj-Napoca, Romania. Babes-Bolyai University is one of the oldest, most prestigious, and influential academic settings in Romania and in the Eastern-Central Europe (<http://www.ubbcluj.ro>). It has been almost four centuries since the beginnings of this university of Cluj-Napoca. The first project to set up a higher education institution in Cluj-Napoca was recorded in 1565, and in 1581 it was set up an academic college, organized and run by Jesuits. Since then, Babes-Bolyai University has become a modern institution which acquired a remarkable fame by its long-time achievements at national and international level. The Institute will invite as affiliates/partners prestigious organizations involved in the development of psychotherapy and applied mental health interventions. IIAS of PAMH aims to promote advanced study and scholarship in applied mental health by research.

The Faculty of IIAS of PAMH will have strategic and consultative responsibilities. It is composed of prestigious professionals in the area of Psychotherapy and Applied Mental Health Interventions:

Dr. David Barlow Boston University, USA
Dr. Pim Cuijpers, VU University Amsterdam,
The Netherlands
Dr. Frank Dattilio, Harvard Medical School, USA
Dr. Raymond DiGiuseppe, St. John's University, USA
Dr. Windy Dryden, London University, UK
Dr. Michael Eysenck, London University, UK
Dr. Michael First, Columbia University, USA
Dr. Joshua Fogel, City University of New York, USA
Dr. David Haaga, American University, USA
Dr. Steven Hayes, University of Nevada, USA

Dr. Phillip Kendal, Temple University, USA
Dr. Scott Lilienfeld, Emory University, USA
Dr. Elisabeth Loftus, University of California-Irvine, USA
Dr. Steven Jay Lynn, SUNY at Binghamton, USA
Dr. Richard McNally, Harvard University, USA
Dr. Donald Meichenbaum, Waterloo University, Canada
Dr. Michael Nash, Tennessee University, USA
Pietro Pietrini, M.D., Ph.D., University of Pisa, Italy
Dr. Ann Vernon, Albert Ellis Institute, USA
Dr. Jeffrey Young, Columbia University, USA

Specific information and more detailed information about IIAS of PAMH may be obtained by visiting our website at: <http://www.psychotherapy.ro> or by emailing us to internationalinstitute@psychology.ro

ISSN: 2360 – 0853

ISSN-L: 2360 – 0853

© 2026, The Editors.

Universitatea Babeș-Bolyai
Presă Universitară Clujeană
Director: Codruța Săcelean
Str. B.P. Hasdeu nr. 51
400371 Cluj-Napoca, România
Tel.: (+40) 744 687 884
E-mail: editura@ubbcluj.ro
editura.ubbcluj.ro | biblioteca.ubbcluj.ro

AN EXPLORATORY COMPARISON OF IEMT- VERSUS EMDR-DIRECTED EYE MOVEMENTS ON CHANGES IN EMOTIONALITY AND DISTRESS DURING RECALL OF NEGATIVE MEMORIES

*Dalena van Heugten – van der Kloet^{1*}, Anouk Boonstra²,
Naima Trouk³, and Anne ten Brinke⁴*

¹ Maastricht University, Faculty of Psychology & Neuroscience, Department of Clinical Psychological Science, Maastricht, the Netherlands

² Maastricht University, Faculty of Health, Medicine & Life Sciences, Department of Psychiatry & Neuropsychology, Mental Health and Neuroscience Research Institute, Maastricht, the Netherlands

³ JeBeNTR, Eindhoven, the Netherlands

⁴ IEMT-training, Ruurlo, the Netherlands

Abstract

Eye movement therapies have shown effectiveness in reducing distress pertaining to past traumatic events. In the clinical field, a variety of eye movement therapies exists, with the directions of eye movements and respective guidelines differentiating one from the other. Anecdotal work field evidence suggests Integral Eye Movement Therapy (IEMT) is as effective as Eye Movement Desensitization and Reprocessing (EMDR) and has less side effects; however, research is scarce at present.

Participants were recruited from the general population, and all received three conditions (IEMT, EMDR, control) in randomized order ($N=33$). Participants were asked to recall a pre-chosen negative memory. Questionnaires were answered pre-, post-intervention, and at one-week follow-up. These self-reports consisted of visual analogue scales measuring emotionality and distress, forming the main outcome of Subjective Units of Distress (SUD).

Blinded to condition, 60.6% preferred the IEMT condition to the EMDR- or the control condition. SUD-scores were significantly lower after both IEMT and EMDR conditions than after the control condition at post-intervention ($p=.001$; $p<.001$, respectively) and at follow-up ($p=.014$; $p=.002$,

* Correspondence concerning this article should be addressed to Dalena van Heugten – van der Kloet, Maastricht University, Faculty of Psychology & Neuroscience, Department of Clinical Psychological Science, Maastricht, the Netherlands.

ORCID: 0000-0002-3404-1279

Email: dalena.vanheugten@maastrichtuniversity.nl

respectively). IEMT and EMDR conditions did not significantly differ at any time point.

This study was the first to compare IEMT and EMDR directed eye movements and found that each significantly lowered the experienced emotionality and distress related to a negative memory, as tested in randomized order in the general population.

Keywords: IEMT, EMDR, eye movement therapy, memory, treatment

Most individuals will experience potentially traumatic events. The lifetime prevalence of potential trauma has been estimated at 80.7% (De Vries & Olff, 2009), and 5.6% for those with Post-Traumatic Stress Disorder (PTSD) (Koenen et al., 2017). As such, trauma has a substantial economic impact: 600-900 million euros of annual societal costs were estimated for the loss of wellbeing due to symptoms of PTSD in The Netherlands (Dijk & Kiernan, 2023). The costs of potentially traumatic events could assumedly be estimated even higher as their prevalence is much higher than that of PTSD. Adverse childhood experiences (ACEs) alone were experienced once by 23.5% and multiple times by 18.7% of individuals, which has been estimated to cost US\$581 billion in Europe (Bellis et al., 2019).

Given the impact of traumatic experiences on the individual, their environment and society in general, effective interventions that counteract the negative consequences of traumatic experiences are essential. Interventions are usually aimed at the reprocessing of emotional memories, based on the rationale that psychopathology occurs when a traumatic experience has not been processed healthily (Emotional Processing Theory; Foa & Kozak, 1986). One treatment avenue for emotional memory reprocessing has been found in the directing of eye movements. Directed eye movements are hypothesized to suppress the activity of the amygdala and strengthen extinction learning whilst reducing the vividness of a memory (de Voogd et al., 2018). The most well-known therapeutic approach involving eye movements is Eye Movement Desensitization and Reprocessing (EMDR; Shapiro, 1989), but other eye movement therapies exist such as Integral Eye Movement Therapy (IEMT), Eye Movement Integration (EMI, Beaulieu, 2003; van der Spuy & DuPlessis van Breda, 2019), and Visual Schema Displacement Therapy (Matthijssen et al., 2019). These interventions differ in the direction of eye movements on different axes and lengths (horizontally, diagonally, and circular) and have various guidelines. EMDR is a recognized treatment form in the mental health care standards of the Netherlands (GGZ Standaarden, 2020), with over 300 studies having examined its clinical application for PTSD (Landin-Romero et al., 2018). Increased evidence for the influence of life events, distress and trauma on the

development of psychopathology has also led to an effective use of EMDR in the treatment of anxiety disorders, mood disorders, psychotic disorders, pain management and dependency (Carletto et al., 2021; Landin-Romero et al., 2018; Lytle et al., 2002). However, more recently, it was stated that eye movements in any form might be effective for emotional reprocessing, provided that the working memory is sufficiently taxed (de Voogd et al., 2018). Indeed, a reduction of emotionality and vividness has been shown when reactivation of memory was paired with other working-memory tasks (e.g., Holmes et al., 2009; James et al., 2015). Specifically, retrieving a memory whilst making eye movements is hypothesized to compete for limited working memory resources, thereby attenuating the vividness and emotionality of a memory (Wadji et al., 2022). The present study is explicitly grounded in the working memory account of eye-movement-based interventions, which posits that the effectiveness of both EMDR and other eye-movement therapies can be explained by competition for limited working memory resources during simultaneous memory recall and eye movements.

IEMT is a therapeutic model developed in 2006 by the British psychotherapist Andrew T. Austin (Austin, 2009, 2015), aimed at dealing with and reducing of intense negative emotional states. This model's origins can be traced back to Andreas's work regarding EMI Therapy (Andreas, 1993) and Shapiro's EMDR therapy (Shapiro, 1989; 2018).

In short, this IEMT model involves the recalling and maintaining of a negative memory regarding an event experienced by the client, which may present itself in the form of an image, sound, smell, feeling, taste, and/or bodily sensation. The client is invited to focus their attention on the memory, while the therapist instructs the client to move their eyes in different but specific directions by pointing them out with a pen, a telescope stick with a ball, or with their finger. The outcome of the therapeutic process is usually the client's loss of the recalled memory's negative emotional impact. The client experiences more distance from the memory and the memory's properties (e.g., its colors) become vaguer. Another important aspect of the IEMT model resides in the fact that the client is not required to disclose the memory's contents to the therapist.

The main hypothesis underlying the therapeutic approach of IEMT, differentiating it from EMDR, is based upon the assumption that eye movements in general can predict the client's recalled experience (Sharot et al., 2008) and that the specific set of eye movements used in the IEMT model is connected to brain areas that regulate memories and emotions, such as the hippocampus and the amygdala (Austin, 2009; Austin, 2015; Ten Brinke, 2023). IEMT therapists are trained to identify 'blinks', defined as tiny involuntary twitches of the eye pupil. The therapist is specifically attentive to these unconscious eye movements during the intervention. The rationale behind this is that, while thinking about the memory, the client may unconsciously try to sidestep or lapse away from particular areas in the visual field.

The blink then signals a changed mental representation within the client and the therapist will alter the course of the directed eye movements accordingly, until the client can follow the movements without disruptions (Austin, 2009; Austin, 2015; Ten Brinke, 2023). During the full IEMT protocol (Austin, 2009; Austin, 2015; Ten Brinke, 2023), the therapist will guide the client through emotionally associated memory traces compared to the focus on a single memory in the EMDR protocol, an attribute falling outside the scope of the current study. Two other main features distinguishing IEMT from EMDR are the length and direction of the eye movements. During EMDR, therapists tend to use narrow and horizontal movements (approximately 60 cm in width), while during IEMT the movements are broader (approximately 150 cm in width) and move horizontally, diagonally, and horizontal eight-shape, see Fig. 1 (Austin, 2009; Austin, 2015; Ten Brinke, 2023).

IEMT is being practiced prolifically in the Netherlands and the United Kingdom, among other countries. Anecdotal evidence from the clinical field suggests that IEMT is experienced by clients as more pleasant and more effective, with less negative side effects than its evidence-based predecessor EMDR. This claim about IEMT's effectiveness contains elements of pseudoscience, particularly its usefulness for a wide variety of mental complaints and the unclear theory about its working mechanism (Cuijpers & Cristea, 2016). More importantly, there is the lack of a solid scientific foundation for this form of treatment (but see Moore & Manea, 2018). Although IEMT currently lacks a strong empirical foundation, it can be conceptually understood within existing theoretical models of eye-movement-based interventions, particularly the working memory account. From this perspective, IEMT does not represent an unexpected or anomalous approach, but rather a variant of dual-attention procedures aimed at taxing working memory during memory retrieval.

Therefore, the authors decided to conduct a critical test to compare the differences in directed eye movement techniques. This study focused on a comparison between eye movement techniques based on IEMT, EMDR and a control condition on diminishing the emotional valence of negative memories, within an experimental laboratory setting. Because previous research demonstrates that a variety of working-memory taxing tasks (including but not limited to eye movements) can reduce emotionality and vividness of memories (e.g., Van de Voogd et al., 2028; Wadji et al., 2022), we hypothesized that IEMT and EMDR would produce comparable effects, as both rely on dual-attention eye-movement tasks during memory retrieval. Furthermore, we hypothesized that both IEMT and EMDR conditions would outperform the control condition.

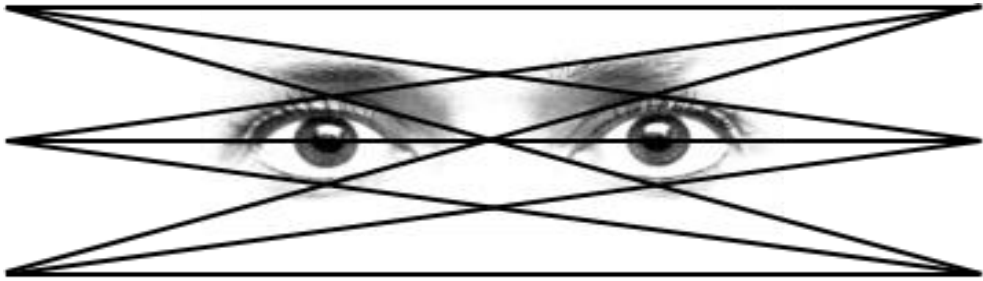


Figure 1. Directions of eye movements during IEMT condition

Methods

Participants

Participants were recruited from the general population through flyers, posters and online promotion. Upon expressed interest, participants received an e-mail with information regarding the exclusion criteria and a secure webpage containing the screening and pre-measurement questionnaire. Exclusion criteria were: (1) having an eye disease; (2) having a neurological disease; (3) having history of head trauma; (4) having to testify in legal procedures; and (5) receiving current treatment for a diagnosis according to the DSM-5. After screening, inclusion held for 33 participants aged 20-60 years ($M=41.5$, $SD=12.1$), of which 25 (75.8%) identified as female. Participants signed an informed consent containing information about study purposes, anonymity, and the potential discomfort in recalling negative memories. Participants were informed about their right to stop the experiment at any moment without repercussions.

Procedure

This prospective study followed a repeated measures design with questionnaires administered at baseline and pre-measurement, which were later averaged to one pre-measurement score, at post-measurement, and at one-week follow-up. After an information e-mail in which the participants chose a two-hour timeslot, the researchers sent an e-mail containing the baseline questionnaire five days before the chosen timeslot. The baseline measurement was administered through a secure online survey system termed Qualtrics. This measure was followed by the request to choose three aversive memories for recall in the experiment (i.e., negative memory). Participants entered a self-chosen keyword for each negative memory in order to remember the choice later and for the researchers to randomly allocate each memory to a condition (IEMT, EMDR, and the control condition).

Participants were blinded to which condition was offered at which time, receiving only a sequence of rooms to enter.

At the day of the experiment, the researchers again briefed the participants, who then signed the informed consent and answered the pre-measurement questionnaire. Following this, all participants experienced all three conditions, each lasting 20 minutes, with a 10-15 minutes resting time between conditions. The therapists set a timer to make sure each condition ended after 20 minutes. Participants had received physical documents that contained the pre- and post-measures directly related to the memories, for each condition, altogether taking maximally half an hour to fill out. At the end of all three conditions, the researchers interviewed the participants about their preferred condition and therapist, still blinded, and enquired about any further remarks.

The IEMT and EMDR conditions were recorded for fidelity. To ensure the procedures were conducted properly, fidelity checks were carried out by supervising video recordings. Recordings were securely saved on a server of Maastricht University. A week after the laboratory session, participants received the follow-up questionnaire via an e-mail in which their keywords were listed again to correctly fill out the questions pertaining to each memory. After completion, participants received the debriefing letter and a monetary reward for participating, a €10,- gift voucher.

Interventions

Detailed study protocols for both the IEMT and EMDR conditions are listed in the appendix. In the IEMT condition, eye movement techniques were conducted according to the existing IEMT-protocols (Austin, 2009; ten Brinke, 2023). The IEMT condition was carried out by three licensed IEMT therapists (EG, CB & JR). After a short introduction with psychoeducation about eye movements, participants were asked to recall the distressing memory. After recalling the memory and rating characteristics on an 11-point Likert scale (Subjective Units of Distress (SUD-scores), see *Instruments*), a set of questions was asked pertaining to the representation and experience of the memory. Next, the IEMT-therapist instructed the participants to sit up straight and follow the IEMT-stick with their eyes without moving their head. Whilst doing this, participants had to recall the memory, continuing to do so whenever the recall became increasingly difficult. After each round of eye movements, lasting approximately 30 seconds, the emotional valence of the memory was scored between 0 and 10 points. Rounds of eye movements were performed until the memory was rated as neutral (score = 0) by the participant or until the end of the condition time.

Mostly ($n=23$), the condition was supplemented at the ending with a positive component for participants: horizontal eight-shaped eye movements, with the instruction to think about a highly positive memory. At the end of the condition, the

participant was asked to re-rate the memory characteristics on the 11-point Likert scale, which were a set of questions pertaining to the representation and experience of the memory.

The EMDR condition was carried out by two licensed EMDR therapists (AV & HD). The EMDR-therapists conducted the directed eye movements according to an abbreviated version of the Dutch standard EMDR-protocol (ten Broeke et al., 2023; De Jongh et al., 2012), see the appendix for details. They asked the participant to recall the memory connected to their selected keyword and to envision the most disturbing part of the memory as the target image. Participants were then asked to rate negative cognition and SUD-scores. Desensitization was performed until condition time ran out or until the SUD-score was rated zero. Depending on the participant process, this could result in either one, or multiple rounds of eye movements. In-between eye movements, the SUD-score was rated again. Some of the sessions ended with focusing on a positive cognition*. At the end of the condition, the participant was asked to re-rate the memory characteristics using the SUD-score.

In the control condition, the participant was asked to recall the memory that was randomly allocated to that room, and to fill out the pre-intervention VAS regarding the memory. Next, the participant was asked to do nothing whilst the researcher sat with them in silence. Participants were requested to think about any topic that came to mind, but to not practice meditation at this time. At eight- and fifteen- minutes' time, participants were asked to recall the memory and fill out the VAS.

Instruments

The *Depression Anxiety Stress Scale 21-R* (DASS-21-R; Henry & Crawford, 2005) is a short version of the original DASS containing 42 items (Lovibond & Lovibond, 1995). It contains 21 items in three subscales, which assess symptoms of depression (items 3, 5, 10, 13, 16, 17, 21; $\alpha = .90$), anxiety (items 2, 4, 7, 9, 15, 19, 20; $\alpha = .68$), and stress (items 1, 6, 8, 11, 12, 14, 18; $\alpha = .81$). The degree to which respondents endorsed the symptoms over the course of the last week is rated on a scale that ranges from 0 (did not apply to me at all) to 3 (applied to me very much or most of the time). Higher scores reflect higher levels of symptom endorsement (Henry & Crawford, 2005). In the present study, the overall reliability of the DASS-21 is rated as excellent ($\alpha = .90$ pre-intervention, $\alpha = .88$ post-intervention).

* Out of 66 video recordings, 14 were (partially) lost due to technical errors. Among the available recordings, 13 EMDR conditions were observed to end with a positive cognition. Based on these recordings, it may be assumed that the majority of EMDR conditions were also concluded with a positive cognition, as therapists were consistent in their methods. However, this cannot be verified for the missing recordings.

Subjective intensity of disturbance. Participants recalled perceived intensity of disturbance or distress of an image or an emotional memory. The score was indicated on an 11-point Likert scale, the Subjective Units of Disturbance (SUD) scale, ranging from 0 ('none at all') to 10 ('maximum distress'). The SUD scale was first introduced by Wolpe (1969) and it is incorporated in the IEMT protocol (ten Brinke, 2023; Austin, 2009), as well as in the EMDR protocol (Shapiro, 2018, Ten Broeke et al., 2023). In the present study, participants were requested to indicate the SUD-score verbally to the therapist during the session and in written form at the start and after each condition, as well as at one-week follow-up.

Data analyses

Statistical analyses were performed using SPSS 27.0. After running descriptive analyses for age, gender, highest finished education and preferred condition, repeated measures analysis of variance (ANOVA) was conducted with SUD-scores as outcome measure. Effects of time (pre-intervention, post-intervention, and follow-up), condition (IEMT, EMDR and control) and their interaction were inspected. In case of significant effects, follow-up multiple comparisons were conducted using separate Bonferroni-corrected ANOVAs split per time and condition. Additionally, exploratory t-tests were run to test differences from pre- to post-intervention in the subscales of depression, anxiety and stress, as measured by the DASS-21-R. All analyses were conducted at $\alpha=.05$. Effect sizes for within-subject changes were calculated using Cohen's d_z , while partial eta squared values were calculated for the ANOVA effects.

Results

Participant characteristics

Of the 33 participants, 25 (75.8%) were female. The mean age was 41.45 years ($SD=2.11$) and ranged from 20 to 60 years. For the majority of the participants, the highest finished education was higher professional education ($n=11$, 33.3%), a university bachelor's education ($n=9$, 27.3%) or university master's education ($n=7$, 21.2%). For others, the highest finished education was a vocational education ($n=1$, 3%), secondary higher-level education ($n=1$, 3%), secondary vocational education ($n=3$, 9.1%) or a PhD ($n=1$, 3%). Based on established DASS-21 severity cut-offs (Lovibond & Lovibond, 1995), the participants scored within the lower end 'normal' range for depression ($M=5.64$, $SD=6.51$), anxiety ($M=4.73$, $SD=4.87$), and stress ($M=10.85$, $SD=6.10$) at baseline, confirming that this was a predominantly non-clinical sample.

Regarding preference for condition (blinded), 20 (60.6%) participants reported to prefer IEMT and nine (27.3%) reported to prefer EMDR. Four participants (12.1%) reported no preference. Reasons for preferring IEMT over EMDR entailed the following: eye movements were more pleasant ($n=5$); IEMT was calmer, softer, and took an emotional perspective ($n=10$); IEMT was more intense but felt more effective ($n=4$); no need to talk during IEMT ($n=4$); having felt changes on a physical level ($n=3$); more in-depth questions during IEMT ($n=1$); feeling of being seen and heard more during IEMT ($n=1$); IEMT brought greater insight, meaning, depth, synthesis, due to its all-encompassing character ($n=10$); and tension in the head, headache and tired eyes during EMDR ($n=4$). Reasons for preferring EMDR over IEMT were: talking during the session helped to find the memory ($n=2$); clear communication and guidance ($n=2$); having to work harder during EMDR ($n=4$); EMDR felt more forceful and therefore effective ($n=1$); EMDR had a stronger effect ($n=3$); and EMDR took a cognitive perspective ($n=1$).

Main outcomes

Repeated measures ANOVA yielded significant main effects of condition ($F[1.619, 50.196] = 24.410, p < .001, \eta^2p = .441$) and time ($F[1.708, 52.951] = 78.992, p < .001, \eta^2p = .718$). A significant interaction effect was found for condition and time ($F[2.551, 79.077] = 9.868, p < .001, \eta^2p = .241$). The profile plot suggested a steeper decrease in SUD-scores for IEMT and EMDR conditions than for the control condition, see Fig 2. This decline held from pre- to post-intervention measurement and was maintained at follow-up. The Bonferroni-corrected 95% CI showed that participants had reliably lower SUD-scores post-IEMT and post-EMDR than after the control condition, as shown in Table 1.

Table 1. SUD-score means, S.E.s and Bonferroni-adjusted 95% CI of ANOVA interaction effect

Condition	Time								
	Pre-intervention			Post-intervention			Follow-up		
	Mean	S.E.	95% CI	Mean	S.E.	95% CI	Mean	S.E.	95% CI
EMDR	61.36	3.02	55.20; 67.52	17.08	3.17	10.62; 23.54	20.80	3.20	14.28; 27.32
IEMT	66.85	2.72	61.31; 72.39	24.16	3.79	16.44; 31.88	23.58	3.69	16.04; 31.11
CONTROL	64.88	2.69	59.39; 70.36	45.59	4.07	37.29; 53.90	35.11	4.53	25.87; 44.35

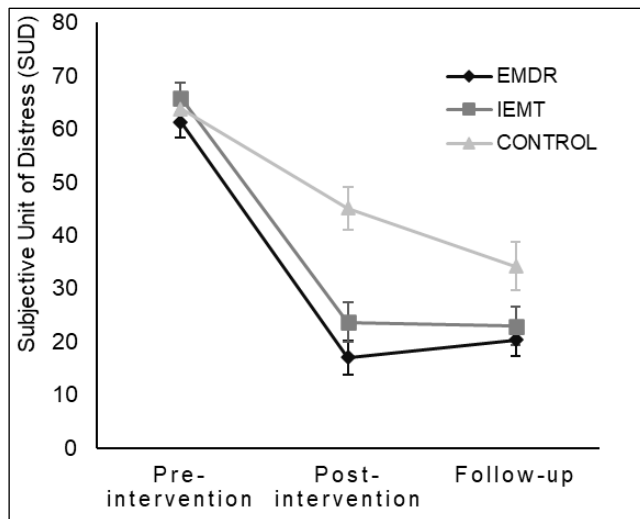


Figure 2. Profile plot SUDs with S.E. error bars, across time and condition

Both univariate and multivariate methods showed significant interactions between condition and time, necessitating follow-up analyses. No pre-intervention differences were found in SUD-scores between the conditions ($p=.113 - p=1.000$). SUD-scores were significantly lower in the IEMT condition than in the control condition post-intervention ($p=.001$) and at follow-up ($p=.014$). Likewise, SUD-scores were significantly lower in the EMDR condition than in the control condition at post-intervention ($p<.001$) and at follow-up ($p=.002$). IEMT and EMDR did not significantly differ at post-intervention or at follow-up, as detailed in Table 2.

Pre- to post-EMDR, the SUD-score significantly decreased by 44.28 points (S.E. = 4.15, [95% CI -54.78; -33.79], $p<.001$, Cohen's $d_z = 1.86$). From the post-EMDR condition to EMDR follow-up, the SUD-score increased non-significantly by 3.72 points (S.E. = 3.00, [95% CI -3.88, 11.32], $p=.675$, Cohen's $d_z = 0.22$). For the IEMT condition, the SUD-score lowered from pre- to post-measurement by 42.70 points (S.E. = 4.08, [95% CI -53.03; -32.36], $p<.001$, Cohen's $d_z = 1.82$), which was maintained at follow-up with a non-significant further decline of .58 points (S.E. = 3.30, [95% CI -8.92; 7.76], $p=1.000$, Cohen's $d_z = 0.03$). Pre- to post-control, the SUD-score significantly decreased by 19.28 points (S.E. = 4.65, [95% CI 7.52; 31.04], $p<.001$, Cohen's $d_z = 0.72$). From the post-control condition to control follow-up, the SUD-score decreased non-significantly by 10.48 points (S.E. = 5.25, [95% CI -2.80, 23.77], $p=.164$, Cohen's $d_z = 0.35$).

Overall, the reduction in SUD scores from pre- to post-intervention in all conditions was large to very large based on the effect sizes, and small for the non-significant reductions from post-intervention to follow-up.

Table 2. Multiple comparisons of the conditions at each time point

Time	Condition comparison	Mean difference	S.E.	<i>p</i> ^a	95% CI ^a
Pre-intervention	EMDR-IEMT	-5.49	2.53	.113	(-11.90, .91)
	Control-IEMT	-1.98	2.21	1.000	(-7.57, 3.62)
	Control-EMDR	3.52	2.14	.334	(-1.91, 8.94)
Post- intervention	EMDR-IEMT	-7.08	3.23	.109	(-15.26, 1.10)
	Control-IEMT	21.44	5.42	.001	(7.71, 35.16)
	Control-EMDR	28.52	3.89	.000	(18.67, 38.36)
Follow-up	EMDR-IEMT	-2.59	2.55	.954	(-9.04, 3.86)
	Control-IEMT	11.26	3.71	.014	(1.88, 20.63)
	Control-EMDR	13.85	3.67	.002	(4.57, 23.13)

^a Bonferroni-adjusted

Sensitivity analyses

Using multivariate methods, no significant differences in SUD-scores were found between therapists within the IEMT condition ($n=19$ for therapist A, $n=8$ for therapist B and $n=6$ for therapist C) at post-measurement (A-B = -6.62, S.E. = 8.93, [95% CI -29.27; 16.04], $p=1.000$; B-C = -7.08, S.E. = 11.45, [95% CI -36.11; 21.95], $p=1.000$; A-C = -13.70, S.E. = 9.93, [95% CI -38.87; 11.47], $p=.533$). The same held at follow-up (A-B = -11.67, S.E. = 8.38, [95% CI -32.91; 9.57], $p=.521$; B-C = -8.00, S.E. = 10.74, [95% CI -35.22; 19.22], $p=1.000$; A-C = -19.67, S.E. = 9.31, [95% CI -43.28; 3.93], $p=.129$). Likewise, no significant differences were found between therapists within the EMDR condition ($n=23$ for therapist D vs. $n=9$ for therapist E) at post-measurement (D-E = -9.94, S.E. = 6.93, [95% CI -24.10; 4.22], $p=.162$) or follow-up (D-E = -5.00, S.E. = 7.17, [95% CI -19.65; 9.65], $p=.491$).

Exploratory analyses DASS-21-R

After experiencing all three conditions, participants scored on average significantly lower on the DASS-21-R subscale *stress* by 2.06 points, as measured at follow-up ($M=8.78$, $SD=6.44$), compared to before the experiment ($M=10.85$, $SD=6.10$), $t(32) = 2.07$, $p=.047$. Furthermore, participants scored 2.97 points lower on the DASS-21-R subscale *depression* at follow-up ($M=2.67$, $SD=3.56$) than at pre-intervention ($M=5.64$, $SD=6.51$), $t(32)=3.91$, $p<.001$). The decrease in DASS-21-R subscale *anxiety* did not significantly deviate from zero, $t(32)=1.64$, $p=.111$, from pre-intervention ($M=4.72$, $SD=4.02$) to post-intervention ($M=3.33$, $SD=4.02$). Both the pre- and post-intervention DASS-21 scores fall within the lower end of the established cut-off ranges and are classified as ‘normal’. Specifically, the normal

range is defined as 0–14 points for the *stress* subscale, 0–9 points for the *depression* subscale, and 0–7 points for the *anxiety* subscale.

Discussion

The present study demonstrated significant reductions in experienced emotionality and distress related to a negative memory through directed eye movements either with an IEMT or an EMDR approach, with no significant differences among these two eye movement techniques. Participants scored significantly better at both post-intervention and follow-up after receiving either eye movements technique than after the control condition. As displayed by the DASS-21-R results, participants experienced less stress and depressive symptoms at follow-up.

The first hypothesis that both IEMT and EMDR conditions would outperform the control condition could thus be confirmed by the present findings. The second hypothesis, that IEMT directed eye movements would perform equally compared with EMDR directed eye movements, could be confirmed as well. This is in line with literature suggesting that the specific shape or form of the eye movements, horizontally or diagonally, may not be as crucial as previously suspected (Van de Voogd et al., 2018). Given the relatively small sample size and experimental nature of the design, findings should be interpreted as exploratory and preliminary.

Interestingly, nearly two thirds of participants preferred the IEMT condition over the EMDR condition despite reductions in SUD-scores being comparable. Those who preferred the IEMT condition most often mentioned that IEMT took a calmer, softer, and more emotional perspective and that it brought greater insight, depth and synthesis, due to its all-encompassing character. Interestingly, obtaining meaningful insights while not having to speak may warrant some reconsiderations regarding the necessity of talking therapies. Experiencing more tension in the head, headaches, and tired eyes during the EMDR condition were also mentioned as reasons for preferring the IEMT condition, reminding us of anecdotal evidence from the clinical field. Participants who preferred the EMDR condition appreciated the cognitive perspective and clear communication, felt that talking during the session helped them to recall the memory, and believed that having to work harder felt more forceful and therefore effective. Potentially, individual differences exist that might explain why certain participants prefer one condition over the other.

The majority of sessions in both experimental conditions ended on a positive note. In the EMDR condition, one focused on a positive cognition, while in the IEMT condition one focused on a positive memory or emotion while following horizontal eight-shaped directed eye movements. In the field of imagery rescripting (ImRs), scientific studies have been conducted to investigate to what degree clinicians use a

positive ending to a session and to what effect, as this varies across protocols and clinicians (Hackmann, 2011). Some explicitly instruct patients to conclude an ImRs session with a positive ending (Van der Wijngaart, 2021), while others do not (e.g., Arntz & Weertman, 1999). While positive emotions can be heightened through a positive component within ImRs, it is as of yet unclear whether this positive component significantly enhances the effectiveness of treatment (Geschwind et al., 2024). The Broaden and Build theory (Fredrickson, 2001) and the undoing hypothesis (Fredrickson et al., 2000) posit that positive emotions may help individuals undo the adverse effects of trauma and build durable biopsychosocial resources. Likewise, for IEMT and EMDR positive endings may enhance treatment effectiveness, as should be investigated in future research.

Interestingly, our findings also showed a significant reduction in experienced emotionality and distress related to a negative memory within the control condition. This reduction of the memory's negative emotional valence might be explained by the requests to think of the negative memory in the presence of a supporting professional, i.e. one of the researchers, and to repeatedly score VAS scales, actions which may lead participants to expect change. The control condition may have also unintentionally functioned as a wakeful rest condition, which has been shown to reduce negative thoughts (Hørlyck et al., 2019). These results reminded us of the potential effects of social buffering (Bratéc et al., 2020), expectancy effects and other common factors (Constantino et al., 2018), and the Hawthorne Effect (McCarney et al., 2007), stipulating the necessity and importance of incorporating an active control condition in experimental designs.

Whilst the present study showed equal outcomes for the IEMT and EMDR conditions, it is yet unknown how IEMT might achieve such results. However, IEMT's effectiveness is hypothesized on the assumption that eye movements, and more specifically purposeful movement along the line of the 'blink', may facilitate change. The blinks are hypothesized to be connected to areas of the brain that are in charge of memories and emotions, such as the hippocampus and the amygdala (Austin, 2009). In the current study, this hypothesis could not be tested. Future studies should entail EEG (electroencephalogram) or MRI (magnetic resonance imaging) techniques to establish changes in specific brain areas before, during, and after IEMT. If underlying processes are established and results are replicated, IEMT may prove a promising avenue in treatment by addressing traumatic experiences.

Strengths and limitations

Notable strengths of this study included the randomization of conditions, the use of a control group, the experimentally controlled setting and the novel insights that were obtained from the data. Participants had to prepare three negative

memories, which were randomized over conditions to avoid participants choosing which memory to work on themselves. Participants were furthermore blind to the conditions.

Several limitations must be noted. First, the sample size of this first exploration was relatively small, and must be expanded upon in future evaluations. Second, while EMDR and IEMT should be beneficial to all persons who experienced adversity, generalizability to clinical groups is presently unclear due to the use of a non-clinical sample. Another important limitation was the restricted timeframe we allowed per session, a maximum of 20 minutes per condition. This meant that therapists in both experimental conditions could not conduct the full protocol of either IEMT or EMDR. The full potential of both therapy conditions could therefore not be established and thoroughly compared through this study, but rather the focus was on the specific eye movement technique. Therefore, conclusions on the forms of therapy should be made with caution as both interventions involve additional and varying components within their protocols. Future studies should therefore employ longitudinal designs with longer follow-ups and more extensive time frames, using the full protocol, for a stronger comparison and to establish the true effectiveness of each therapy condition.

Furthermore, the design of the study meant that participants went from one condition to the other with a short break in between. As such, one may speculate that treatment effects could have therefore crossed over from one session to the next. However, the applied randomization of conditions should control for this potential risk.

A final limitation is the absence of an active working-memory control task (e.g., Tetris; Holmes et al., 2009). Future studies should include such conditions to more directly test whether observed effects are attributable specifically to eye movements or to working-memory taxation more broadly.

Conclusion

This study is, to the best of our knowledge, the first to compare IEMT directed eye movements with EMDR directed eye movements. Results showed that participants experienced a similar decrease in emotionality and distress related to a negative memory in both conditions. This leads to the cautious conclusion that directed eye movements in general, regardless of their specific directions or axes, may be responsible for the decrease in distress via taxation of the limited working memory capacity. The majority of participants preferred IEMT as a treatment condition, however, it remains unclear how IEMT yields its positive effects. Replication of these results is warranted, both in non-clinical and clinical samples.

Authors' note

Statements and Declarations: This study is funded by a SWOL grant, Universiteitsfonds Limburg, the Netherlands, awarded to the first author. This study has been ethically approved by the Ethical Review Commission of Psychology and Neuroscience at Maastricht University (Date: 16-02-2023; ERCPN-263_18_02_2023). Informed consent was obtained from all individual participants included in the study.

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. DH, AtB and NT designed the study and collected data; AB and DH analyzed the data. All authors contributed to writing the article.

Acknowledgments: We would like to thank Ruud Koolen for this help with data collection and organization during the study. We would like to thank Anouk Velthuisen, Henriëtte Decker, Ellen Nicolaides-de Groot, Carla Bloemen and Jeanine Rijcks for their invaluable input during the data collection.

References

- Andreas, S. (1993). *Eye Movement Integration Booklet*. Boulder CO: Real People Press.
- Arntz, A., & Weertman, A. (1999). Treatment of childhood memories: theory and practice. *Behaviour Research and Therapy*, 37(8), 715–740.
- Austin, T. A. (2009). *Integral Eye Movement Therapy (IEMT) Practitioner Manual*. UK: The Association for IEMT Practitioners.
- Austin, T. A. (2015). Integral Eye Movement Therapy. In E. S. Neukrug (Ed.) *The SAGE Encyclopedia of Theory in Counseling and Psychotherapy* (p. 539-541). Thousand Oaks, CA: SAGE Publications, Inc.
- Beaulieu, D. (2003). *Eye Movement Integration Therapy: The Comprehensive Clinical Guide*. Williston, VT: Crown House Publishing.
- Bellis, M.A., Hughes, K., Ford, K., Rodriguez, R.G., Sethi, D., & Passmore, J. (2019). Life course health consequences and associated annual costs of adverse childhood experiences across Europe and North America: a systematic review and meta-analysis. *The Lancet Public Health*, 4(10), e517 - e528. [https://doi.org/10.1016/S2468-2667\(19\)30145-8](https://doi.org/10.1016/S2468-2667(19)30145-8)
- Mulej Bratec, S., Bertram, T., Starke, G., Brandl, F., Xie, X., & Sorg, C. (2020). Your presence soothes me: a neural process model of aversive emotion regulation via social buffering. *Social cognitive and affective neuroscience*, 15(5), 561–570. <https://doi.org/10.1093/scan/nsaa068>

- Carletto, S., Malandrone, F., Berchialla, P., Oliva, F., Colombi, N., Hase, M., Hofmann, A., & Ostacoli, L. (2021). Eye movement desensitization and reprocessing for depression: a systematic review and meta-analysis. *European Journal of Psychotraumatology*, 12(1), 1894736. <https://doi.org/10.1080/20008198.2021.1894736>
- Constantino, M. J., Višlá, A., Coyne, A. E., & Boswell, J. F. (2018). A meta-analysis of the association between patients' early treatment outcome expectation and their posttreatment outcomes. *Psychotherapy (Chicago, Ill.)*, 55(4), 473–485. <https://doi.org/10.1037/pst0000169>
- Cuijpers, P., & Cristea, I. A. (2016). How to prove that your therapy is effective, even when it is not: A guideline. *Epidemiology and Psychiatric Sciences*, 25, 428–435. <https://doi.org/10.1017/s2045796015000864>
- de Voogd, L. D., Kanen, J. W., Neville, D. A., Roelofs, K., Fernández, G., & Hermans, E. J. (2018). Eye-Movement Intervention Enhances Extinction via Amygdala Deactivation. *The Journal of Neuroscience : the official journal of the Society for Neuroscience*, 38(40), 8694–8706. <https://doi.org/10.1523/JNEUROSCI.0703-18.2018>
- Jongh, A. de, & Broeke, E. ten (2012). *Handboek EMDR. Een geprotocolleerde behandelmethode voor de gevolgen van psychotrauma*. Amsterdam: Pearson
- De Vries, G.-J., & Olf, M. (2009). The lifetime prevalence of traumatic events and posttraumatic stress disorder in the Netherlands. *Journal of Traumatic Stress*, 22(4), 259–267. <https://doi.org/10.1002/jts.20429>
- Dijk, H.H., & Kiernan, C.K. (2023). Maatschappelijke kosten van trauma [Societal costs of trauma]. *Tijdschrift voor Psychiatrie*, 65(9), 563–567.
- GGZ Standaarden (2020). *Zorgstandaard Psychotrauma- en stressorgerelateerde stoornissen*. Published online on 28-11-2024. Retrieved at <https://www.ggzstandaarden.nl/zorgstandaarden/psychotrauma-en-stressorgerelateerde-stoornissen/introductie>
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: Exposure to corrective information. *Psychological Bulletin*, 99(1), 20–35. <https://doi.org/10.1037/0033-2909.99.1.20>
- Fredrickson B. L. (2001). The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. *The American Psychologist*, 56(3), 218–226. <https://psycnet.apa.org/doi/10.1037/0003-066X.56.3.218>
- Fredrickson, B. L., Mancuso, R. A., Branigan, C., and Tugade, M. M. (2000). The Undoing Effect of Positive Emotions. *Motivation and Emotion*, 24(4), 237–258. <https://doi.org/10.1023/A:1010796329158>
- Geschwind, N., Keasberry, E., Voncken, M., Lobbestael, J., Peters, M., Rijkeboer, M., & van Heugten-van der Kloet, D. (2024). Imagery rescripting: The value of an added positive emotion component. *Journal of Behavior Therapy and Experimental Psychiatry*, 84, 101958. <https://doi.org/10.1016/j.jbtep.2024.101958>
- Hackmann, A. (2011). Imagery rescripting in posttraumatic stress disorder. *Cognitive and Behavioral Practice*, 18(4), 424–432.

- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 44(2), 227-239. <https://doi.org/10.1348/014466505X29657>
- Holmes, E.A., James, E.L., Coode-Bate, T., & Deeprose, C. (2009). Can playing the computer game “Tetris” reduce the build-up of flashbacks for trauma? A proposal from cognitive science. *PLoS One*, 4, e4153. <https://doi.org/10.1371/journal.pone.0004153>
- Hørlyck, L. D., Bisby, J. A., King, J. A., & Burgess, N. (2019). Wakeful rest compared to vigilance reduces intrusive but not deliberate memory for traumatic videos. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-49634-8>
- James, E.L., Bonsall, M.B., Hoppitt, L., Tunbridge, E.M., Geddes, J.R., Milton, A.L., & Holmes, E.A. (2015). Computer game play reduces intrusive memories of experimental trauma via reconsolidation-update mechanisms. *Psychological Science*, 26, 1201–1215. <https://doi.org/10.1177/0956797615583071>
- Koenen, K. C., Ratanatharathorn, A., Ng, L., McLaughlin, K. A., Bromet, E. J., Stein, D. J., Karam, E. G., Meron Ruscio, A., Benjet, C., Scott, K., Atwoli, L., Petukhova, M., Lim, C. C. W., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Bunting, B., Ciutan, M., de Girolamo, G., Degenhardt, L., ... Kessler, R. C. (2017). Posttraumatic stress disorder in the World Mental Health Surveys. *Psychological Medicine*, 47(13), 2260–2274. <https://doi.org/10.1017/S0033291717000708>
- Landin-Romero, R., Moreno-Alcazar, A., Pagani, M., & Amann, B. L. (2018). How Does Eye Movement Desensitization and Reprocessing Therapy Work? A Systematic Review on Suggested Mechanisms of Action. *Frontiers in Psychology*, 9, 1395. <https://doi.org/10.3389/fpsyg.2018.01395>
- Lovibond, P. F., and Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335-343. [https://doi.org/10.1016/0005-7967\(94\)00075-u](https://doi.org/10.1016/0005-7967(94)00075-u)
- Lytle, R. A., Hazlett-Stevens, H., & Borkovec, T. D. (2002). Efficacy of Eye Movement Desensitization in the treatment of cognitive intrusions related to a past stressful event. *Journal of anxiety disorders*, 16(3), 273–288. [https://doi.org/10.1016/s0887-6185\(02\)00099-3](https://doi.org/10.1016/s0887-6185(02)00099-3)
- McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne Effect: a randomised, controlled trial. *BMC medical research methodology*, 7, 30. <https://doi.org/10.1186/1471-2288-7-30>
- Matthijssen, S.J.M.A., van Beerschoten, L.M., de Jongh, A., Klugkist, I.G., & van den Hout, M. (2019). Effects of “Visual Schema Displacement Therapy” (VSDT), an abbreviated EMDR protocol and a control condition on emotionality and vividness of aversive memories: Two critical analogue studies. *Journal of Behavior Therapy and Experimental Psychiatry*, 63, 48-56. <https://doi.org/10.1016/j.jbtep.2018.11.006>

- Moore, J. M., and Manea, A. I. (2018). The use of Integral Eye Movement Therapy (IEMT) in the treatment of psoriasis- induces skin eruption - A case study. *Journal of Experiential Psychotherapy*, 21 (3), 71-81.
- Shapiro, F. (1989). Efficacy of the eye movement desensitization procedure in the treatment of traumatic memories. *Journal of Traumatic Stress*, 2(2), 199-223. <https://doi.org/10.1002/jts.2490020207>
- Shapiro, F. (2018). Eye movement desensitization and reprocessing: Basic principles, protocols, and procedures (3rd ed.). New York: Guilford Press.
- van der Spuy, C., & DuPlessis van Breda, A. (2019) An exploratory study on the use of Eye Movement Integration Therapy for treating trauma in early childhood in South Africa. *Child Care in Practice*, 25(2), 157-174.
- Ten Brinke, A. (2023). *Handboek IEMT*. Eindhoven: IEMT-training.
- Ten Broeke, E., De Jongh, A. & Hornsveld, H. (2023). EMDR standaardprotocol. Published on www.emdr.nl
- Van der Wijngaart, R. (2021). *Imagery Rescripting: Theory and Practice*. Pavilion Publishing and Media Ltd.
- Wadji, D. L., Martin-Soelch, C., & Camos, V. (2022). Can working memory account for EMDR efficacy in PTSD?. *BMC psychology*, 10(1), 245. <https://doi.org/10.1186/s40359-022-00951-0>
- Wolpe, J. (1969). *The practice of behavior therapy (2nd ed.)*. New York: Pergamon Press.

FEASIBILITY AND USABILITY OF A VR-BASED TRAINING FOR THERAPEUTIC RELATIONSHIP SKILLS IN FUTURE PSYCHOTHERAPISTS: A PILOT STUDY

Dana-Andreea Cristolțan^{1, 2*}

¹ The International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health, Babeș-Bolyai University, Cluj- Napoca, România

² Kind Strategies SRL, Cluj-Napoca, Romania, 400445

Abstract

Empathy and therapeutic alliance are key components of evidence-based psychotherapy and strong predictors of positive outcomes. Yet future psychologists often lack practical experience, underscoring the need for training that integrates theory with practice. Virtual reality (VR) offers a safe, controlled environment to rehearse relational skills. This pilot single-group pre-post study examined the feasibility and usability of the *Future Psychotherapists* VR application, designed to enhance empathy and alliance-building skills. Fifty-five participants, predominantly students, completed one VR-based training session interacting with a virtual client in a predefined scenario. Simulator sickness was assessed pre- and post-session, whereas usability and perceived cognitive workload were measured post-intervention. As an exploratory aim, state empathy (cognitive, affective, and associative components) was assessed before and after the session. Simulator sickness showed a slight increase from pre to post, but remained low overall. Usability ratings were high and perceived cognitive workload was moderate. For the exploratory aim, significant pre-post improvements emerged across all empathy components. Expert evaluations supported the scenario's relevance for therapeutic skill development. Overall, the *Future Psychotherapists* VR tool appears to be a feasible and usable method for training empathy and alliance-related skills in psychology students and may help bridge the gap between academic training and clinical practice by offering realistic opportunities to practice core therapeutic competencies.

* Correspondence concerning this article should be addressed to Dana-Andreea Cristolțan, The International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health, Babeș-Bolyai University, Str. Republicii 37, 400015 Cluj-Napoca, Romania.

ORCID: 0009-0003-5402-012X

Email: dana.cristoltan@ubbcluj.ro

Keywords: Virtual reality; Empathy; Therapeutic alliance; VR training; Future psychotherapists; Skills development

A robust body of evidence highlights the therapeutic relationship as the strongest determinant of successful psychotherapy outcomes, with an effectiveness that is both transtheoretical and largely transdiagnostic, extending across therapeutic modalities and a wide range of clinical problems (Norcross & Karpiak, 2025). These findings place the development of relational competencies at the center of psychotherapist training, pointing to the necessity of immersive and experiential approaches that mirror real-world therapeutic interactions. Research on psychotherapy outcomes indicates that client improvement is influenced by extra-therapeutic factors, expectancy effects, specific treatment techniques, and the client–therapist relationship. Among these, the therapeutic relationship—often referred to as a common factor—has been identified as the most significant contributor to positive outcomes from the therapist’s side (Lambert & Barley, 2001). The APA’s Interdivisional Task Force on Evidence-Based Relationships and Responsiveness highlights that the therapeutic relationship is as critical to successful outcomes as the specific treatment techniques employed (Norcross & Lambert, 2011). In 2018, Norcross and Lambert revisited the 2011 findings and reaffirmed the presence of core relationship factors, including therapeutic alliance collaboration, empathy, goal consensus, positive regard and affirmation, collecting and delivering client feedback and cohesion in group therapy (Norcross & Lambert, 2018). Particular attention is given in this paper to the therapeutic alliance and empathy, because previous studies have shown promising results and because a clearer hierarchical model of therapeutic-relationship elements has been proposed, in which the therapeutic alliance, empathy, and cohesion are superordinate constructs (Horvath et al., 2016, as cited in Norcross & Lambert, 2018).

Empathy is defined by the APA Dictionary of Psychology (2023) as “the ability to understand the thoughts, feelings, motivations, or behaviors of another person from their perspective.” Similarly, therapeutic alliance is defined as “the cooperative working relationship between client and therapist, including shared goals, consensus on tasks, and a strong interpersonal bond.” A meta-analysis by Elliott et al. (2018) shows that empathy is a moderately strong predictor of psychotherapy outcomes, and the therapeutic alliance exhibits a solid positive association with outcomes (Flückiger et al., 2018). Although empathy and the therapeutic alliance are highlighted as central to the success of the therapeutic process, research shows that developing these core elements in psychologists under supervision requires more than conceptual understanding. A meta-analysis found that didactic instruction alone has a significant effect on empathy development, and that combining didactic components with rehearsal and observation yields even stronger results (Ngo et al., 2022). Similarly, although training programs report awareness of the alliance’s importance, only a few programs include evidence-based

methods in their curricula. Best practices would integrate manuals, alliance-focused courses, observation of recorded sessions, and active feedback (Constantino et al., 2017).

However, despite such well-documented recommendations, educational psychology programs offer limited practical experience, resulting in novice therapists who feel inadequately prepared for real-world clinical challenges (Rosén, 2019). In this context, Rosén (2019) also emphasizes the importance of deliberate practice. Considering the recommendations made by researchers to prioritize active learning and behavioral rehearsal in psychotherapist training (Beidas & Kendall, 2010; Sudak & Brenner, 2024), we developed a VR-based training application for developing empathic skills and strengthening the therapeutic alliance during individual training, in order to reduce the gap that aspiring psychologists often feel between theory and practice. We chose virtual reality because it offers the opportunity to recreate the real world or to create different worlds that allow users to understand concepts and to perform certain tasks repeatedly, all within a safe environment (Chittaro & Ranon, 2007). Studies show that VR environments provide an effective way to build skills in mental-health interviewing and diagnosis through role-play (Lowell & Alshammari, 2019), enhance students' understanding of age-related health challenges and foster greater empathy toward older adults experiencing vision or hearing impairments or Alzheimer's disease (Dyer, Swartzlander, & Gugliucci, 2018), and improve communication confidence post-training (Sapkaroski et al., 2022), among other benefits. Effective VR learning tools should include feedback and reflection opportunities (Huang et al., 2024). We therefore intend to combine summative and formative feedback, which has been shown to improve learner outcomes (Ismail et al., 2022). Virtual reality interventions have been demonstrated in recent meta-analyses to effectively train healthcare professionals in the assessment and treatment of mental health disorders (Steen et al., 2024), outperforming traditional and other digital training methods in improving post-intervention knowledge and skills (Kyaw et al., 2019; Martingano et al., 2021). We identified two meta-analyses examining VR-based empathy training, although neither focuses specifically on psychotherapists or other healthcare professionals. One reported a statistically significant improvement in perspective-taking following VR exposure, with no notable change in overall empathy levels (Ventura et al., 2020), while the other showed that VR enhances emotional but not cognitive empathy—evoking compassionate feelings without encouraging users to adopt others' perspectives (Martingano et al., 2021).

Regarding the therapeutic alliance, we found no studies directly examining its development in VR-based training. However, existing research highlights its importance even within virtual environments (Varšová & Juřík, 2024; Meyerbröker & Emmelkamp, 2008) and demonstrates that therapeutic presence in VR can positively influence the working alliance (Im, Jo, & Lee, 2024). Despite evidence supporting empathy and therapeutic alliance as key to effective psychotherapy, and

the promise of VR for skill development, no studies have yet focused on using VR to train future psychotherapists in these essential relational skills. In this manuscript, the term future psychotherapists refers to psychology trainees, including university students and early-career therapists in supervised practice, whereas the VR application evaluated in this study is titled *Future Psychotherapists*.

Accordingly, the present pilot study had two aims. The first aim was to assess feasibility and usability of the VR application in future psychotherapists. The second aim was exploratory and examined pre–post changes in state empathy after a single VR session. Although the application targets both empathy and alliance-building skills, therapeutic alliance was not assessed with a standardized questionnaire in this pilot sample. Instead, alliance-related components were examined through expert content validation.

In line with the first aim, we expected low cybersickness, high perceived usability, and low-to-moderate perceived workload. For the secondary, exploratory aim, we expected pre–post changes in state empathy.

Methods

Design

This pilot study used a single-group pretest–posttest design to evaluate the virtual reality application *Future Psychotherapists*. The primary outcomes were feasibility and usability. Feasibility was operationalized as cybersickness measured both before and after the VR session using the Simulator Sickness Questionnaire (SSQ). Usability was evaluated post-session using the System Usability Scale (SUS) and the NASA Task Load Index (NASA-TLX) for perceived workload. The secondary outcome was exploratory and examined pre–post change in state empathy measured with the State Empathy Scale following a single VR session. We did not administer a therapeutic alliance measure, but alliance components were addressed during the expert validation phase.

Ethical Consideration

This study was approved by the Research Ethics Subcommittee of the University (Approval No. 470, dated 16.05.2024). Informed consent was obtained from all participants. All procedures involving human participants were conducted in accordance with the APA Ethical Principles of Psychologists and Code of Conduct.

Participants

The sample included 55 participants comprising undergraduate and master's-level psychology students as well as a small number of early-career therapists in supervised practice. Participants were recruited from the Faculty of Psychology based on the following inclusion criteria: (a) aged 18 or older; (b) currently enrolled in undergraduate or master's-level psychology programs or engaged in supervised practice as early-career therapists. Exclusion criteria included: (c) history of epilepsy, photosensitivity, or severe mental disorders; and (d) significant discomfort during VR tasks. The final sample consisted of 55 participants who completed all study stages.

Procedure

Participant recruitment was conducted through in-class announcements and a post on the Faculty's official communication channels, and interested individuals completed a Google Forms link that supported two stages: the scheduling phase and the participant selection phase. After registering for the study by providing informed consent and meeting the eligibility criteria, participants began with a brief familiarization session with the virtual reality application conducted in the research lab. They then viewed a standardized introductory clip within the VR-based training application (i.e. the first segment of the training; replayed at scenario onset) in which a virtual client presented the reason for attending a psychotherapy session. At this stage, participants only listened to the virtual client and then proceeded to the next step. Following this, the participants completed a set of baseline (T1) questionnaires: the Simulator Sickness Questionnaire (Kennedy et al., 1993) and the State Empathy Scale (Shen, 2010). Participants completed these questionnaires outside (desktop computer in the research lab) VR after the introductory clip to ensure all participants received the same client information before the baseline (T1) assessment and before returning to VR for the interactive task. The next step, they engaged in an experimental task in VR. This task consisted of a psychotherapy session in which participants interacted with a virtual client through an interactive scenario. At the end of the session, participants received feedback with both summative and formative components, based on their answers. If an incorrect answer was selected, the correct option was revealed to support learning. Finally, participants completed the post-intervention set of questionnaires: the Simulator Sickness Questionnaire (Kennedy et al., 1993), the State Empathy Scale (Shen, 2010), the NASA Task Load Index (Hart & Staveland, 1988), and the System Usability Scale (Brooke, 1996). There was a single VR session, and the procedure lasted 30-45 minutes in total.

Virtual reality–based intervention training description

Participants experienced the VR scenario through a Meta Quest 2 (Meta Platforms, Inc., 2020) VR headset. The virtual environment simulates an individual therapy session. The client presents mild symptomatology (mild depression and relational conflicts). The clinical scenario was presented as pre-recorded 2D video segments displayed in the head-mounted display against a minimal 360° black background (cinema-like). Response alternatives were presented as 3D interface elements in front of the user at each decision node. A minimalist background was used to standardize exposure and minimize irrelevant contextual information.

The application's script is implemented as an interactive branching narrative via discrete video segments. In each segment, the virtual patient describes their symptomatology, and at each decision node the user selects the response they deem most appropriate for the client from the predefined options on screen (see Figure 1). For an example of a client dialogue and corresponding response options, see Appendix A. After they choose a response, the virtual client will give them a predefined response and move to the next video segment.

The responses were designed to target empathic exploration, empathic understanding, empathic affirmation, and therapeutic alliance–building tasks. In total, there were 16 decision nodes: 9 focused on empathy-building and 7 on therapeutic alliance development. After users completed all decision nodes, they received summative feedback detailing the number of correct responses for each targeted skill (see Figure 2). The application also provided a formative feedback component: users could reopen the corresponding video clip where an incorrect response was selected, listen to it again, and view the expert-informed correct option (see Figure 3). In the review screen, the expert-informed correct option was highlighted in green and the participant's selected incorrect option in red, enabling direct comparison. After reviewing this feedback, users could restart the scenario and repeat it as many times as they wished to support deliberate practice, although in the present study, participants completed only a single VR session to assess initial feasibility. However, the application did not provide an explicit textual rationale explaining why one option was preferable and users were expected to derive the rationale by comparing their choice with the exemplar response. The application was developed using Unity (version 2019) and the clips were filmed in 4K at 30 fps with spatialized audio.

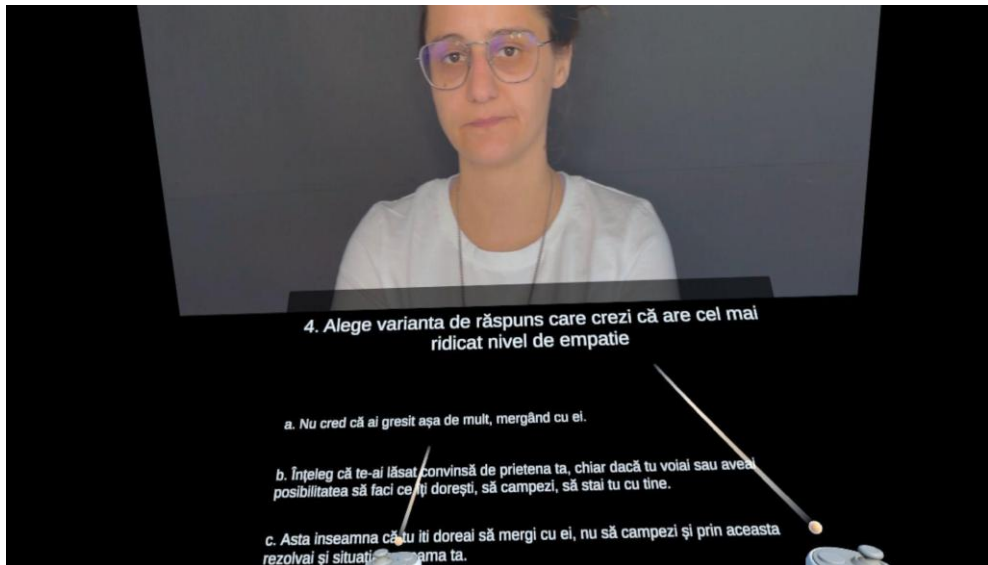


Figure 1. Example of a decision node in the branching narrative. The virtual patient presents a symptom, and the user selects the most appropriate response from multiple predefined options. The interface is in Romanian.

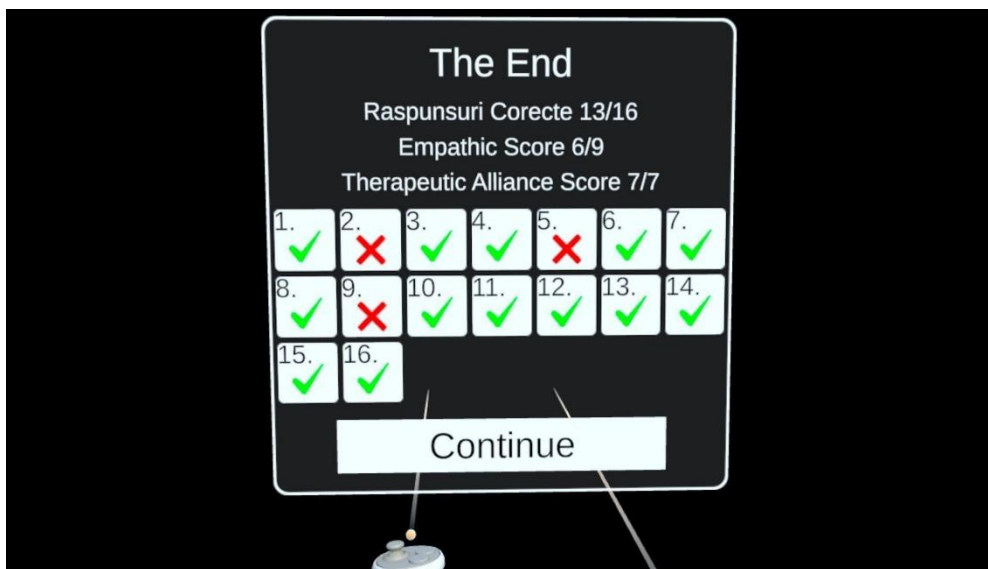


Figure 2. Summative feedback screen displaying the number of correct responses for each targeted skill.



Figure 3. Formative component feedback screen allowing users to review video segments in which incorrect responses were selected, along with the correct choices. The interface is in Romanian.

Expert contributors

In a first phase, two expert psychologists consulted on the most important elements of the therapeutic relationship that would be suitable for inclusion in a VR-based intervention targeting the development of relational therapeutic skills. A group of five expert psychologists then completed a two-part questionnaire to evaluate the application's textual content (the virtual client's speech and the response options). Following the rationale for expert content validation highlighted by different researchers (Palter et al., 2012; Ceberio et al., 2025), expert evaluations were included to establish a professional “gold standard” for the training scenarios. Since students are novices, expert validation was essential to ensure that the branching narrative options were clinically sound and provided a benchmark for proficiency that students cannot establish alone. Part A focused on content validation of the scenario and response options, whereas Part B captured experts' judgments regarding the application's educational/training potential rather than training efficacy.

Measures

We measured cybersickness using the *Simulator Sickness Questionnaire* (SSQ; Kennedy et al., 1993). It is the most commonly used instrument for measuring sickness induced by virtual reality environments (Bimberg et al., 2020). The SSQ comprises 16 items rated on a 4-point Likert scale (0 = None to 4 = Severe) and is grouped into three subscales—nausea, oculomotor disturbance, and disorientation—and a total score is calculated (Kennedy et al., 1993). We calculated internal consistency with Cronbach's alpha for each subscale. The results show good reliability: $\alpha = .72$ for nausea, $\alpha = .77$ for oculomotor, $\alpha = .83$ for disorientation, and $\alpha = .87$ for total score.

For assessing the usability of the system, the participants were asked to complete the *System Usability Scale* (SUS; Brooke, 1996), which has 10 items that measure the perceived usability on a 5-point Likert scale (1 = strongly disagree; 5 = strongly agree). SUS is a reliable tool for assessing the usability of a wide variety of products and services and generates an overall score on a scale from 0 to 100. We didn't calculate internal consistency because SUS is designed to provide a single overall usability score. The results were interpreted using the cut-off score of 68.

NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) was used to measure the cognitive load of participants, thus enabling us to see whether the application is easy to use from this point of view. It has 6 subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration, and a very specific method of calculating the score that requires good knowledge in using the instrument. It is one of the most widely used scales for measuring cognitive load (Rutledge et al., 2009). We did not calculate Cronbach's alpha for the NASA-TLX, as the scale is designed to measure multiple distinct aspects of workload, rather than a single underlying construct. Although the NASA-TLX does not include standardized cut-off points, previous research suggests that workload scores can be interpreted as follows: 0–20 = very low; 21–40 = low; 41–60 = moderate; 61–80 = high; and 81–100 = very high (Febiyani et al., 2021).

Participants' empathy was measured with the *State Empathy Scale* (SES; Shen, 2010). The scale includes three subscales—affective empathy, cognitive empathy, and associative empathy—and containing 12 items rated on a 5-point Likert scale (1 = not at all; 5 = completely). The instrument has high psychometric properties (Shen, 2010). Higher scores reflected a stronger level of empathy in participants. Internal consistency was evaluated with Cronbach's alpha for each subscale. The results present good reliability: $\alpha = .84$ for affective empathy, $\alpha = .84$ for cognitive empathy, and $\alpha = .72$ for associative empathy.

For measuring the quality of the application content's text section, we developed an instrument composed of two parts:

Part A: Evaluation of the therapeutic response texts on a 5-point Likert scale (1 = not at all, 5 = very much), measuring the degree to which each response exemplifies empathic skills (exploration, understanding, affirmation) or alliance-building behaviors (validation, collaboration, affective support) (Example item: “Read the text below and indicate to what extent the psychotherapist’s response reflects an empathic response or a response that can build the therapeutic alliance.”).

Part B: Global assessment of content clarity, relevance, and pedagogical utility through five Likert-scale items (1–5) addressing overall text adequacy, dialogue realism, scenario relevance, and potential for skill development (Example item: “The dialogues in the script are clear and easy to understand.”).

We calculated internal consistency using Cronbach’s alpha for each subscale. The results show different levels of reliability: $\alpha = .724$ for empathy, $\alpha = .78$ for therapeutic alliance, and $\alpha = .833$ for global assessment from part B.

Data Analysis

All data were processed using SPSS (version 20). Descriptive analyses (i.e., means, SDs) were conducted for all variables. Feasibility (SSQ), measured pre- and post-session, and usability outcomes (SUS and NASA-TLX), measured post-session only, were summarized descriptively. For the exploratory outcome, we used paired-samples *t*-tests to examine pre–post changes in state empathy. The significance level was set at $p < .05$. The expert evaluation, inter-rater agreement was assessed using Kendall’s coefficient of concordance (Kendall’s *W*), and expert ratings were summarized descriptively.

Results

Feasibility (Cybersickness)

Descriptive statistics for cybersickness subscales at pre-intervention (T1) and post-intervention (T2) are presented in Table 1. We can observe only slight increases for cybersickness (nausea, oculomotor disturbance, disorientation, and total SSQ score). To examine whether these changes were statistically significant, we conducted paired-samples *t*-tests. As shown in Table 1, there were no significant changes in cybersickness. Specifically, the scores for nausea ($t(54) = -0.16$, $p = .873$), oculomotor disturbance ($t(54) = -1.62$, $p = .111$), disorientation ($t(54) = -0.96$, $p = .340$), and the total SSQ score ($t(54) = -0.98$, $p = .334$) showed no significant differences.

Table 1. Means and Standard Deviations at T1 and T2 and Paired-Samples t-Tests (SSQ)

Subscale	T1 <i>M</i> (<i>SD</i>)	T2 <i>M</i> (<i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i>
Nausea (SSQ)	10.23 (17.40)	10.93 (28.29)	-0.16 (54)	.873
Oculomotor Disturbance (SSQ)	15.02 (18.13)	20.53 (25.04)	-1.62 (54)	.111
Disorientation (SSQ)	20.25 (35.82)	26.32 (46.61)	-0.96 (54)	.340
Total Cybersickness Score (SSQ)	16.86 (23.23)	21.49 (34.17)	-0.98 (54)	.334

Note. T1 = Pre-intervention; T2 = Post-intervention; *SD* = Standard Deviation; SSQ = Simulator Sickness Questionnaire (Kennedy et al., 1993). Higher scores indicate higher levels on each construct. Negative *t* values indicate higher post-intervention scores.

Usability and Workload (Post-intervention)

Furthermore, descriptive statistics for the post-intervention measures (SUS and NASA-TLX) are presented in Table 2. Usability (SUS) was rated very well, with a mean score of $M = 85.09$ ($SD = 9.54$), with scores above 68 (cut-off) indicating that the application is perceived as well-designed and easy to use. Regarding participants' perceived workload, we used the NASA-TLX. The results show that most participants reported low to moderate workload levels ($M = 32.50$, $SD = 17.13$).

Table 2. Descriptive Statistics for Post-Intervention Measures (T2; SUS, NASA-TLX)

Variable	<i>N</i>	Minimum	Maximum	<i>M</i>	<i>SD</i>
System Usability Scale (SUS)	55	57.50	100.00	85.09	9.54
NASA Task Load Index (NASA-TLX)	55	4.00	65.00	32.50	17.13

Note. SUS = System Usability Scale (Brooke, 1996); NASA-TLX = NASA Task Load Index (Hart & Staveland, 1988); *SD* = Standard Deviation. Higher SUS scores indicate better perceived usability. Lower NASA-TLX scores indicate lower perceived workload.

Expert Content Evaluation (Content validity)

The content validity of the application was assessed by five independent experts. Our purpose was expert consensus, not statistical inference. They evaluated empathy-related content, therapeutic alliance content, and general content quality. Table 3 presents descriptive statistics. The empathy-related content received a mean score of $M = 4.40$ ($SD = 0.30$). These results indicate high perceived relevance and clarity of the script used to foster state empathy. Moreover, the standard deviation indicates a high level of agreement between experts. The evaluation of the therapeutic alliance content also received a high mean score ($M = 4.00$, $SD = 0.17$), indicating that the VR application may facilitate the development of a therapeutic alliance. The content quality (Part B) received the highest ratings with a mean of M

= 4.76 ($SD = 0.33$). This suggests that experts considered the application content to be appropriate, coherent, and clinically relevant. Given the small variability across raters and consistently high scores across all categories, these results support the strong content validity of the application from a professional standpoint. This conclusion is reinforced by the high inter-rater agreement, Kendall's $W = .96$, $\chi^2(2) = 9.58$, $p = .008$, indicating consistent expert ratings across dimensions.

Table 3. Descriptive Statistics for Empathy, Therapeutic Alliance, and Content Evaluation

Variable	<i>N</i>	Minimum	Maximum	<i>M</i>	<i>SD</i>
Total Empathy	5	4.00	4.78	4.40	0.30
Total Therapeutic Alliance	5	3.86	4.29	4.00	0.17
Content Score	5	4.20	5.00	4.76	0.33

Note. SD = Standard Deviation. Scores are based on a 1–5 Likert scale where higher values indicate more favorable evaluations.

Empathy Outcomes (Exploratory)

Table 4 summarizes empathy scores at pre-intervention (T1) and post-intervention (T2). Mean scores increased from T1 to T2 for all three empathy components. To examine whether these changes were statistically significant, we conducted paired-samples t -tests. As shown in Table 4, data indicated significant improvements in all three dimensions of empathy: affective empathy ($t(54) = -4.11$, $p < .001$), cognitive empathy ($t(54) = -4.71$, $p < .001$), and associative empathy ($t(54) = -3.56$, $p = .001$).

Table 4. Means and Standard Deviations at T1 and T2 and Paired-Samples t -Tests (SES)

Subscale	T1 <i>M</i> (<i>SD</i>)	T2 <i>M</i> (<i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i>
Affective Empathy (SES)	14.20 (3.12)	16.22 (3.10)	-4.11 (54)	< .001
Cognitive Empathy (SES)	17.56 (2.35)	19.13 (1.28)	-4.71 (54)	< .001
Associative Empathy (SES)	15.02 (2.86)	16.69 (2.48)	-3.56 (54)	.001

Note. T1 = Pre-intervention; T2 = Post-intervention; SD = Standard Deviation; SES = State Empathy Scale (Shen, 2010). Higher scores indicate higher levels on each construct. Negative t values indicate higher post-intervention scores.

Discussion

The present study investigated the feasibility and usability of the Future Psychotherapists application, and secondarily examined changes in state empathy

from pre-intervention (T1) to post-intervention (T2). The VR-based training was developed to support future psychotherapists in cultivating therapeutic relationship skills.

The low level of cybersickness observed in our application aligns with findings from the literature indicating that increasing cybersickness can compromise in-VR performance and, by extension, the feasibility of VR-based training. Symptom severity has been associated with poorer visuospatial working memory and degraded psychomotor performance (e.g., slower reaction times) during immersion (Kourtesis et al., 2024). Moreover, evidence suggests that VR sickness can reduce attention and impair task performance (Mimnaugh et al., 2023). In addition, VR sickness is associated with non-completion of exposure: a systematic review and meta-analysis reported a mean dropout rate of 15.6% attributable to VR sickness (Saredakis et al., 2020), which may reduce effective exposure to training sessions. Taken together, the low cybersickness observed in our sample suggests that simulator sickness is unlikely to have meaningfully interfered with task performance or engagement, supporting the feasibility of delivering this VR training in a single-session format.

Regarding the usability of our application, participants reported high usability and low-to-moderate cognitive workload. High scores on the System Usability Scale (>68) suggest that participants found the interface easy to use, which is essential for user engagement in virtual learning environments. Moreover, the fact that participants reported low to moderate workload aligns with the research of Kirschner (2002), which affirms that an appropriate workload when using VR is preferable because it helps participants optimize cognitive load for learning.

Moreover, expert evaluations confirmed the application's potential to train empathy and therapeutic alliance. From a content perspective, expert validation highlighted the application's clarity, relevance, and clinical usefulness. Kendall's $W = .96$ indicates excellent agreement among raters.

In relation to our exploratory objective regarding empathy, a single VR training session significantly improved state empathy (cognitive, affective, associative) without increasing cybersickness. Regarding cognitive and affective empathy, our results align with previous research showing that VR can enhance perspective-taking, and at the same time, with the findings reported in the meta-analysis by Martingano et al. (2021). As for associative empathy, this dimension has been less studied. However, from a functional perspective, it plays an important role in fostering social bonding and the development of relationships (Davis, 1994), acting as a bridge between perception and action in the emergence of state empathy (Shen, 2010). Looking at all these results, we can conclude that the current VR-based training application can facilitate multiple layers of empathic response, which is likely to be effective in developing core relational skills.

The VR environment was intentionally designed in a minimalist manner to reduce extraneous processing and keep attention on the instructional objective (i.e.,

to choose the best response option in terms of empathy and therapeutic alliance), in line with multimedia learning principles such as the coherence principle (Fiorella & Mayer, 2021). We therefore prioritized a clear and well-timed presentation of task-relevant information over contextual environmental realism (Fiorella & Mayer, 2021). Recent evidence from educational VR, including technical and vocational contexts, suggests that within immersive 3D VR environments, increasing environmental realism does not necessarily improve learning performance relative to more minimalist designs and may even support motivation by reducing distractions (Larmuseau et al., 2025). Moreover, CAMIL suggests that learning outcomes depend on interactions among presence/agency, affective responses, and cognitive load, rather than on visual fidelity alone (Makransky & Petersen, 2021). Although our implementation relied primarily on video presented in an HMD, this minimalist approach was expected to influence the training effects by reducing competing stimuli and helping participants focus on identifying and selecting the most appropriate therapeutic response.

Considering the results of this study regarding the feasibility and usability of the application, as well as the experts' evaluation of the application's content and the observed change from T1 to T2, these findings contribute to the expanding body of literature supporting the use of VR-based interventions in mental health education (Steen et al., 2024; Kyaw et al., 2019) for training key elements of the psychotherapeutic relationship, such as empathy and therapeutic alliance, among students or novice practitioners. Both formative and summative feedback components reinforce the idea of an application capable of fostering empathy and the therapeutic alliance among future psychologists and early-career practitioners, while also enabling deliberate practice through repeated scenario completion. For example, Sapkaroski et al. (2022) found that immersive VR training improved students' confidence and communication abilities, while Dyer et al. (2018) showed increased empathy in medical students after exposure to aging simulations. Practically, this method can serve as a complementary tool in the training of psychotherapists, responding to current gaps in experiential learning and supporting the development of relational competencies through active engagement in simulated clinical scenarios.

Limitations and future directions

The study has several limitations that should be acknowledged. One limitation is the fact that the pretest-posttest design does not include a control group. This was a deliberate decision, as the main objective at this exploratory phase was to assess the VR application's usability—according to ISO 9241-11 (International Organization for Standardization [ISO], 2018) and its overall feasibility for single-session VR training, rather than to test its efficacy compared to alternative interventions. In line with this definition, our evaluation prioritized feasibility and

user experience, assessing whether users could successfully engage with and complete the training tasks, rather than providing evidence of superior training outcomes. Hence, causal inferences should be made with caution. The second limitation refers to the fact that the intervention was conducted through a single VR session, and therefore, the durability of the effects over time remains unknown. Another limitation is that, although deliberately minimalist, the video-based VR format—consisting primarily of a 2D video presented in an HMD against a black 360° background—may yield lower immersion/presence and fewer contextual cues than a shared, fully modeled virtual therapy room. This should be considered when interpreting the extent to which the observed effects generalize to spontaneous, real-world therapist–client interaction, and should be examined in future research. An additional limitation is that the feedback is primarily evaluative (correct/incorrect), accompanied by an expert-informed exemplar, without a detailed rationale explaining why one option is preferable. Although this structure supports guided practice through repeated scenario completion and provides both formative and summative feedback on empathy- and alliance-consistent responding, deeper reflective learning and transfer to open-ended clinical communication were not established in this pilot. Although the application was designed to foster both empathy and therapeutic alliance, we did not include a standardized measure for alliance. This decision was primarily based on the fact that most participants were psychology students without direct client experience. Although a small number of early-career therapists in supervised practice were included, this subgroup was too small to support separate analyses. Therefore, the findings should be interpreted as primarily reflecting psychology students. Since existing alliance questionnaires typically require respondents to reflect on prior therapeutic interactions, they would not have been appropriate for this population. Finally, future studies should incorporate more behaviorally grounded outcome measures (e.g., observer-rated performance, behavioral coding of responses in role-plays or standardized patient interactions, and performance-based tasks requiring open-ended responding) to evaluate training efficacy beyond self-report measures. These limitations point toward the need for future research in which VR-based training is used in more complex studies, such as randomized controlled trials, and potentially in a longitudinal design as well.

Conclusions

This study demonstrates the feasibility and usability of the VR-based training application for relational skills in future psychotherapists (i.e., psychology students). The application was rated as usable, engaging, and with a low to moderate level of participants' workload. Moreover, the VR intervention produced significant

changes in empathy across all three dimensions (cognitive, affective, and associative). All these aspects suggest that the current VR-based training application can complement traditional methods, so that future psychotherapists have more learning opportunities for the skills that currently represent a significant gap between theory and practice. Given that the results are preliminary, as this is a pilot study, future research should include larger and more diverse samples, control groups, and longitudinal follow-up to evaluate the long-term impact of VR-based relational skills training. To the best of our knowledge, this study introduces a novel educational tool that can serve as a model for future innovations in psychotherapy education and professional development.

Author's Notes

Competing Interests: The authors declare no conflicts of interest. Dana Cristolțan is the administrator of Kind Strategies SRL. The company had no roles in the design of the study, in the collection, analysis or interpretation of data; in the writing of the manuscript, or in the decision to publish the articles. The paper reflects the views of the scientists and not the company.

Funding: The application used in this research was supported by Kind Strategies SRL.

References

- American Psychological Association. (2023). Empathy. In *APA Dictionary of Psychology*. <https://dictionary.apa.org/empathy>
- American Psychological Association. (2023). Therapeutic alliance. In *APA Dictionary of Psychology*. <https://dictionary.apa.org/therapeutic-alliance>
- Beidas, R. S., & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, 17(1), 1–30. <https://doi.org/10.1111/j.1468-2850.2009.01187.x>
- Bimberg, P., Weissker, T., & Kulik, A. (2020). On the usage of the Simulator Sickness Questionnaire for virtual reality research. In *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)* (pp. 464–467). IEEE. <https://doi.org/10.1109/VRW50115.2020.00098>
- Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, I. L. McClelland, & B. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). Taylor & Francis.

- Ceberio, I., Al-Rashaida, M., García, M., Lopez Paz, J. F., Salgueiro, M., Passi, N., Pavel, H., & Amayra, I. (2025). Content and face validity in virtual reality with children: A validation in five steps+1 of a wheelchair basketball game. *Frontiers in Virtual Reality*, 5, 1505630. <https://doi.org/10.3389/frvir.2024.1505630>
- Chittaro, L., & Ranon, R. (2007). Web3D technologies in learning, education and training: Motivations, issues, opportunities. *Computers & Education*, 49(1), 3–18. <https://doi.org/10.1016/j.compedu.2005.06.002>
- Constantino, M. J., Morrison, N. R., Coyne, A. E., & Howard, T. (2017). Exploring therapeutic alliance training in clinical and counseling psychology graduate programs. *Training and Education in Professional Psychology*, 11(4), 219–226. <https://doi.org/10.1037/tep0000157>
- Davis, D. (1994). *Empathy: A social psychological approach*. Brown & Benchmark Publishers.
- Dyer, E., Swartzlander, B. J., & Gugliucci, M. R. (2018). Using virtual reality in medical education to teach empathy. *Journal of the Medical Library Association*, 106(4), 498–500. <https://doi.org/10.5195/jmla.2018.518>
- Elliott, R., Bohart, A. C., Watson, J. C., & Murphy, D. (2018). Therapist empathy and client outcome: An updated meta-analysis. *Psychotherapy*, 55(4), 399–410. <https://doi.org/10.1037/pst0000175>
- Febiyani, A., Febriani, A., & Ma'Sum, J. (2021). Calculation of mental load from e-learning student with NASA TLX and SOFI method. *Jurnal Sistem dan Manajemen Industri*, 5(1), 35–42. <https://doi.org/10.30656/jsmi.v5i1.2789>
- Fiorella, L., & Mayer, R. E. (2021). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 185–198). Cambridge University Press. <https://doi.org/10.1017/9781108894333.019>
- Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, 55(4), 316–340. <https://doi.org/10.1037/pst0000172>
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). North-Holland.
- Huang, Y.-M., Wang, W.-S., Lee, H.-Y., Lin, C.-J., & Wu, T.-T. (2024). Empowering virtual reality with feedback and reflection in hands-on learning: Effect of learning engagement and higher-order thinking. *Journal of Computer Assisted Learning*, 40, 1413–1427. <https://doi.org/10.1111/jcal.12959>
- Im, S., Jo, D., & Lee, S. M. (2024). Exploring the impact of therapeutic presence on working alliance in Metaverse counseling. *The Counseling Psychologist*, 52(5), 698–724. <https://doi.org/10.1177/00110000241234630>
- International Organization for Standardization. (2018). *Ergonomics of human-system interaction—Part 11: Usability: Definitions and concepts (ISO 9241-11:2018)*. <https://www.iso.org/standard/63500.html>

- Ismail, S. M., Rahul, D. R., Patra, I., & Rezvani, E. (2022). Formative vs. summative assessment: Impacts on academic motivation, attitude toward learning, test anxiety, and self-regulation skill. *Language Testing in Asia*, 12(1), Article 40. <https://doi.org/10.1186/s40468-022-00191-4>
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology*, 3(3), 203–220.
- Kirschner, P. A. (2002). Cognitive load theory: Implications of cognitive load theory on the design of learning. *Learning and Instruction*, 12(1), 1–10. [https://doi.org/10.1016/S0959-4752\(01\)00014-7](https://doi.org/10.1016/S0959-4752(01)00014-7)
- Kourtesis, P., Papadopoulou, A., & Roussos, P. (2024). Cybersickness in virtual reality: The role of individual differences, its effects on cognitive functions and motor skills, and intensity differences during and after immersion. *Virtual Worlds*, 3(1), 62–93. <https://doi.org/10.3390/virtualworlds3010004>
- Kyaw, B., Saxena, N., Posadzki, P., Vsetecková, J., Nikolaou, C., George, P., Divakar, U., Masiello, I., Kononowicz, A., Zary, N., & Tudor Car, L. (2019). Virtual reality for health professions education: Systematic review and meta-analysis by the Digital Health Education Collaboration. *Journal of Medical Internet Research*, 21(1), e12959. <https://doi.org/10.2196/12959>
- Lambert, M. J., & Barley, D. E. (2001). Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy: Theory, Research, Practice, Training*, 38(4), 357–361. <https://doi.org/10.1037/0033-3204.38.4.357>
- Larmuseau, C., De Leersnijder, L., Rotsaert, T., Boel, C., Demanet, J., & Schellens, T. (2025). Beyond realism: Rethinking VR design for optimal learning in technical and vocational secondary education. *Computers & Education: X Reality*, 6, 100098. <https://doi.org/10.1016/j.cexr.2025.100098>
- Lowell, V. L., & Alshammari, A. (2019). Experiential learning experiences in an online 3D virtual environment for mental health interviewing and diagnosis role-playing: A comparison of perceived learning across learning activities. *Educational Technology Research and Development*, 67(4), 825–854. <https://doi.org/10.1007/s11423-018-9632-8>
- Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 33, 937–958. <https://doi.org/10.1007/s10648-020-09586-2>
- Martingano, A. J., Herrera, F., & Konrath, S. (2021). Virtual reality improves emotional but not cognitive empathy: A meta-analysis. *Technology, Mind, and Behavior*, 2(1). <https://doi.org/10.1037/tmb0000034>
- Meta Platforms, Inc. (2020). *Meta Quest 2* [Virtual reality headset]. <https://www.meta.com/quest/>
- Meyerbröcker, K., & Emmelkamp, P. M. (2008). Therapeutic processes in virtual reality exposure therapy: The role of cognitions and the therapeutic alliance. *CyberTherapy & Rehabilitation*, 1(3), 247–257.

- Mimnaugh, K. J., Center, E. G., Suomalainen, M., Becerra, I., Lozano, E., Murrieta-Cid, R., Ojala, T., LaValle, S. M., & Federmeier, K. D. (2023). Virtual reality sickness reduces attention during immersive experiences. *IEEE Transactions on Visualization and Computer Graphics*, 29(11), 4394–4404. <https://doi.org/10.1109/TVCG.2023.3320222>
- Ngo, H. L., Sokolović, N., & Jenkins, J. M. (2022). Teaching empathy to mental health practitioners and trainees: Pairwise and network meta-analyses. *Journal of Consulting and Clinical Psychology*, 90(11), 851–860. <https://doi.org/10.1037/ccp0000773>
- Norcross, J. C., & Karpiak, C. P. (2025). Psychotherapy relationships. In S. Hupp & D. F. Tolin (Eds.), *Science-based therapy: Raising the bar for empirically supported treatments* (pp. 381–398). Cambridge University Press.
- Norcross, J. C., & Lambert, M. J. (2011). Evidence-based therapy relationships. In J. C. Norcross (Ed.), *Psychotherapy relationships that work: Evidence-based responsiveness* (2nd ed., pp. 3–22). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199737208.003.0001>
- Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy*, 55(4), 303–315. <https://doi.org/10.1037/pst0000193>
- Palter, V. N., Graafland, M., Schijven, M. P., & Grantcharov, T. P. (2012). Designing a proficiency-based, content validated virtual reality curriculum for laparoscopic colorectal surgery: A Delphi approach. *Surgery*, 151(3), 391–397. <https://doi.org/10.1016/j.surg.2011.07.084>
- Rosén, E. (2019). *Learning how to learn: Psychology trainees' self-compassion while implementing deliberate practice, with FIT at a psychotherapy training clinic* [Master's thesis, Umeå University].
- Rutledge, T., Stucky, E., Dollarhide, A., Shively, M., Jain, S., Wolfson, T., Weinger, M. B., & Dresselhaus, T. (2009). A real-time assessment of work stress in physicians and nurses. *Health Psychology*, 28(2), 194–200. <https://doi.org/10.1037/a0013145>
- Sapkaroski, D., Mundy, M., & Dimmock, M. R. (2022). Immersive virtual reality simulated learning environment versus role-play for empathic clinical communication training. *Journal of Medical Radiation Sciences*, 69(1), 56–65. <https://doi.org/10.1002/jmrs.555>
- Saredakis, D., Szpak, A., Birckhead, B., Keage, H. A. D., Rizzo, A., & Loetscher, T. (2020). Factors associated with virtual reality sickness in head-mounted displays: A systematic review and meta-analysis. *Frontiers in Human Neuroscience*, 14, 96. <https://doi.org/10.3389/fnhum.2020.00096>
- Shen, L. (2010). On a scale of state empathy during message processing. *Western Journal of Communication*, 74(5), 504–524. <https://doi.org/10.1080/10570314.2010.512278>
- Steen, C. W., Söderström, K., Stensrud, B., Nylund, I. B., & Siqueland, J. (2024). The effectiveness of virtual reality training on knowledge, skills and attitudes of health care professionals and students in assessing and treating mental health

- disorders: A systematic review. *BMC Medical Education*, 24(1), 480. <https://doi.org/10.1186/s12909-024-05423-0>
- Sudak, D. M., & Brenner, A. M. (2024). The future of psychotherapy training. *Academic Psychiatry*, 48(5), 511–515. <https://doi.org/10.1007/s40596-024-02022-5>
- Varšová, K., & Juřík, V. (2024). Fostering therapeutic alliance and long-term benefits through virtual collaboration in VRET. *International Psychological Applications Conference and Trends (InPACT2024)*.
- Ventura, S., Badenes-Ribera, L., Herrero, R., Cebolla, A., Galiana, L., & Baños, R. (2020). Virtual reality as a medium to elicit empathy: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 23(10), 667–676. <https://doi.org/10.1089/cyber.2019.0681>

MATERNAL DEPRESSION AND CHILDREN'S SOCIO-EMOTIONAL DEVELOPMENT IN THE FIRST TWO YEARS OF LIFE: A META-ANALYSIS

*Mădălina Ruxandra Costin^{1*}, Diana Tăut¹,
Adriana Smaranda Băban¹, and Sebastian Pinteă¹*

¹ Department of Psychology, Babeş-Bolyai University

Abstract

This meta-analysis synthesizes evidence on whether maternal depressive symptoms during pregnancy and the postpartum period are associated with children's socio-emotional development in the first two years of life and examines sources of variability across studies.

Following PRISMA guidelines, we searched PubMed, Web of Science, and Scopus up to 01 December 2025. We included quantitative studies reporting maternal depression and child socio-emotional outcomes assessed at ≤ 24 months. Random-effects models were used, and we tested potential moderators to account for between-study heterogeneity.

Twenty-nine independent studies ($N = 22,655$ mother-child dyads) were included. Maternal depression showed a reliable overall small-to-moderate association with less favorable early socio-emotional development. Associations were observed for both prenatal and postnatal depression. Effects were generally stronger when outcomes were mother-reported rather than observational, stronger for socio-emotional problems than for competencies, and tended to increase when outcomes were assessed later within the first two years. Studies with older, more educated, and partnered samples typically reported smaller associations.

Maternal depressive symptoms are reliably linked to children's socio-emotional development in the first two years. Differences across studies indicate that measurement approach, outcome type, timing, and family context shape effect sizes, underscoring the value of perinatal screening and early intervention.

Keywords: Maternal depression, perinatal depression, socio-emotional development, early childhood, meta-analysis

* Correspondence concerning this article should be addressed to Mădălina Ruxandra Costin, Department of Psychology, Babeş-Bolyai University, Str. Republicii, nr. 37, Cluj-Napoca, Cluj, Romania.

Email: madalina.costin@ubbcluj.ro

Children's first life experiences are fundamental for their subsequent short- and long-term development (Likhar et al., 2022). In particular, the first two years of life are not only marked by rapid and major physical changes (Ilyka et al., 2021) but also by cognitive, social, and emotional development, which, in turn, are cornerstones for later psychological adjustment and well-being in the following years. In early childhood, socio-emotional development encompasses three main areas of development: emotion regulation, other-oriented socio-emotional processes, and self-oriented socio-emotional processes (Denham & Zinsser, 2014).

These elements are connected and influence each other (Tervahartiala et al., 2024). Children with harmonious socio-emotional development manage to (1) understand the feelings and behaviors of those around them and respond to the demands of the environment; (2) successfully regulate their own emotions in relation to personal needs and in alignment with interpersonal relational goals; and (3) understand and develop functional relationships with attachment figures and with other persons (Paley & Hajal, 2022; Thümmeler et al., 2022; Speidel et al., 2023; Malik & Marwaha, 2024).

In summary, the development of these skills from an early age supports children in forming stronger interpersonal relationships, collaborating effectively with others, maintaining better overall mental health, developing adaptive coping mechanisms, and, more broadly, enhancing their resilience and engagement in prosocial behaviors (Jones et al., 2015; Oberle & Schonert-Reichl, 2017; Izett et al., 2021). Moreover, these skills facilitate greater access to educational and economic opportunities (Frosch et al., 2021) and can serve as protective factors against a range of difficulties, including disruptive behaviors, depression, anxiety, emotion dysregulation, and both externalizing and internalizing problems (Izett et al., 2021; Colomeischi et al., 2022; Anderson et al., 2023; Zięba-Kołodziej & Oliiynyk, 2023; Smorti et al., 2024).

In the present study, we operationalize socio-emotional development in the first two years of life as a general measure of early socio-emotional adjustment that spans both strengths and difficulties. This approach is consistent with early-childhood reviews showing that socio-emotional development is typically described in terms of both adaptive competencies and emerging difficulties, alongside the development of self-regulation (Halle & Darling-Churchill, 2016). Specifically, we treat socio-emotional competencies (e.g., emotion regulation, social competence, and prosocial/other-oriented functioning) and socio-emotional problems (e.g., internalizing and externalizing symptoms, withdrawal, or broader socio-emotional difficulties) as two related facets of early socio-emotional development.

Many social determinants, extending beyond biological and genetic aspects, significantly impact children's socio-emotional development (Kirby, 2020; Thijssen et al., 2023). From an ecological perspective (Bronfenbrenner & Morris, 2007), children's development is shaped by interrelated systems that range from broad socio-cultural structures to their most immediate family interactions. Within the microsystem, children are directly influenced by their home and play environments, the immediate settings in which they grow, learn, and play. A particularly crucial component of this system is the parent-child relationship (Frosch et al., 2021). This relationship begins even before birth and continues to shape children's socio-emotional growth across development (Bianciardi et al., 2023). Given its role, any factor affecting this dynamic can have substantial developmental consequences. One such factor is maternal depression in the perinatal period (Goeglein & Yatchmink, 2020; Milgrom et al., 2023; Morales et al., 2023). We operationalize maternal depression to include both depressive symptoms assessed with validated instruments and clinically diagnosed depression in the perinatal period.

During the perinatal period, which includes pregnancy and the postpartum phase, women undergo substantial biological, psychological, and social changes (Balaji et al., 2007; Schaffir, 2016; Atkinson & Teychenne, 2019; Wilson & Lee, 2023). Prenatally, depressive symptoms can alter the intrauterine environment and disrupt the hypothalamic-pituitary-adrenal (HPA) axis, which regulates stress hormone production (Beijers et al., 2014; Brummelte et al., 2017). These effects can make the fetus hypersensitive to stressors in the environment (Bind et al., 2022). In addition, high cortisol levels in depressed mothers during pregnancy can lead to adverse birth outcomes such as prematurity, low birth weight, and intrauterine growth restriction, all of which increase long-term cognitive, emotional, and socio-emotional deficits in childhood (Traylor et al., 2020; Ghimire et al., 2021). Furthermore, depressive symptoms can impair a mother's ability to prepare cognitively and emotionally for motherhood, affecting her sensitivity and attunement to her child postpartum (Pearson et al., 2012; McAndrew, 2019; Hildingsson & Rubertsson, 2022).

Postnatally, mothers may withdraw, become less physically and emotionally responsive to their child's needs, misinterpret the child's cues, and encounter difficulties in establishing a warm and nurturing relationship (Letourneau et al., 2010; Sidor et al., 2011; McAndrew, 2019).

These adverse dynamics have a major impact on the first two years of children's lives, the stage in which they develop their limbic system and the right hemisphere of the brain (Glynn & Baram, 2019). Previous studies have shown that infants of depressed mothers frequently show signs of anger, defensive coping strategies, and passivity as a result of inadequate maternal support and responsiveness (Field et al., 2006; Tronick & Reck, 2009). These contexts can have

overall negative effects on children, impacting their growth and well-being, diminishing their capacity to develop effective emotion regulation skills, and increasing their risk of developing internalizing and externalizing problems (Goodman et al., 2011; Letourneau et al., 2019; Urizar & Muñoz, 2022; Morales et al., 2023).

While research has concentrated on the influence of postnatal depression on child development, more recent studies have also begun to examine the effects of prenatal depression on children's development (Takács et al., 2020). In this review, prenatal refers to depressive symptoms assessed during pregnancy, whereas postnatal refers to symptoms assessed after childbirth.

Following this direction is both promising and challenging because the relationship between maternal depression across the perinatal period and children's socio-emotional development is multifaceted (Junge et al., 2017). It involves a range of factors, including cultural, socioeconomic, relational, and biological ones (Cui et al., 2018; Chorbajian et al., 2020; Harris & Santos, 2020). At the same time, this brings new challenges. It is more difficult to come to clear conclusions from individual studies when they differ in design, sampling, how and when maternal depression is assessed, and which socio-emotional outcomes are targeted (Cuijpers et al., 2010; Morales et al., 2023). Furthermore, the strength of associations may be influenced by the child's age at the time of assessment, contextual risks (e.g., socioeconomic disadvantage), and cultural norms (Rogers et al., 2020; Jaramillo et al., 2017).

Consequently, we need to synthesize findings to emphasize overall patterns, identify potential moderators, and evaluate whether the timing of the depression is associated with differences in children's socio-emotional development during their first two years of life (Takács et al., 2020).

Through this meta-analysis, we aim to systematically review existing research on how maternal depression impacts the socio-emotional development of children during their first two years of life. We conceptualize socio-emotional development through three fundamental components: emotion regulation, other-oriented socio-emotional processes, and self-oriented socio-emotional processes. By synthesizing empirical evidence from both prenatal and postnatal periods, we aim to 1) identify and analyze the specific impact of maternal depression on children's socio-emotional development during the first two years of life; 2) differentiate the effects of prenatal and postnatal depression on children's socio-emotional development; and 3) explore potential methodological and sociodemographic moderators that shape the relationship between maternal depression and children's socio-emotional development.

Method

Search strategy

We followed the guidelines outlined in PRISMA for reporting systematic reviews and meta-analyses (Page et al., 2021). A comprehensive search was conducted up to 01 December 2025 in three sources: PubMed, Web of Science, and Scopus for published studies. The search strategy included combinations of the following terms: “postnatal depression” OR “postpartum depression” OR “maternal depression” OR “prenatal depression” OR “perinatal depression” OR “mother’s mental health.” We combined these terms with the Boolean operator “AND” with terms related to child socio-emotional development, such as “child development” OR “child/infant/young child socio-emotional development” OR “child/infant/young child social development” OR “child/infant/young child emotional development.” Truncation symbols were used in searches when appropriate. Two authors assessed the final eligibility of studies, with the senior author settling any disagreements. The full database-specific search strings and Boolean operators are provided in Supplementary Material Appendix 1.

Inclusion and Exclusion Criteria

For this meta-analysis, studies were included if they met the following criteria: (1) written in English and published as a scientific article; (2) contained at least one analysis investigating the relationship between maternal depression and the child's socio-emotional development; (3) employed validated measurements; and (4) reported adequate statistics to allow for the calculation of effect sizes necessary for the meta-analysis.

Studies were excluded based on the following criteria: (1) systematic reviews and book chapters; (2) articles that assessed socio-emotional development outcomes in children older than 24 months; (3) studies that examined cognitive or physical development and health outcomes in children, rather than socio-emotional outcomes; (4) studies that investigated other mental health problems in mothers (e.g. combined distress, depression scores without separable depression effects) or other family members (e.g., fathers, grandparents); (5) articles focusing on the socio-emotional development of adopted children; and (6) studies where socio-emotional development was only mentioned in the text or discussion, but not directly measured.

The selection process is illustrated in the flowchart diagram (see Figure 1).

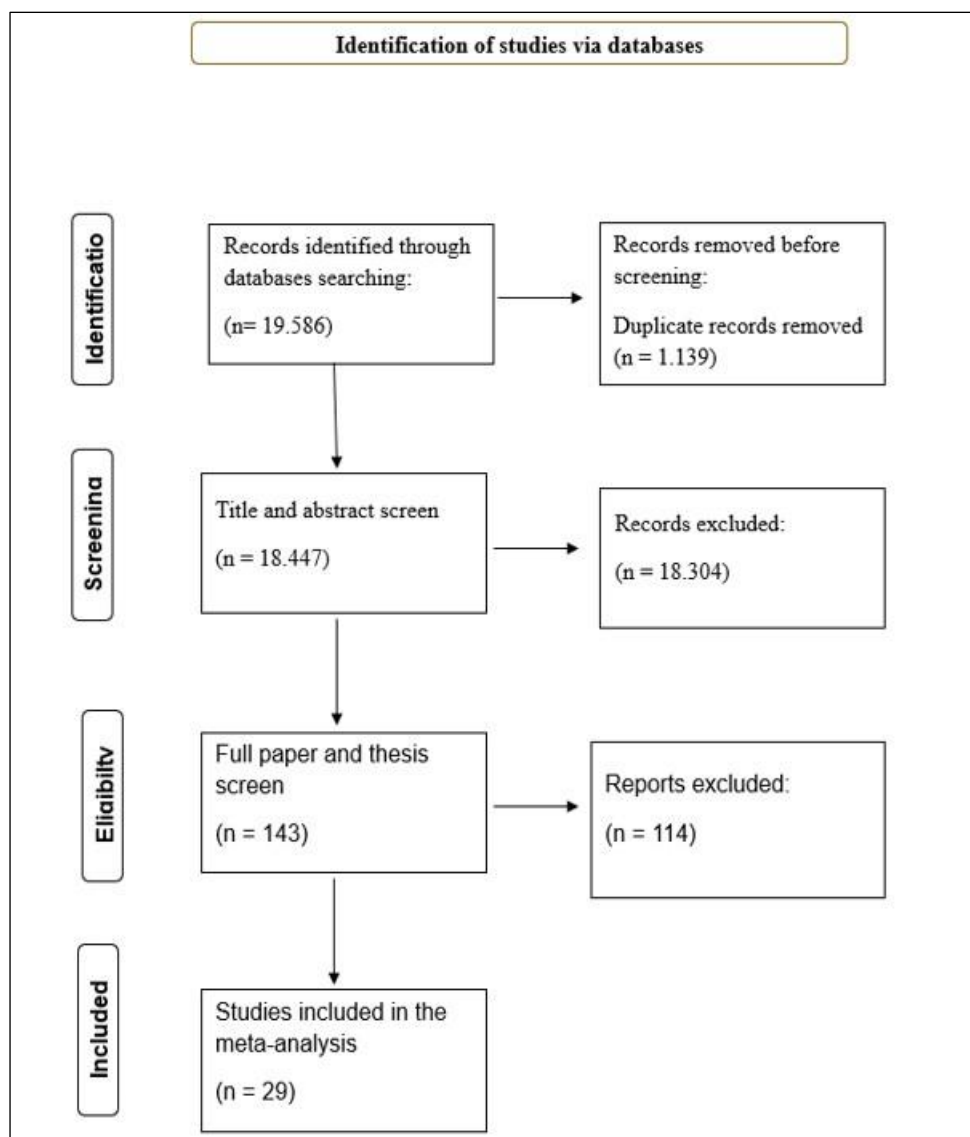


Figure 1. The PRISMA flow chart diagram describing the selection of studies.

Coding procedure

To systematically analyze the studies included in this meta-analysis, we extracted and coded a comprehensive set of variables from each study. The variables coded were: study ID (authorship, year of publication), study design, sample size, and the reported effect size(s) and related statistical information (including odds ratios, confidence intervals, p-values, and other effect-size-related metrics) needed for synthesis and comparability across studies.

Depression variables

Depression-related variables were coded to specify whether depression was prenatal or postnatal, alongside the instrument used to assess depression and the type of measurement (e.g., self-reported). We also coded the timing of depression assessment, recorded in weeks, whenever this information was available.

Children's socio-emotional development variables

For outcomes related to children's socio-emotional development, we coded the instruments used, the type of measurement (e.g., observational vs. reported by the mother), and the timing of the outcome assessment, recorded in weeks, when reported.

Coding of moderators

We initially coded a broad array of potential moderators. However, we focused our analyses on moderators that were reported in a sufficient number of studies and for which the distribution across studies was relatively balanced. The moderators retained for analysis were coded as follows:

Mother's age

This variable was coded as a continuous measure, using the mean age (in years) of mothers reported for each study.

Mother's Education (> 12 years)

Education was coded as a continuous measure reflecting the percentage of mothers with more than 12 years of education. When studies reported education in categories, we extracted the category that best matched >12 years and recorded the corresponding proportion when available.

Mothers living with a partner

Relational status was coded as a continuous measure representing the percentage of mothers living with a partner, typically operationalized as married and/or cohabiting participants, depending on how each study reported relationship status.

Mother's primigravida status

Primigravida (women in their first pregnancy) status was coded as a continuous measure reflecting the percentage of primigravida mothers reported in each study sample.

Study quality assessment

Study quality was assessed using the Newcastle–Ottawa Scale (NOS) (Wells et al., 2013). We used the NOS cohort version for longitudinal/cohort designs and an adapted NOS version for cross-sectional designs. The cohort NOS captures three domains (Selection, Comparability, Outcome; total score range 0–9), and the cross-sectional adaptation captures analogous domains (Selection, Comparability, Outcome; total score range 0–8 in our coding sheet). Each item was coded as 0/1 according to prespecified criteria. We expressed study quality as a percentage of the maximum score: $(\text{NOS total} / \text{NOS max}) \times 100$, where higher values indicate higher quality. Study quality was also examined as a continuous measure.

Coding of outcome direction

During the coding process, we observed that the children's socio-emotional development variables were typically either child competencies, such as emotional regulation, or child problems, such as internalizing and externalizing behaviors. To maintain consistency in our overall analyses, for all overall analyses, we recoded the direction of correlations such that associations congruent with the existing literature received a positive sign, whereas associations incongruent with the literature received a negative sign. This decision was made to prevent positive correlations with child problems from canceling out negative correlations with child competencies if the original signs were retained. The original signs of the correlations were only preserved in analyses where the type of variable (competencies vs. problems) was treated as a moderating variable.

The studies included in this meta-analysis are described in Table 1.

Table 1

The studies included in the meta-analysis and their main features.

Study	Country	Design	N	Depression phase	Depression measure	Depression time (wks)	Outcome domain	Outcome measure	Outcome method	Outcome time (wks)	Mother age (M)	Edu >12y (%)	Partner (%)	Primigravida (%)	Quality (%)
Schavo & Perosa, 2020	Brazil	prospective longitudinal	139	Postnatal	BDI	26–61	Personal social delay	DDST-II	Observational	26–61	NS	66	83	40	77.78
Anizan & Kunun, 2021	Turkey	cross-sectional	537	Postnatal	BSI	101	Internalizing behaviors; Externalizing behaviors	CBCL	Reported by mother	101	32.03	74	NS	NS	60.0
Bates et al., 2020	The United States	prospective longitudinal	142	Postnatal	EPDS	48	Self-regulation	IBQ-R VSF	Reported by mother	48	26.3	96.9	50.6	NS	100.0
Behrendt et al., 2019	Germany	cross-sectional (correlational)	61	Postnatal	BDI	57	Social Emotional problems; Social Emotional competencies	BITSEA	Reported by mother	57	31.89	45	NS	61	60.0
Black et al., 2007	Bangladesh	longitudinal	221	Postnatal	CES-D	26–52	Orientation/Engagement; Emotional Regulation	BSID-II	Observational	26–52	28.1	NS	98.2	27.5	100.0
DiPietro et al., 2006	The United States	longitudinal	94	Prenatal	POMS; CES-D	24–32	Orientation/Engagement; Emotional Regulation	BSID-II	Observational	24.8	32.0	NS	95.7	NS	100.0
Dugnay et al., 2022	Canada	longitudinal	468	Prenatal; Postnatal	EPDS	25.5; 10.5	Social Emotional development	ASQSE-2	Reported by mother	10.5	30.0	NS	98.5	63.9	55.56
Edwards & Hans, 2016	The United States	longitudinal	196	Prenatal; Postnatal	CES-D	27.6; 104	Social Emotional problems	BITSEA	Reported by mother	104	18.3	NS	69	NS	77.78

Fredrikse et al., 2019	Norway	longitudinal	387	Prenatal, Postnatal	EPDS	NS	Internalizing problems; Externalizing problems	ITSEA	Reported by mother	78	30.3	77.1	95.9	54.8	77.78
Gao et al., 2007	New Zealand	longitudinal	1021	Postnatal	EPDS; GHQ-12	6-104	Internalizing problems; Externalizing problems	CBCL	Reported by mother	104	NS	63.6	80.8	27.4	77.78
Gravener et al., 2012	The United States	cross sectional	198	Postnatal	BDI	88	Internalizing problems; Externalizing problems	CBCL	Reported by mother	88	31.68	198	87.9	NS	40.0
Guyon-Harris et al., 2016	The United States	longitudinal	120	Prenatal, Postnatal	EPDS; BDI	33.5; 14-104	Internalizing problems; Externalizing problems; Social Emotional problems	BITSEA	Reported by mother	104	26.2	95.8	27.5	25	88.89
Koutra et al., 2013	Greece	longitudinal	470	Prenatal, Postnatal	EPDS	28-32; 8	Social Emotional development	BSID-III	Reported by mother	8-78	30.31	40.8	99.6	42.7	88.89
Haile et al., 2018	The United States	longitudinal	650	Postnatal	CES-D	104	Internalizing problems; Externalizing problems	CBCL	Reported by mother	104	NS	73	NS	NS	77.78
Hummel & Kiel, 2015	The United States	longitudinal	91	Postnatal	CES-D	104	Internalizing problems	ITSEA	Reported by mother	104	31.81	100	97	NS	66.67
Kamal & Kalita, 2010	India	longitudinal	10	Postnatal	EPDS	26	Social development	CSBS-DP	Reported by mother	26	28.05	61.1	100	27.8	44.44
Marucci et al., 2021	Italy	cross sectional	74	Postnatal	EPDS	74	Social Emotional development	BSID-III	Reported by mother	NS	NS	NS	88.5	NS	40.0
Mason et al., 2011	The United States	longitudinal	219	Postnatal	EPDS	12	Social Emotional development	ASQ SE-2	Reported by mother	30	23.5	77.7	22.7	NS	44.44

McDonald et al., 2018	Canada	longitudinal	1596	Postnatal	CES-D	104	Social Emotional delay	BITSEA	Reported by mother	104	NS	77.5	96.24	NS	77.78
Nix et al., 2021	Ireland	longitudinal	61	Prenatal; Postnatal	HAM-D	NA; 112	Social Emotional development; Social adaptive behaviour	BSID-III	Reported by mother	112	36.0	98.4	91.8	NS	66.67
Okinundu-Luwa et al., 2021	Congo	cross sectional	458	Postnatal	GDAS	38	Social withdrawal behavior	ADBB	Observational	38	28.0	98	100	NS	40.0
Pacheco & Figueiredo, 2012	Portugal	prospective	110	Prenatal	EPDS	NS	Regulation of state	NBAS	Observational	1	NS	NS	92.8	100	77.78
Palmer et al., 2013	The United States	longitudinal	697	Postnatal	EPDS	4–52	Social Emotional problems	BITSEA	Reported by mother	52	NS	36.8	41	NS	100.0
Smith-Nielsen et al., 2019	Copenhagen	longitudinal	69	Postnatal	EPDS	17	Social withdrawal behavior	ADBB	Observational	17	NS	100	97.1	100	55.56
The APON Team et al., 2019	Canada	longitudinal	634	Postnatal	EDS	13	Social withdrawal behavior; Aggressive behaviour; Externalizing behaviour; Internalizing behaviour	CBCL	Reported by mother	13	33.2	92.5	98	68	88.89
Almeida et al., 2012	Portugal	longitudinal	204	Prenatal	IACLIDE	NS	Social Emotional development	BITSEA	Reported by mother	52	29.0	24.8	71.5	NS	66.67
Korja et al., 2024	Finland	longitudinal	59	Prenatal; Postnatal	EPDS	NS	Externalizing behaviour; Internalizing behaviour	BITSEA	Reported by mother	107	31.7	96	NS	53	77.78

Rogers et al., 2023	Australia	longitudinal	1539	Prenatal, Postnatal	EPDS	8	Social Emotional development	BSID-III	Reported by mother	53	32.89	80	92	56	100.0
Chen et al., 2025	China	prospective	1213	Prenatal	PHQ-9	7–34	Personal social delay	ASQ-3	Reported by mother	23	30.35	83.4	NS	69.64	88.89

Note. Study = author, year; N = mother-child dyads; Depression phase = prenatal vs postnatal; Depression measure = depression instrument; Depression time (wks) = timing of depression assessment in weeks; Outcome domain = socio-emotional development domain; Outcome measure(s) = outcome instrument(s); Outcome method = observational vs reported by mother; Outcome time (wks) = timing of outcome assessment in weeks; Mother age (M) = mothers' mean age; Edu >1.2y (%) = percentage of mothers with >1.2 years of education; Partner (%) = percentage of mothers living with a partner; Primigravida (%) = percentage of primigravida mothers; Quality (%) = study quality percentage. For studies including both prenatal and postnatal depression, values separated by semicolons correspond to prenatal; postnatal assessments. Ranges (e.g., 26–61) indicate the time window reported in the original study. NS = not specified. BDI = Beck Depression Inventory; BSI = Brief Symptom Inventory; CES-D = Center for Epidemiologic Studies Depression Scale; EDS = Edinburgh Depression Scale; EPDS = Edinburgh Postnatal Depression Scale; GDAS = (as reported in the original study); GHQ-12 = General Health Questionnaire-12; HAM-D = Hamilton Depression Rating Scale; LACLIDE = (as reported in the original study); PHQ-9 = Patient Health Questionnaire-9; POMS = Profile of Mood States; ADBB = Alarm Distress Baby Scale; ASQ-SE-2 = Ages and Stages Questionnaire: Social-Emotional, Second Edition; ASQ-3 = Ages and Stages Questionnaire, Third Edition; BITSEA = Brief Infant-Toddler Social and Emotional Assessment; BSID-II/BSID-III = Bayley Scales of Infant Development, Second/Third Edition; CBCL = Child Behavior Checklist; CSBS-DP = Communication and Symbolic Behavior Scales-Developmental Profile; DDST-II = Denver Developmental Screening Test II; IBQ-R YSF = Infant Behavior Checklist; ITSEA = Infant-Toddler Social and Emotional Assessment; NBAS = Neonatal Behavioral Assessment Scale.

Data analysis

Statistical analyses were conducted using Comprehensive Meta-Analysis software, version 2. Given the heterogeneity of the studies presented below, we applied a random-effects model.

To evaluate publication bias, we employed two approaches: Rosenthal's Classic Fail-safe N and Begg and Mazumdar's rank correlation test. Rosenthal's Classic Fail-safe N estimates the number of unidentified or unpublished studies that would be required to render the meta-analysis non-significant ($p > 0.05$) (Rosenthal, 1979). Begg and Mazumdar's rank correlation test ranks studies by effect size and standard error and then calculates the correlation between these two variables (Begg & Mazumdar, 1994). This method helps determine if larger studies are consistently included regardless of effect size, whereas smaller studies may be included only if they show a significant effect size.

For moderator analysis, we used two strategies: heterogeneity tests for categorical variables (e.g., depression phase) and meta-regression tests for continuous variables (e.g., mothers' age), following the guidelines of Borenstein and colleagues (Borenstein et al., 2021).

Results

Study Characteristics

The present meta-analysis included 29 independent studies in total. The sample sizes of the included studies varied, ranging from 10 to 12,130 participants, comprising a total sample of 22,655 mother-child dyads. Regarding the study design, the majority ($n = 24$) utilized a longitudinal design, monitoring the same dyads over time, while 5 studies employed a cross-sectional design.

The included studies were taken in a diverse range of geographical contexts, although Western samples predominated. Most studies were conducted in North America ($n = 12$; primarily the United States and Canada) and Europe ($n = 9$; Denmark, Finland, Germany, Greece, Ireland, Italy, Norway, and Portugal), with fewer studies from Asia ($n = 3$; Bangladesh, China, India), Oceania ($n = 2$; Australia, New Zealand), South America ($n = 1$; Brazil), Africa ($n = 1$; Congo), and Turkey ($n = 1$).

Maternal depression was assessed exclusively during the prenatal period in 4 studies, during the postnatal period in 17 studies, and 8 studies included assessments at both time points. When the timing of assessment was reported in weeks, prenatal assessments ranged from 7 to 34 gestational weeks (median = 29.8), and postnatal assessments ranged from 4 to 112 weeks postpartum (median = 52).

All studies used standardized self-report depression measures. For prenatal depression ($n = 12$), EPDS was most common ($n = 7$), followed by CES-D ($n = 2$), with other instruments used in single studies (HAM-D, POMS, IACLIDE, and PHQ-9 variants). For postnatal depression, EPDS was again most frequently used ($n = 13$), followed by CES-D ($n = 5$) and BDI ($n = 4$), with additional measures used in single studies (GHQ-12, BSI, GDAS, and HAM-D).

Child socio-emotional development variables were assessed between 1 and 112 weeks post-birth (median = 57 weeks). Most studies relied on maternal-report measures ($n = 23$), whereas fewer used observational or clinician-administered assessments ($n = 6$). The most frequently used measure was BITSEA ($n = 7$), followed by CBCL ($n = 5$) and the Bayley Scales (BSID-III: $n = 4$; BSID-II: $n = 2$). Other measures included ADBB ($n = 2$), ASQ:SE-2 ($n = 2$), and ITSEA ($n = 2$), while ASQ-3, CSBS-DP, DDST-II, IBQ-R VSF, and NBAS were each used in one study.

Regarding sociodemographic characteristics, reporting varied across studies. Mean maternal age was available in 21 studies, ranging from 18.3 to 36.0 years, with a mean of 29.60 across reporting studies. The proportion of mothers with >12 years of education was reported in 24 studies (range 24.8%–100.0%); in 21 studies, more than 50% of the sample had >12 years of education. The proportion with <8 years of education was reported in 19 studies (range 0.0%–55.6%). The proportion of mothers living with a partner was reported in 24 studies (range 22.7%–100.0%), and the proportion of primigravida mothers was reported in 15 studies (range 25.0%–100.0%).

Heterogeneity analysis

The heterogeneity analysis performed for the distribution of effect sizes in our meta-analysis indicated a significant heterogeneity, $Q(28) = 369.45$, $p < .001$, $I^2 = 92.42$. Taking into account also the diversity related to the research designs and participants, we decided to perform all the data analysis using a random-effects model.

The overall effect size

In Figure 2, the forest plot of all studies included in the meta-analysis is presented. As can be seen, the correlation varies from -0.24 (Nix et al., 2021) to 0.914 (Kamal & Kalita, 2010). The overall effect size is $r = 0.204$, 95% $CI [0.147, 0.260]$, and statistically significant ($Z = 6.86$, $p < 0.001$), indicating a small to moderate effect size between maternal depression and children's socio-emotional development during the first two years of life.

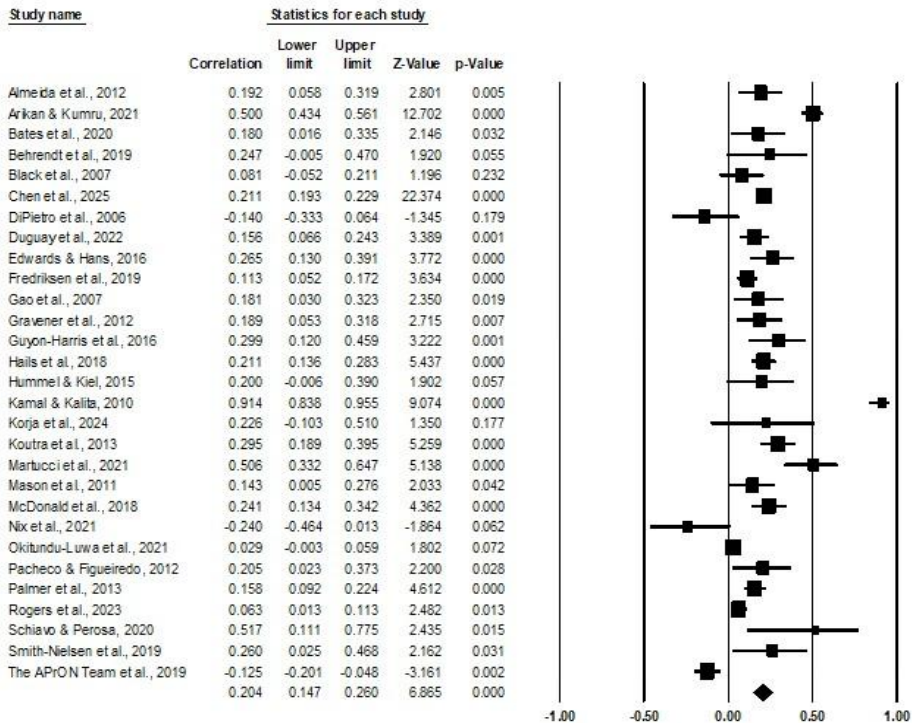


Figure 2. The forest plot of all studies included in the analysis and the overall effect size.

Publication bias analysis

To evaluate publication bias, we employed three approaches: Rosenthal's Classic Fail-safe N , Begg and Mazumdar's rank correlation test, and visual inspection of a funnel plot.

The Classic Fail-safe N of Rosenthal. Our meta-analysis incorporates data from 29 studies, which yields a z value of 19.70 and a corresponding two-tailed p value of 0.0001. The fail-safe N is 2902. This means that we would need to locate and include 2902 'null' studies in order for the combined two-tailed p value to exceed 0.05. In other words, there would be needed 100.1 missing studies with a null effect for every observed study for our effect to be nullified.

Begg and Mazumdar's rank correlation. This test was concerned with the potential relationship between the size of the studies and the effect size obtained by each one. The results revealed a nonsignificant Kendall's tau b of 0.21, with a two-tailed p value of 0.106 (based on continuity-corrected normal approximation), suggesting there is no tendency for studies that are more precise (and implicitly larger) to generate larger effect sizes.

Funnel plot inspection

The funnel plot appears largely consistent with an overall symmetrical distribution of studies around the pooled effect size, particularly among the more precise studies clustered at the top. While there is a slight visual imbalance, with somewhat more studies on the right side, this asymmetry is relatively modest and may be attributable to natural heterogeneity or random variation rather than clear evidence of publication bias. The majority of studies fall within the expected funnel boundaries, and although a few points lie further from the center, such dispersion is not uncommon in meta-analyses and does not necessarily indicate systematic bias. Overall, the plot suggests that the findings are reasonably robust, with no strong indication of substantial publication bias.

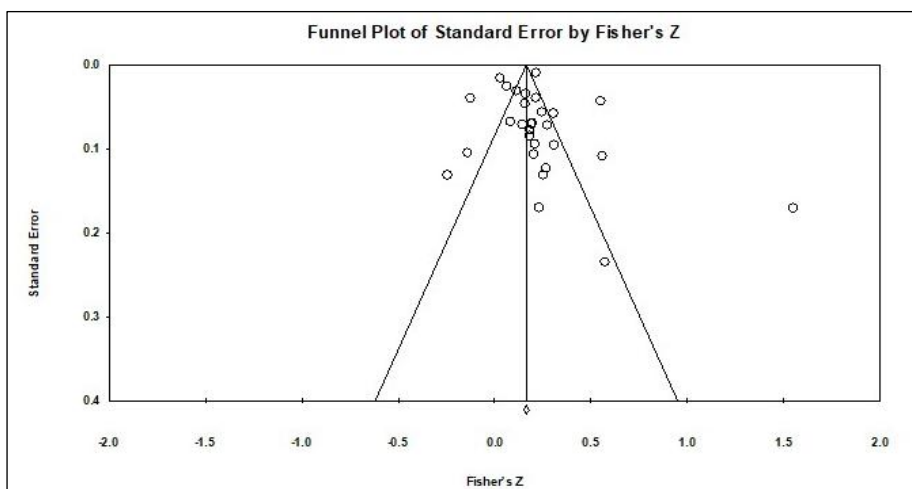


Figure 3. Funnel plot for the studies included in the meta-analysis.

Moderators analysis

Depression phase. In Table 2, we present the comparison between the overall effect size obtained in the prenatal phase and the overall effect size obtained in the postnatal phase of maternal depression.

Table 2. Comparison of effect sizes between postnatal and prenatal maternal depression

Phase of measurement	No. of studies	r	Limits of confidence interval (95%)		Z	p
			Lower	Upper		
Postnatal	25	0.229	0.157	0.298	6.12	0.001
Prenatal	12	0.141	0.076	0.204	4.26	0.001

As can be seen in Table 2, the effect sizes obtained in the prenatal and the postnatal phase are both statistically significant ($r = 0.14$, $p = 0.001$) for the prenatal phase and ($r = 0.22$, $p = 0.001$) for the postnatal phase, with no statistically significant differences between them $Q(1) = 3.027$, $p = 0.070$. Consequently, we conclude that the depression phase does not moderate the effect size.

Socio-emotional development measurement (observational vs reported by mother).

In Table 3, we present the analysis of the effect sizes as a function of the outcome measurement.

Table 3. The analysis of the effect sizes as a function of the children's socio-emotional development measurement method (observational vs. reported by mother).

Children's socio-emotional development measurement	No. of studies	r	Limits of confidence interval (95%)		Z	p
			Lower	Upper		
Observational	6	0.101	-0.012	0.212	1.75	0.079
Reported by mother	23	0.225	0.161	0.286	6.81	<.001

As Table 3 shows, when the outcomes were measured by observation, the effect size is small and non-significant ($r = 0.101$, $p = 0.079$). When children's socio-emotional development was reported by mothers, the effect size was moderate and statistically significant ($r = 0.225$, $p < .001$). The heterogeneity test between categories revealed marginally statistically significant differences, $Q(1) = 3.60$, $p = 0.058$. Consequently, we conclude that the way the outcomes were measured (observational vs. reported by mothers) has a tendency to moderate its relationship with maternal depression.

Children's socio-emotional development time (weeks post-birth)

The moderation analysis for outcome time was performed using a meta-regression procedure in which the outcome time was treated as a predictor and the effect size as a criterion. A total of 28 studies reported the timing of the outcome, and to standardize the data, we transformed all reported times into weeks. The analysis revealed that the outcome time is a significant positive predictor of the effect size ($b = 0.004$, $Z = 2.95$, $p = 0.003$). This finding suggests that as the time of outcome measurement increases, the strength of the relationship between maternal depression and child outcomes also increases. Therefore, outcome time plays a significant moderating role in this relationship.

Mother's age

Because most of the studies did not report their results as a function of age, we quantified the mean age of mothers and tested through a meta-regression to see if it predicted the effect size from each study. Our analysis revealed that age was a significant negative predictor of the effect sizes ($b = -0.015$, $Z = -6.37$, $p < .001$).

This finding suggests that as the age of the mother increases, the strength of the relationship between maternal depression and children's socio-emotional outcomes decreases. Thus, the mother's age is an important moderating variable in understanding this relationship.

Percentage of mothers with more than 12 years of education

The moderation analysis for the percentage of mothers with more than 12 years of education was conducted using a meta-regression procedure, with education level as the predictor and effect size as the criterion. The analysis revealed that the percentage of mothers with higher education was a significant negative predictor of the effect size ($b = -0.0009$, $Z = -3.99$, $p < .001$). Thus, as the percentage of mothers with more than 12 years of education increases, the effect size decreases. Therefore, the level of maternal education moderates the relationship between depression and child socio-emotional development.

Percentage of mothers living with a partner

We also conducted a meta-regression to assess the percentage of mothers living with a partner as a potential moderator. The percentage of partnered mothers was treated as the predictor and effect size as the criterion. The results showed that this variable was a significant negative predictor of the effect size ($b = -0.002$, $Z = -8.80$, $p < .001$). This suggests that as the percentage of mothers living with a partner increases, the effect size decreases. Consequently, the relationship between maternal depression and child outcomes is moderated by whether the mother is living with a partner.

Percentage primigravida mothers

The meta-regression analysis for the percentage of primigravida mothers (treated this variable as the predictor and effect size as the criterion). The analysis found that the percentage of primigravida mothers was not a significant predictor of the effect size ($b = 0.0002$, $Z = 0.61$, $p = 0.539$). Therefore, we found no evidence that primigravida status moderates the relationship between maternal depression and child socio-emotional outcomes.

Competencies vs problems

Table 4. The effect size as a function of competencies vs problems

Competences vs problems	No. of studies	r	Limits of confidence interval (95%)		Z	p
			Lower	Upper		
Competences	13	-0.142	-0.278	-0.001	-1.97	0.048
Problems	17	0.206	0.134	0.276	5.54	<.001

As Table 4 shows, when competencies were correlated with maternal depression, the effect size was small and significant ($r = -0.142, p = 0.048$). When problems were associated with depression, the effect size was small to moderate and statistically significant ($r = 0.206, p < .001$). The heterogeneity test between categories revealed statistically significant differences, $Q(1) = 18.56, p < .001$. Consequently, we will state that the type of outcome (competencies vs. problems) moderates its relationship with depression.

Study quality

The meta-regression analysis indicated that study quality was a statistically significant predictor of effect size, with a positive regression coefficient ($B = 0.001, Z = 4.49, p < .001$). This suggests that higher-quality studies tended to report slightly larger effect sizes.

Discussion

We aimed to examine the relationship between maternal depression and the socio-emotional development of children during their first two years of life and explore the potential moderators explaining the variability of the results. We analyzed a total of 29 studies. Our results indicated a statistically significant, small to moderate effect size. These results are consistent with previous findings in other research papers, showing that both prenatal and postnatal depression can negatively impact children's socio-emotional development (Goodman et al., 2011; Letourneau et al., 2019). Although both phases of maternal depression were detrimental and the postnatal effect size was numerically larger, the difference between postnatal and prenatal effects was not statistically significant. Consequently, our findings support the idea that depressive symptoms across the whole perinatal period represent a relevant risk for early socio-emotional development, rather than clearly indicating one period as more harmful than the other. One possible explanation is the continuity of depressive symptoms across pregnancy and postpartum, such that early symptoms may signal a more persistent vulnerability. In addition, prenatal and postnatal depression may operate through partially different pathways that converge on similar outcomes (e.g., prenatal biological sensitivity and postnatal caregiving disruptions) (Beijers et al., 2014; Bind et al., 2022). In particular, when depressive symptoms persist into the postpartum period, the risk may be expressed more directly through the early caregiving environment. During this challenging period of significant changes, mothers experiencing depressive symptoms may struggle to provide nurturing and supportive interactions or establish healthy routines, essential aspects for a child's development (Milgrom et al., 2023; Saharoy et al., 2023). These findings emphasize the need for comprehensive support throughout the perinatal period to support the balanced development of children and the well-being of the mothers.

The moderator analysis accentuated the complexity of the relationship between maternal depression and children's socio-emotional development, being influenced by several factors.

One significant factor was the method used to assess socio-emotional development. When mothers reported their children's socio-emotional development, the association with maternal depression was stronger; in contrast, observational assessments had smaller, non-significant results. This may suggest a negative bias in mothers' reporting, possibly influenced by depressive symptoms, which can affect attention and cognitive processes (Li et al., 2023). To increase the validity of the measurements and have a more comprehensive understanding of children's socio-emotional development, the studies should use both self-report and observational methods.

Another significant factor was the timing of socio-emotional development measurement, the effect of depressive symptoms having a greater impact as the children get older. As children grow older, the cumulative exposure to maternal depressive symptoms and prolonged disruptions in caregiving interactions may lead to socio-emotional difficulties becoming more observable and stable, particularly in domains such as emotion regulation and externalizing behavior. This aligns with Morales and colleagues (Morales et al., 2023), who found that mothers with a higher chronicity of depressive symptoms reported more significant externalizing and internalizing problems in their children.

Additionally, maternal age also moderated the relationship. The older age of mothers buffered the adverse effects of depressive symptoms on the socio-emotional development of children. Previous research supports this direction; advancing in age brings greater life experience, psychological maturity, and resources, important aspects that help parenting skills (Trillingsgaard & Sommer, 2018; Morales et al., 2023).

Similarly, mothers' education plays a moderating role. Mothers with higher levels of education showed weaker associations between depression and socio-emotional development of children. More educated mothers may have increased access to resources and knowledge about children's needs and development, and they are better informed about mental health, allowing them to seek professional help when they need it (Augustine & Crosnoe, 2010).

Living with a partner decreased the effects of maternal depression on the children's socio-emotional development. In the analyzed studies, data only captured whether the mothers were living with their partner, but did not include information on the level of support provided by them. However, the findings still align with previous research, emphasizing that partners can emotionally support mothers and practically assist them, creating a healthier mental environment for both mother and child (Martin & Brock, 2023).

Interestingly, primigravida status did not significantly moderate the association between maternal depression and children's socio-emotional development in the present meta-analysis, suggesting it did not explain the variability in effect sizes across studies and may not be a consistent vulnerability marker.

The type of socio-emotional development outcomes also influenced the strength of the association with maternal depression. The relationship was stronger when examining child problems, such as internalizing and externalizing behaviors, than child competencies like emotion regulation. This pattern suggests that maternal depressive symptoms are more consistently associated with elevations in children's socio-emotional difficulties than with reductions in competencies, implying a stronger link with problem-focused indicators than with positive socio-emotional skills during the first two years of life.

While this meta-analysis offers valuable insights, several limitations must be acknowledged. Many studies used mothers' self-reports to measure the children's socio-emotional development, which can introduce bias. Mothers with depressive symptoms, in particular, tend to evaluate things more negatively and passively, perhaps, than objectively.

Additionally, the heterogeneity of the studies in terms of design, measurement tools, and sample characteristics may limit the generalizability of the findings. Although we accounted for some variability through moderator analyses, future research should continue to explore other contextual factors, including social support and cultural and socioeconomic influences.

Looking forward, longitudinal studies monitoring children from the prenatal period through early childhood are needed to fully understand the temporal dynamics of the relationship between maternal depression and child socio-emotional development. Furthermore, intervention studies should address maternal mental health issues in more vulnerable populations, such as younger, less-educated, or first-time mothers. Another important factor to explore is the role of fathers and other caregivers in reducing the effects of maternal depression and supporting a healthy environment for the children.

Conclusion

This meta-analysis highlights the significant impact of maternal depression on the socio-emotional development of children during their first two years of life. The findings underscore the need for early identification and intervention for mothers experiencing depression, as well as monitoring young children whose mothers have experienced perinatal mental health difficulties. By addressing maternal mental health problems as early as possible, we can support children in developing in the most balanced and healthy way from the very beginning.

Authors' Notes

Conflict of Interest: All authors have no conflicts of interest to disclose.

Authors' Contributions:

Mădălina-Ruxandra Costin: Conceptualization; Methodology; Formal analysis; Investigation; Resources; Writing – original draft; Writing – review & editing.
Diana Tăut: Data curation; Writing – review & editing.
Adriana Smaranda Băban: Writing – review & editing.
*Sebastian Pinte*a: Formal analysis.

References

- Almeida, C. P. de, Sá, E., Cunha, F. F., & Pires, E. P. (2012). Common mental disorders during pregnancy and baby's development in the first year of life. *Journal of Reproductive and Infant Psychology*, 30(4), 341–351. <https://doi.org/10.1080/02646838.2012.736690>
- Anderson, V., Darling, S., Hearps, S., Darby, D., Dooley, J., McDonald, S., Turkstra, L., Brown, A., Greenham, M., Crossley, L., Charalambous, G., & Beauchamp, M. H. (2023). Deep phenotyping of socio-emotional skills in children with typical development, neurodevelopmental disorders, and mental health conditions: Evidence from the PEERS. *PloS One*, 18(10), e0291929. <https://doi.org/10.1371/journal.pone.0291929>
- Arikan, G., & Kumru, A. (2021). Patterns of Associations Between Maternal Symptoms and Child Problem Behaviors: The Mediating Role of Mentalization, Negative Intentionality, and Unsupportive Emotion Socialization. *Child Psychiatry & Human Development*, 52(4), 640–653. <https://doi.org/10.1007/s10578-020-01046-w>
- Atashbahar, O., Sari, A. A., Takian, A., Olyaeemanesh, A., Mohamadi, E., & Barakati, S. H. (2022). The impact of social determinants of health on early childhood development: A qualitative context analysis in Iran. *BMC Public Health*, 22(1), 1149. <https://doi.org/10.1186/s12889-022-13571-5>
- Atkinson, L., & Teychenne, M. (2019). Psychological, Social and Behaviour Changes During Pregnancy: Implications for Physical Activity and Exercise. In R. Santos-Rocha (Ed.), *Exercise and Sporting Activity During Pregnancy: Evidence-Based Guidelines* (pp. 19–43). Springer International Publishing. https://doi.org/10.1007/978-3-319-91032-1_2
- Augustine, J. M., & Crosnoe, R. (2010). Mothers' Depression and Educational Attainment and their Children's Academic Trajectories. *Journal of Health and Social Behavior*, 51(3), 274–290. <https://doi.org/10.1177/0022146510377757>
- Balaji, A. B., Claussen, A. H., Smith, D. C., Visser, S. N., Morales, M. J., & Perou, R. (2007). Social Support Networks and Maternal Mental Health and Well-Being. *Journal of Women's Health*, 16(10), 1386–1396. <https://doi.org/10.1089/jwh.2007.CDC10>
- Barnes, J., & Theule, J. (2019). Maternal depression and infant attachment security: A meta-analysis. *Infant Mental Health Journal*, 40(6), 817–834. <https://doi.org/10.1002/imhj.21812>

- Bates, R. A., Salsberry, P. J., Justice, L. M., Dynia, J. M., Logan, J. A. R., Gugi, M. R., & Purtell, K. M. (2020). Relations of Maternal Depression and Parenting Self-Efficacy to the Self-Regulation of Infants in Low-Income Homes. *Journal of Child and Family Studies*, 29(8), 2330–2341. <https://doi.org/10.1007/s10826-020-01763-9>
- Begg, C. B., & Mazumdar, M. (1994). Operating Characteristics of a Rank Correlation Test for Publication Bias. *Biometrics*, 50(4), 1088–1101. <https://doi.org/10.2307/2533446>
- Behrendt, H. F., Scharke, W., Herpertz-Dahlmann, B., Konrad, K., & Firk, C. (2019). Like mother, like child? Maternal determinants of children's early social-emotional development. *Infant Mental Health Journal*, 40(2), 234–247. <https://doi.org/10.1002/imhj.21764>
- Beijers, R., Buitelaar, J. K., & de Weerth, C. (2014). Mechanisms underlying the effects of prenatal psychosocial stress on child outcomes: Beyond the HPA axis. *European Child & Adolescent Psychiatry*, 23(10), 943–956. <https://doi.org/10.1007/s00787-014-0566-3>
- Bianciardi, E., Ongaretti, F., De Stefano, A., Siracusano, A., & Niu, C. (2023). The Mother-Baby Bond: Role of Past and Current Relationships. *Children (Basel, Switzerland)*, 10(3), 421. <https://doi.org/10.3390/children10030421>
- Bind, R. H., Sawyer, K., & Pariante, C. (2022). Depression in Pregnancy: Biological, Clinical, and Psychosocial Effects. In M. Percudani, A. Bramante, V. Brenna, & C. Pariante (Eds.), *Key Topics in Perinatal Mental Health* (pp. 3–21). Springer International Publishing. https://doi.org/10.1007/978-3-030-91832-3_1
- Black, M. M., Baqui, A. H., Zaman, K., McNary, S. W., Le, K., Arifeen, S. E., Hamadani, J. D., Parveen, M., Yunus, Md., & Black, R. E. (2007). Depressive symptoms among rural Bangladeshi mothers: Implications for infant development. *Journal of Child Psychology and Psychiatry*, 48(8), 764–772. <https://doi.org/10.1111/j.1469-7610.2007.01752.x>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to Meta-Analysis*. John Wiley & Sons.
- Brummelte, S., Mc Glinchey, E., Bonnin, A., & Oberlander, T. F. (2017). Developmental changes in serotonin signaling: Implications for early brain function, behavior and adaptation. *Neuroscience, Early Adversity and Brain Development*, 342, 212–231. <https://doi.org/10.1016/j.neuroscience.2016.02.037>
- Bronfenbrenner, U., & Morris, P. A. (2007). The Bioecological Model of Human Development. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology* (1st ed.). Wiley. <https://doi.org/10.1002/9780470147658.chpsy0114>
- Chorbadjian, T. N., Deavenport-Saman, A., Higgins, C., Chao, S. M., Yang, J. H., Koolwijk, I., & Vanderbilt, D. L. (2020). Maternal Depressive Symptoms and Developmental Delay at Age 2: A Diverse Population-Based Longitudinal Study. *Maternal and Child Health Journal*, 24(10), 1267–1277. <https://doi.org/10.1007/s10995-020-02990-8>
- Colomeischi, A. A., Ursu, A., Bogdan, I., Ionescu-Corbu, A., Bondor, R., & Conte, E. (2022). Social and Emotional Learning and Internalizing Problems among

- Adolescents: The Mediating Role of Resilience. *Children*, 9(9), Article 9. <https://doi.org/10.3390/children9091326>
- Cui, J., Mistur, E. J., Wei, C., Lansford, J. E., Putnick, D. L., & Bornstein, M. H. (2018). Multilevel factors affecting early socioemotional development in humans. *Behavioral Ecology and Sociobiology*, 72(10), 172. <https://doi.org/10.1007/s00265-018-2580-9>
- Cuijpers, P., Li, J., Hofmann, S. G., & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: A meta-analysis. *Clinical Psychology Review*, 30(6), 768–778. <https://doi.org/10.1016/j.cpr.2010.06.001>
- Denham, S., & Zinsser, K. (2014). *Social and Emotional Learning During Early Childhood* (pp. 926–935). https://doi.org/10.1007/978-1-4614-5999-6_144
- DiPietro, J. A., Novak, M. F. S. X., Costigan, K. A., Atella, L. D., & Reusing, S. P. (2006). Maternal Psychological Distress During Pregnancy in Relation to Child Development at Age Two. *Child Development*, 77(3), 573–587. <https://doi.org/10.1111/j.1467-8624.2006.00891.x>
- Duguay, G., Garon-Bissonnette, J., Lemieux, R., Dubois-Comtois, K., Mayrand, K., & Berthelot, N. (2022). Socioemotional development in infants of pregnant women during the COVID-19 pandemic: The role of prenatal and postnatal maternal distress. *Child and Adolescent Psychiatry and Mental Health*, 16(1), 28. <https://doi.org/10.1186/s13034-022-00458-x>
- Edwards, R. C., & Hans, S. L. (2016). Prenatal Depressive Symptoms and Toddler Behavior Problems: The Role of Maternal Sensitivity and Child Sex. *Child Psychiatry & Human Development*, 47(5), 696–707. <https://doi.org/10.1007/s10578-015-0603-6>
- Field, T., Hernandez-Reif, M., & Diego, M. (2006). Intrusive and withdrawn depressed mothers and their infants. *Developmental Review - DEVELOP REV*, 26, 15–30. <https://doi.org/10.1016/j.dr.2005.04.001>
- Fredriksen, E., von Soest, T., Smith, L., & Moe, V. (2019). Parenting Stress Plays a Mediating Role in the Prediction of Early Child Development from Both Parents' Perinatal Depressive Symptoms. *Journal of Abnormal Child Psychology*, 47(1), 149–164. <https://doi.org/10.1007/s10802-018-0428-4>
- Frosch, C. A., Schoppe-Sullivan, S. J., & O'Banion, D. D. (2021). Parenting and Child Development: A Relational Health Perspective. *American Journal of Lifestyle Medicine*, 15(1), 45–59. <https://doi.org/10.1177/1559827619849028>
- Gao, W., Paterson, J., Abbott, M., Carter, S., & Iusitini, L. (2007). Maternal Mental Health and Child Behaviour Problems at 2 Years: Findings from the Pacific Islands Families Study. *Australian & New Zealand Journal of Psychiatry*, 41(11), 885–895. <https://doi.org/10.1080/00048670701634929>
- Ghimire, U., Papabathini, S. S., Kawuki, J., Obore, N., & Musa, T. H. (2021). Depression during pregnancy and the risk of low birth weight, preterm birth and intrauterine growth restriction- an updated meta-analysis. *Early Human Development*, 152, 105243. <https://doi.org/10.1016/j.earlhumdev.2020.105243>

- Glynn, L. M., & Baram, T. Z. (2019). The Influence of Unpredictable, Fragmented Parental Signals on the Developing Brain. *Frontiers in Neuroendocrinology*, 53, 100736. <https://doi.org/10.1016/j.yfrne.2019.01.002>
- Goeglein, S. K., & Yatchmink, Y. E. (2020). Maternal Depression Is a Public Health Crisis: The Time to Act Is Now. *Pediatrics*, 146(3), e2020010413. <https://doi.org/10.1542/peds.2020-010413>
- Goldfeld, S., Woolcock, G., Katz, I., Tanton, R., Brinkman, S., O'Connor, E., Mathews, T., & Giles-Corti, B. (2015). Neighbourhood Effects Influencing Early Childhood Development: Conceptual Model and Trial Measurement Methodologies from the Kids in Communities Study. *Social Indicators Research*, 120(1), 197–212. <https://doi.org/10.1007/s11205-014-0578-x>
- Goodman, S. H., Rouse, M. H., Connell, A. M., Broth, M. R., Hall, C. M., & Heyward, D. (2011). Maternal depression and child psychopathology: A meta-analytic review. *Clinical Child and Family Psychology Review*, 14(1), 1–27. <https://doi.org/10.1007/s10567-010-0080-1>
- Guyon-Harris, K., Huth-Bocks, A., Lauterbach, D., & Janisse, H. (2016). Trajectories of maternal depressive symptoms across the birth of a child: Associations with toddler emotional development. *Archives of Women's Mental Health*, 19(1), 153–165. <https://doi.org/10.1007/s00737-015-0546-8>
- Hails, K. A., Reuben, J. D., Shaw, D. S., Dishion, T. J., & Wilson, M. N. (2018). Transactional Associations Among Maternal Depression, Parent–Child Coercion, and Child Conduct Problems During Early Childhood. *Journal of Clinical Child & Adolescent Psychology*, 47(sup1), S291–S305. <https://doi.org/10.1080/15374416.2017.1280803>
- Halle, T. G., & Darling-Churchill, K. E. (2016). Review of measures of social and emotional development. *Journal of Applied Developmental Psychology*, *Measuring Social and Emotional Development in Early Childhood*, 45, 8–18. <https://doi.org/10.1016/j.appdev.2016.02.003>
- Harris, R. A., & Santos, H. P., Jr. (2020). Maternal depression in Latinas and child socioemotional development: A systematic review. *PLOS ONE*, 15(3), e0230256. <https://doi.org/10.1371/journal.pone.0230256>
- Hildingsson, I., & Rubertsson, C. (2022). Postpartum bonding and association with depressive symptoms and prenatal attachment in women with fear of birth. *BMC Pregnancy and Childbirth*, 22(1), 66. <https://doi.org/10.1186/s12884-021-04367-3>
- Hummel, A. C., & Kiel, E. J. (2015). Maternal Depressive Symptoms, Maternal Behavior, and Toddler Internalizing Outcomes: A Moderated Mediation Model. *Child Psychiatry & Human Development*, 46(1), 21–33. <https://doi.org/10.1007/s10578-014-0448-4>
- Ilyka, D., Johnson, M. H., & Lloyd-Fox, S. (2021). Infant social interactions and brain development: A systematic review. *Neuroscience and Biobehavioral Reviews*, 130, 448–469. <https://doi.org/10.1016/j.neubiorev.2021.09.001>
- Izett, E., Rooney, R., Prescott, S. L., De Palma, M., & McDevitt, M. (2021). Prevention of Mental Health Difficulties for Children Aged 0–3 Years: A

- Review. *Frontiers in Psychology*, 11, 500361. <https://doi.org/10.3389/fpsyg.2020.500361>
- Jaramillo, J. M., Rendón, M. I., Muñoz, L., Weis, M., & Trommsdorff, G. (2017). Children's Self-Regulation in Cultural Contexts: The Role of Parental Socialization Theories, Goals, and Practices. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00923>
- Jones, D. E., Greenberg, M., & Crowley, M. (2015). Early social-emotional functioning and public health: The relationship between kindergarten social competence and future wellness. *American Journal of Public Health*, 105(11), 2283–2290. <https://doi.org/10.2105/AJPH.2015.302630>
- Junge, C., Garthus-Niegel, S., Slinning, K., Polte, C., Simonsen, T. B., & Eberhard-Gran, M. (2017). The Impact of Perinatal Depression on Children's Social-Emotional Development: A Longitudinal Study. *Maternal and Child Health Journal*, 21(3), 607–615. <https://doi.org/10.1007/s10995-016-2146-2>
- Kamal, D., & Kalita, K. (2010). Developmental Profile of Infants Born to Mothers with Postpartum Depression and Anxiety: A Comparative Study. *J. Indian Assoc. Child Adolesc. Ment. Health*, 6, 3–12. <https://doi.org/10.1177/0973134220100102>
- Kirby, J. N. (2020). Nurturing Family Environments for Children: Compassion-Focused Parenting as a Form of Parenting Intervention. *Education Sciences*, 10(1), Article 1. <https://doi.org/10.3390/educsci10010003>
- Komro, K. A., Flay, B. R., & Biglan, A. (2011). Creating Nurturing Environments: A Science-Based Framework for Promoting Child Health and Development within High-Poverty Neighborhoods. *Clinical Child and Family Psychology Review*, 14(2), 111–134. <https://doi.org/10.1007/s10567-011-0095-2>
- Koutra, K., Chatzi, L., Bagkeris, M., Vassilaki, M., Bitsios, P., & Kogevinas, M. (2013). Antenatal and postnatal maternal mental health as determinants of infant neurodevelopment at 18 months of age in a mother–child cohort (Rhea Study) in Crete, Greece. *Social Psychiatry and Psychiatric Epidemiology*, 48(8), 1335–1345. <https://doi.org/10.1007/s00127-012-0636-0>
- Letourneau, N., Leung, B., Ntanda, H., Dewey, D., Deane, A. J., Giesbrecht, G. F., & APrON Team. (2019). Maternal and paternal perinatal depressive symptoms associate with 2- and 3-year-old children's behaviour: Findings from the APrON longitudinal study. *BMC Pediatrics*, 19(1), 435. <https://doi.org/10.1186/s12887-019-1775-1>
- Letourneau, N., Salmani, M., & Duffett-Leger, L. (2010). Maternal Depressive Symptoms and Parenting of Children From Birth to 12 Years. *Western Journal of Nursing Research*, 32, 662–685. <https://doi.org/10.1177/0193945909359409>
- Li, J., Ma, H., Yang, H., Yu, H., & Zhang, N. (2023). Cognitive bias modification for adult's depression: A systematic review and meta-analysis. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.968638>
- Likhar, A., Baghel, P., & Patil, M. (2022). Early Childhood Development and Social Determinants. *Cureus*, 14(9), e29500. <https://doi.org/10.7759/cureus.29500>

- Malik, F., & Marwaha, R. (2024). Developmental Stages of Social Emotional Development in Children. In *StatPearls*. StatPearls Publishing. <http://www.ncbi.nlm.nih.gov/books/NBK534819/>
- Martin, R. C. B., & Brock, R. L. (2023). The importance of high-quality partner support for reducing stress during pregnancy and postpartum bonding impairments. *Archives of Women's Mental Health*, 26(2), 201–209. <https://doi.org/10.1007/s00737-023-01299-z>
- Martucci, M., Fava, G., Giacchetti, N., Aceti, F., Galeoto, G., Panfili, M., & Sogos, C. (2021). Perinatal depression as a risk factor for child developmental disorders: A cross-sectional study. *Rivista di Psichiatria*, 56(6), 321–327.
- Mason, Z. S., Briggs, R. D., & Silver, E. J. (2011). Maternal attachment feelings mediate between maternal reports of depression, infant social-emotional development, and parenting stress. *Journal of Reproductive and Infant Psychology*, 29(4), 382–394. <https://doi.org/10.1080/02646838.2011.629994>
- McAndrew, A. J. (2019). Maternal perinatal depression: Emotional development in offspring from infancy to adolescence. *Early Child Development and Care*, 189(1), 168–177. <https://doi.org/10.1080/03004430.2017.1292398>
- McDonald, S. W., Kehler, H. L., & Tough, S. C. (2018). Risk factors for delayed social-emotional development and behavior problems at age two: Results from the All Our Babies/Families (AOB/F) cohort. *Health Science Reports*, 1(10), e82. <https://doi.org/10.1002/hsr2.82>
- Milgrom, J., Hirshler, Y., Holt, C., Skouteris, H., Galbally, M., East, C., Glover, V., Reece, J., O'Donnell, K. J., Walker, S. P., Malloy, S., & Gemmill, A. W. (2023). Early intervention to prevent adverse child emotional and behavioural development following maternal depression in pregnancy: Study protocol for a randomised controlled trial. *BMC Psychology*, 11(1), 222. <https://doi.org/10.1186/s40359-023-01244-w>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Morales, M. F., Girard, L.-C., Raouna, A., & MacBeth, A. (2023). The association of different presentations of maternal depression with children's socio-emotional development: A systematic review. *PLOS Global Public Health*, 3(2), e0001649. <https://doi.org/10.1371/journal.pgph.0001649>
- Nix, L., Nixon, E., Quigley, J., & O'Keane, V. (2021). Perinatal depression and children's developmental outcomes at 2 years postpartum. *Early Human Development*, 156, 105346. <https://doi.org/10.1016/j.earlhumdev.2021.105346>
- O'Connor, E., Senger, C. A., Henninger, M., Gaynes, B. N., Coppola, E., & Weyrich, M. S. (2019). Accuracy of Screening Instruments to Predict Perinatal Depression. In *Interventions to Prevent Perinatal Depression: A Systematic Evidence Review for the U.S. Preventive Services Task Force* [Internet]. Agency for Healthcare Research and Quality (US). <https://www.ncbi.nlm.nih.gov/books/NBK537807/>

- Oberle, E., & Schonert-Reichl, K. A. (2017). Social and emotional learning: Recent research and practical strategies for promoting children's social and emotional competence in schools. In *Handbook of social behavior and skills in children* (pp. 175–197). Springer International Publishing/Springer Nature. https://doi.org/10.1007/978-3-319-64592-6_11
- Okitundu-Luwa, D., Imbula Essam, B., Sombo Ayanne, M.-T., Ndjukendi Omba, A., Otete Djamba, F., Kayembe Kalula, T., Kamanga Mbuyi, T., Guedeney, A., & Kashala-Abotnes, E. (2021). Sustained social withdrawal behavior and 'difficult' temperament among infants, in relation to maternal affectivity in a low-income area of Kinshasa, Democratic Republic of Congo. *European Child & Adolescent Psychiatry*. <https://doi.org/10.1007/s00787-021-01873-9>
- Pacheco, A., & Figueiredo, B. (2012). Mother's depression at childbirth does not contribute to the effects of antenatal depression on neonate's behavioral development. *Infant Behavior and Development*, 35(3), 513–522. <https://doi.org/10.1016/j.infbeh.2012.02.001>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, 105906. <https://doi.org/10.1016/j.ijssu.2021.105906>
- Paley, B., & Hajal, N. J. (2022). Conceptualizing Emotion Regulation and Coregulation as Family-Level Phenomena. *Clinical Child and Family Psychology Review*, 25(1), 19–43. <https://doi.org/10.1007/s10567-022-00378-4>
- Palmer, F. B., Anand, K. J. S., Graff, J. C., Murphy, L. E., Qu, Y., Völgyi, E., Rovnaghi, C. R., Moore, A., Tran, Q. T., & Tylavsky, F. A. (2013). Early Adversity, Socioemotional Development, and Stress in Urban 1-Year-Old Children. *The Journal of Pediatrics*, 163(6), 1733–1739.e1. <https://doi.org/10.1016/j.jpeds.2013.08.028>
- Pearson, R. M., Melotti, R., Heron, J., Joinson, C., Stein, A., Ramchandani, P. G., & Evans, J. (2012). Disruption to the development of maternal responsiveness? The impact of prenatal depression on mother–infant interactions. *Infant Behavior and Development*, 35(4), 613–626. <https://doi.org/10.1016/j.infbeh.2012.07.020>
- Quansah, E., Ohene, L. A., Norman, L., Mireku, M. O., & Karikari, T. K. (2016). Social Factors Influencing Child Health in Ghana. *PLOS ONE*, 11(1), e0145401. <https://doi.org/10.1371/journal.pone.0145401>
- Rogers, A., Obst, S., Teague, S. J., Rossen, L., Spry, E. A., Macdonald, J. A., Sunderland, M., Olsson, C. A., Youssef, G., & Hutchinson, D. (2020). Association Between Maternal Perinatal Depression and Anxiety and Child and Adolescent Development: A Meta-analysis. *JAMA Pediatrics*, 174(11), 1082–1092. <https://doi.org/10.1001/jamapediatrics.2020.2910>

- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>
- Saharoy, R., Potdukhe, A., Wanjari, M., & Taksande, A. B. (2023). Postpartum Depression and Maternal Care: Exploring the Complex Effects on Mothers and Infants. *Cureus*, 15(7), e41381. <https://doi.org/10.7759/cureus.41381>
- Schaffir, J. (2016). Biological Changes During Pregnancy and the Postpartum Period. In A. Wenzel (Ed.), *The Oxford Handbook of Perinatal Psychology* (p. 0). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199778072.013.23>
- Schiavo, R. de A., & Perosa, G. B. (2020). Child Development, Maternal Depression and Associated Factors: A Longitudinal Study. *Paidéia (Ribeirão Preto)*, 30, e3012. <https://doi.org/10.1590/1982-4327e3012>
- Sidor, A., Kunz, E., Schweyer, D., Eickhorst, A., & Cierpka, M. (2011). Links between maternal postpartum depressive symptoms, maternal distress, infant gender and sensitivity in a high-risk population. *Child and Adolescent Psychiatry and Mental Health*, 5(1), 7. <https://doi.org/10.1186/1753-2000-5-7>
- Smith-Nielsen, J., Lange, T., Wendelboe, K. I., Wower, R. K., & Væver, M. S. (2019). Associations Between Maternal Postpartum Depression, Infant Social Behavior With a Stranger, and Infant Cognitive Development. *Infancy*, 24(4), 663–670. <https://doi.org/10.1111/infa.12287>
- Smorti, M., Milone, A., Fanciullacci, L., Ciaravolo, A., & Berrocal, C. (2024). Parenting and Emotional and Behavioral Difficulties in a General Population Sample of Adolescents: The Mediating Role of Emotional Dysregulation. *Children*, 11(4), Article 4. <https://doi.org/10.3390/children11040435>
- Speidel, R., Wong, T. K. Y., Al-Janaideh, R., Colasante, T., & Malti, T. (2023). Nurturing child social-emotional development: Evaluation of a pre-post and 2-month follow-up uncontrolled pilot training for caregivers and educators. *Pilot and Feasibility Studies*, 9(1), 148. <https://doi.org/10.1186/s40814-023-01357-4>
- Takács, L., Kandrnal, V., Kaňková, Š., Bartoš, F., & Mudrák, J. (2020). The effects of pre- and post-partum depression on child behavior and psychological development from birth to pre-school age: A protocol for a systematic review and meta-analysis. *Systematic Reviews*, 9, 146. <https://doi.org/10.1186/s13643-019-1267-2>
- Tervahartiala, K., Nolvi, S., Kataja, E.-L., Seppälä, M., Autere, T.-A., Hakanen, H., Karlsson, H., Carter, A., Karlsson, L., & Korja, R. (2024). Childcare context and socio-emotional development in toddlers – a quantitative report from the FinnBrain Birth Cohort Study, Finland. *European Early Childhood Education Research Journal*, 32(1), 71–85. <https://doi.org/10.1080/1350293X.2023.2234110>
- The APRON Team, Letourneau, N., Leung, B., Ntanda, H., Dewey, D., Deane, A. J., & Giesbrecht, G. F. (2019). Maternal and paternal perinatal depressive symptoms associate with 2- and 3-year-old children's behaviour: Findings from the APRON longitudinal study. *BMC Pediatrics*, 19(1), 435. <https://doi.org/10.1186/s12887-019-1775-1>

- Thijssen, S., Riem, M. M. E., Vijayakumar, N., Cima, M. J., & Whittle, S. (2023). Editorial: Nurturing the brain: Associations between family environment and child brain development. *Frontiers in Neuroscience*, 17. <https://doi.org/10.3389/fnins.2023.1119838>
- Thümmler, R., Engel, E.-M., & Bartz, J. (2022). Strengthening Emotional Development and Emotion Regulation in Childhood—As a Key Task in Early Childhood Education. *International Journal of Environmental Research and Public Health*, 19(7). <https://doi.org/10.3390/ijerph19073978>
- Traylor, C. S., Johnson, J. D., Kimmel, M. C., & Manuck, T. A. (2020). Effects of psychological stress on adverse pregnancy outcomes and nonpharmacologic approaches for reduction: An expert review. *American Journal of Obstetrics & Gynecology Mfm*, 2(4), 100229. <https://doi.org/10.1016/j.ajogmf.2020.100229>
- Trillingsgaard, T., & Sommer, D. (2018). Associations between older maternal age, use of sanctions, and children's socio-emotional development through 7, 11, and 15 years. *European Journal of Developmental Psychology*, 15(2), 141–155. <https://doi.org/10.1080/17405629.2016.1266248>
- Tronick, E., & Reck, C. (2009). Infants of depressed mothers. *Harvard Review of Psychiatry*, 17(2), 147–156. <https://doi.org/10.1080/10673220902899714>
- Urizar, G. G., & Muñoz, R. F. (2022). Role of Maternal Depression on Child Development: A Prospective Analysis from Pregnancy to Early Childhood. *Child Psychiatry and Human Development*, 53(3), 502–514. <https://doi.org/10.1007/s10578-021-01138-1>
- Wall-Wieler, E., Roos, L. L., & Gotlib, I. H. (2020). Maternal Depression in Early Childhood and Developmental Vulnerability at School Entry. *Pediatrics*, 146(3), e20200794. <https://doi.org/10.1542/peds.2020-0794>
- Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2013). The Newcastle–Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
- Wilson, N., & Lee, J. J. (2023). Psychological Changes During and After Pregnancy. In M. Gomes-Ferreira & J. Olivas-Menayo (Eds.), *Post-maternity Body Changes: Obstetric Fundamentals and Surgical Reshaping* (pp. 75–97). Springer International Publishing. https://doi.org/10.1007/978-3-030-43840-1_5
- Zięba-Kołodziej, B., & Oliinyk, M. (2023). Helping Children Overcome Socio-Emotional and Behavioral Problems Based on the Kids' Skills Method. *Revista Romaneasca Pentru Educatie Multidimensionala*, 15, 31–43. <https://doi.org/10.18662/rrem/15.3/751>

PERCEIVED PARENTAL BLAME AND EMOTION REGULATION AS PREDICTORS OF ADULT PSYCHOLOGICAL DISTRESS: DEVELOPMENT AND VALIDATION OF THE PERCEIVED PARENTAL BLAME SCALE

*Nesrullah Okan¹**

¹ Associate Professor Dr., Department of Educational Sciences, Guidance and Psychological Counseling, Firat University/ Elazığ/ Turkey

Abstract

The present study developed and validated the Perceived Parental Blame Scale (PPBS) and examined the underlying psychological mechanisms that connect perceived parental blame experienced during childhood to emotional distress in adulthood. The present study is anchored in attachment theory and emotion regulation frameworks. Its objective is to explore how early perceptions of parental blame may shape long-term emotional functioning and vulnerability to psychological difficulties. Specifically, a mediation model was tested in which difficulties in emotion regulation were hypothesized to explain the relationship between parental blame and adult symptoms of depression, anxiety, and stress. Two independent adult samples were recruited to ensure methodological robustness. The initial sample (N = 308) was utilised to conduct exploratory and confirmatory factor analyses, which substantiated a unidimensional and psychometrically sound factor structure of the PPBS with high internal consistency. The second sample (N = 576) was used for structural equation modelling, which confirmed that perceived parental blame significantly predicted elevated levels of psychological distress. Furthermore, it was demonstrated that this association was partially mediated by emotion regulation deficits. The findings emphasise the long-lasting impact of early family interactions on adult emotional well-being, suggesting that experiences of blame from caregivers may result in the internalisation of maladaptive emotional schemas that persist into adulthood. The PPBS thus offers a novel and reliable instrument for future research and clinical assessment of family-based emotional dynamics. The results of the study provide valuable

* Correspondence concerning this article should be addressed to Associate Professor Dr. Nesrullah Okan, Department of Educational Sciences, Guidance and Psychological Counseling, Firat University/ Elazığ/ Turkey
ORCID: <https://orcid.org/0000-0002-9496-6417>
E-mail: nesrokan@gmail.com

implications for prevention and intervention programs aiming to strengthen emotion regulation skills and address maladaptive parental patterns within family and systemic therapy contexts.

Keywords: perceived parental blame, emotion regulation, adult psychological distress, attachment theory, family dynamics, scale development

It is an established fact that early relational experiences have a significant impact on psychological well-being throughout the lifespan. Among these, perceived parental blame represents a specific dimension of negative parent–child interaction that has received limited empirical attention (Bowlby, 1988; Knappe et al., 2019). Perceived parental blame is defined as a child's experience of being unfairly or repeatedly held responsible by caregivers for mistakes or failures. This construct is conceptually related to parental criticism and expressed emotion, which have been extensively studied as risk factors for psychopathology (Hooley & Parker, 2006; Renshaw, 2008). However, extant measures such as the Parental Bonding Instrument (Parker, Tupling, & Brown, 1979), the Perceived Criticism Measure (Renshaw, 2008), or expressed emotion scales (Hooley & Parker, 2006) primarily capture general caregiving attitudes or critical communication styles rather than the child's retrospective perception of being personally blamed by parents.

Such experiences of perceived parental blame have been shown to disrupt the development of self-worth and emotional security, resulting in heightened vulnerability to stress, impaired socio-emotional development, and long-term psychological difficulties, including cognitive distortions, reduced self-efficacy, and difficulties in emotion regulation (Brown & Harris, 2020; Kline et al., 2022; Monroe & Simons, 1991; Morris et al., 2017). Despite these associations, there is currently no standardized tool specifically designed to measure perceived parental blame as a distinct relational adversity. The present study introduces the Perceived Parental Blame Scale (PPBS), a retrospective self-report instrument developed to capture emotionally salient parent–child dynamics from the adult's perspective.

In addition to validating the instrument, the study investigates the long-term psychological consequences of perceived parental blame, particularly its associations with depression, anxiety, and stress. The present study draws upon attachment theory and emotion regulation frameworks to hypothesise that emotion regulation, defined as the ability to modulate one's emotional responses in challenging interpersonal contexts (Gross & Jazaieri, 2014; Aldao et al., 2010), serves as a mediating factor in these relationships. The study's objective is twofold: firstly, to elucidate the role of emotion regulation, and secondly, to contribute to a

more nuanced understanding of how parental blame experiences shape adult psychological functioning.

Parental Blame and Psychological Health

Perceived parental blame refers to relational patterns in which children are persistently criticised, reminded of failures, or implicitly held responsible for familial problems (Knappe et al., 2019). This form of interpersonal dynamic has conceptual overlap with expressed emotion and perceived criticism, constructs which emphasise the impact of critical and blaming communication on children's psychological functioning (Hooley & Parker, 2006; Renshaw, 2008). From a developmental and attachment perspective, repeated experiences of parental blame have the potential to undermine children's self-concept and emotional security, contributing to vulnerabilities such as low self-esteem, reduced resilience, and difficulties in regulating emotions (Bowlby, 1988; Morris et al., 2017). Empirical studies suggest that individuals exposed to such relational patterns are more likely to develop insecure attachment orientations, particularly anxious and avoidant styles, which heighten risks for depression and anxiety across the lifespan (Mikulincer & Shaver, 2016; Dagan et al., 2018).

Furthermore, children who internalise parental blame may adopt maladaptive cognitive styles, including self-blame, helplessness, and negative automatic thoughts. These have been established as risk factors for affective disorders (Ingram et al., 2006; Chang et al., 2021). These cognitive-emotional vulnerabilities frequently persist into adulthood, manifesting as challenges in trust, boundary-setting, and self-compassion—domains closely tied to family functioning and relational well-being (Scharf & Goldner, 2018). Psychological distress in adulthood is commonly measured with instruments such as the Depression Anxiety Stress Scale-21 (DASS-21), which has been demonstrated to have strong psychometric validity in assessing symptomatology linked to adverse relational experiences (Lovibond & Lovibond, 1995; Henry & Crawford, 2005). It is evident from previous studies that early relational adversities, including perceived parental blame, are associated with elevated levels of distress and may also contribute to symptoms reminiscent of trauma. This underscores the importance of assessing this construct in both research and clinical practice (Steele et al., 2018; Infurna & Luthar, 2018).

Parental Blame and Emotion Regulation

Emotion regulation is defined as the individual's capacity to monitor, evaluate, and modify emotional responses in order to adapt to distressing or

challenging situations (Gross & John, 2003; Gratz & Roemer, 2004). Adaptive strategies, such as cognitive reappraisal, have been demonstrated to be consistently linked to psychological resilience. Conversely, maladaptive approaches, including avoidance or suppression, have been shown to be associated with increased vulnerability to emotional distress (Aldao et al., 2010; Joormann & Stanton, 2016; Schäfer et al., 2017). Within the family context, parental behaviours are of particular significance in the socialisation of children's regulatory abilities. Supportive parenting, characterised by validation and guidance, has been shown to foster the development of adaptive regulation strategies in children. Conversely, critical or blame-oriented interactions have been demonstrated to impede children's ability to manage emotions effectively (Morris, Silk, Steinberg, Myers, & Robinson, 2007; Luyten et al., 2020). It is hypothesised that, over time, repeated exposure to parental blame may reinforce maladaptive regulation patterns, leaving individuals more vulnerable to psychological distress in adulthood.

Utilising the theoretical framework outlined above, the present study hypothesises that emotion regulation serves as a mediator in the association between perceived parental blame and the subsequent development of psychological symptoms in adults. Specifically, individuals reporting higher levels of parental blame are expected to demonstrate greater emotional difficulties, which in turn contribute to elevated levels of depression, anxiety, and stress. The testing of this mediating mechanism provides a more precise understanding of how family dynamics translate into adult psychological functioning.

Hypotheses, Contribution and Importance of the Research

The present study investigates the long-term psychological implications of perceived parental blame in childhood, and whether emotion regulation mediates these associations. Although previous studies have documented the adverse effects of parental criticism and negative family interactions, few studies have explicitly conceptualised and measured perceived parental blame as a distinct construct (Hooley & Parker, 2006; Renshaw, 2008). Existing approaches have frequently relied on broad measures of expressed emotion or perceived criticism. While these measures are informative, they do not fully capture the specific relational pattern of parental blame. The present study makes an original contribution to the assessment of childhood relational experiences by developing and validating the Perceived Parental Blame Scale (PPBS). The following hypotheses underpin the study:

H1: Perceived parental blame will be positively associated with adult depression, anxiety, and stress.

H2: Emotion regulation will mediate the relationship between perceived parental blame and adult psychological distress.

H3: Higher levels of emotion regulation will be associated with lower levels of psychological distress, even in the presence of perceived parental blame.

The present research makes a contribution to the literature on intergenerational family dynamics, emotion regulation, and mental health by introducing a psychometrically validated tool and empirically supported model. The objective of the present study is twofold: firstly, to advance theoretical understanding of how specific parental communication patterns influence long-term psychological outcomes, and secondly, to provide a foundation for future applied research in family psychology and counselling.

Method

Research Design

This study was conducted using a two-stage quantitative research design to investigate the psychological consequences of perceived parental blame and the mediating role of emotion regulation. In the first stage, the Perceived Parental Blame Scale (PPBS) was developed and subjected to a series of psychometric evaluations. The process included item generation based on theoretical literature and expert consultation, followed by exploratory factor analysis (EFA) to identify the underlying factor structure and confirmatory factor analysis (CFA) to assess construct validity. Reliability was evaluated using internal consistency coefficients. In the second stage, the validated scale was employed in a broader sample to examine the associations among perceived parental blame, emotion regulation, and psychological health outcomes. Structural equation modeling (SEM) was used to test hypothesized relationships and the mediating role of emotion regulation in the link between parental blame and adult psychological distress (i.e., depression, anxiety, and stress).

The study adopted a cross-sectional and descriptive design, allowing for the identification of associations between variables at a single point in time. However, consistent with methodological guidelines, it should be noted that such a design does not permit causal inference regarding the directionality of the observed relationships (Bryman, 2016; Maxwell & Cole, 2007).

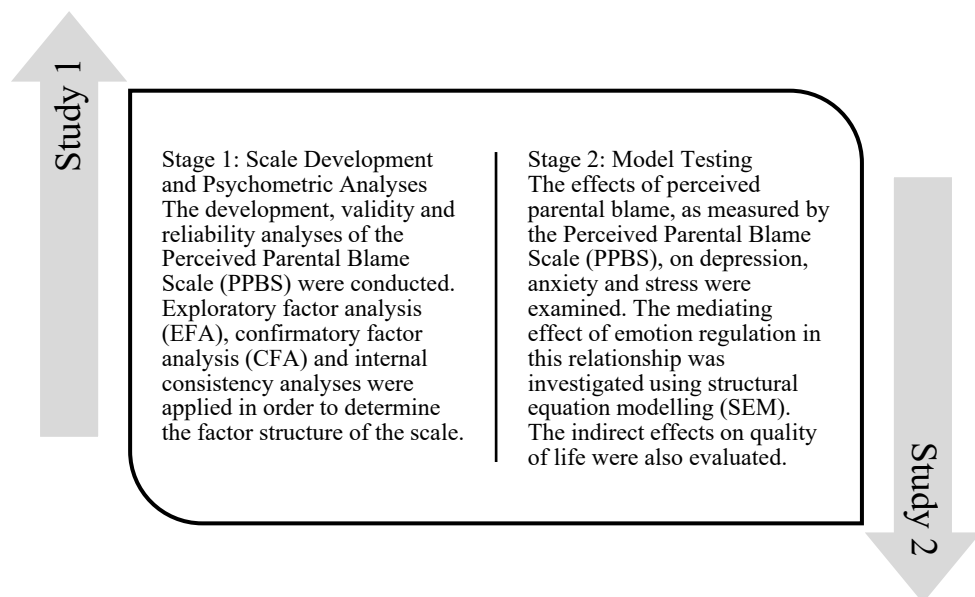


Figure 1. Research Design Stages

Figure 1 illustrates the two-stage design of the study. In Study 1, the PPBS was developed and validated using psychometric analyses. In Study 2, SEM was applied to test the mediating role of emotion regulation in the relationship between perceived parental blame and adult psychological outcomes.

Participants

The study was conducted in two phases.

Scale development phase: 308 participants (198 females, 110 males).

Model testing phase: 576 participants (272 females, 304 males).

All participants were **Turkish adults**, and those in the model testing phase were specifically **married individuals**. The rationale for selecting married adults was based on the assumption that childhood relational experiences, such as perceived parental blame, may resurface or influence functioning within marital relationships, making this group particularly relevant for the study of intergenerational family dynamics.

Participants were recruited through an online survey platform (Google Forms) disseminated via social media and professional counseling networks. Inclusion criteria were being 18 years or older and currently married. Exclusion criteria included being divorced, widowed, or under the age of 18.

The demographic characteristics of the model testing sample are presented in Table 1. As shown, gender distribution was relatively balanced (47.2% female,

52.8% male). The largest subgroup had been married for 11–15 years (26.9%). Participants were predominantly well educated, with 40.3% holding a postgraduate degree and 36.1% holding a bachelor's degree. Most participants identified themselves as belonging to the middle socioeconomic level (81.9%). While this educational profile strengthens the reliability of self-reported data, it also suggests that the findings may not be fully generalizable to populations with lower educational attainment.

Table 1. Demographic Characteristics of the Participants

Variable	Category	Frequency (n)	Percentage (%)
Gender	Female	272	47.2
	Male	304	52.8
Marital Duration	0-5 years	86	14.9
	6-10 years	142	24.7
	11-15 years	155	26.9
	16-25 years	113	19.6
	26 years and above	80	13.9
	Primary School	20	3.5
Educational Level	Middle School	12	2.1
	High School	56	9.7
	Associate Degree	48	8.3
	Bachelor's Degree	208	36.1
	Postgraduate	232	40.3
	Low	8	1.4
Perceived Socioeconomic Status	Middle	472	81.9
	High	96	16.7

Data Collection Process

The data were collected using an online survey method. A convenience sampling strategy was adopted, and participants were informed about the purpose and scope of the research prior to participation. The data collection process was executed through the utilisation of Google Forms, while the survey link was disseminated via social media platforms (e.g., Facebook, Instagram) and professional counselling networks, with the objective of reaching a diverse adult population. Prior to commencing the questionnaire, all participants provided written informed consent. At the conclusion of the survey, participants confirmed that the information they provided could be used anonymously and analysed solely for scientific purposes. No personal identifying data were collected, and all procedures strictly adhered to ethical guidelines for research involving human subjects (American Psychological Association, 2017).

Measurement Tools

The instruments utilised in this study underwent a rigorous evaluation process to ascertain their psychometric reliability and validity. Reliability analyses, incorporating Cronbach's alpha coefficients, were conducted for all scales. The newly developed measure was subjected to exploratory factor analysis (EFA) and confirmatory factor analysis (CFA).

Perceived Parental Blaming Scale (PPBS)

The Perceived Parental Blame Scale (PPBS) was developed in this study to assess adults' retrospective perceptions of parental blame experienced during childhood. The generation of items was informed by a thorough review of the extant literature on parental criticism and blame constructs (see, for example, Hooley & Parker, 2006; Renshaw, 2008), in addition to expert assessments provided by family psychology specialists. The preliminary pool comprised 22 items, which were reduced to 15 items following expert review and pilot testing. The following sample items are provided for your consideration: It has been asserted that “My parents often reminded me of my past mistakes” and that “I was blamed for family problems even when they were not my fault”. The items were evaluated using a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores denoting an increased perception of parental blame. The EFA supported a unidimensional structure with strong exploratory power (KMO = 0.94; Bartlett's test $\chi^2 = 9098.604$, $df = 105$, $p < .001$; total explained variance = 67.60%). The factor loadings ranged from 0.646 to 0.912, indicating the satisfactory performance of the items. CFA confirmed the construct validity of the PPBS, with good model fit indices (GFI = 0.94; CFI = 0.96; TLI = 0.95; RMSEA = 0.06; SRMR = 0.05). These indices were consistent with recommended thresholds (Hu & Bentler, 1999). The internal consistency of the final 15-item scale was found to be excellent ($\alpha = 0.965$).

Emotional Regulation Scale

Emotion regulation was measured using the Turkish adaptation (Yiğit & Güzey, 2019) of the Emotion Regulation Questionnaire (Gross & John, 2003). The ERQ comprises 10 items that assess two emotion regulation strategies: cognitive reappraisal (six items; e.g., “When I want to feel less negative emotion, I change the way I think about the situation”) and expressive suppression (four items; e.g., “I keep my emotions to myself”). The items are evaluated using a 7-point Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). In the present study, the Turkish version demonstrated excellent reliability (Cronbach's $\alpha = 0.91$). Scoring and composite construction. The ERQ indexes current/typical emotion-regulation strategies rather than retrospective use (Gross & John, 2003). For the present

analyses, we created an Emotion-Regulation Difficulties (ERD) composite so that higher scores reflect greater difficulty. Specifically, cognitive reappraisal items were reverse-scored, and the two ERQ subscale scores (reappraisal-reversed and suppression) were z-standardized and averaged to yield ERD (higher = more difficulty). The composite demonstrated excellent internal consistency in the current sample ($\alpha = .91$). This scoring decision aligns the direction of the ER variable with distress indicators and facilitates interpretation in the mediation model.

Stress, Depression and Anxiety Scales (DASS-21)

The psychological distress of the participants was assessed using the Turkish adaptation (Sarıçam, 2018) of the Depression Anxiety Stress Scale–21 (DASS-21; Lovibond & Lovibond, 1995). The scale comprises 21 items, which are distributed across three subscales: The presence of depression was indicated by statements such as “I felt that life was meaningless”. Anxiety was indicated by statements such as “I experienced trembling in my hands”. Stress was indicated by statements such as “I found it hard to wind down”. Responses are given on a 4-point Likert scale ranging from 0 (did not apply to me at all) to 3 (applied to me very much or most of the time). In this study, the internal consistency coefficients were found to be high. The alpha coefficients for stress, depression and anxiety were found to be 0.89, 0.91 and 0.87 respectively.

Data Analysis

The analysis was conducted in several sequential stages to ensure methodological rigor:

The following statistics are descriptive in nature. Descriptive analyses were performed to summarise the demographic characteristics of participants and to examine the means, standard deviations, skewness, and kurtosis values for all items. The assumption of normality was evaluated using skewness and kurtosis values, with thresholds suggested by Kline (2015).

Scale Development Analyses. In the initial phase, exploratory factor analysis (EFA) was utilised to investigate the underlying factor structure of the Perceived Parental Blame Scale (PPBS). The suitability of the data for factor analysis was evaluated using the Kaiser–Meyer–Olkin (KMO) statistic and Bartlett's test of sphericity. In the second phase, confirmatory factor analysis (CFA) was applied in order to validate the factor structure. The model's fit was evaluated through multiple fit indices, including the chi-squared divided by degrees of freedom (χ^2/df), the goodness of fit index (GFI), the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR) (Hu & Bentler, 1999). The internal consistency reliability of the scale was assessed using Cronbach's α and composite reliability values.

Structural Equation Modeling (SEM). The present study employed Structural Equation Modelling (SEM) to examine the hypothesised relationships between perceived parental blame, emotion regulation, and psychological distress (i.e. depression, anxiety, and stress). The mediation effect of emotion regulation was tested within the Structural Equation Modelling (SEM) framework. The evaluation of the model followed conventional criteria for fit indices, and standardized path coefficients were reported in order to assess the strength of the associations (Kline, 2015). Bootstrapping Procedures. In order to test the indirect effects and mediation hypotheses, non-parametric bootstrapping with 5,000 resamples was employed (Preacher & Hayes, 2008). Bias-corrected confidence intervals were calculated, and mediation was considered to be significant when the 95% confidence interval did not include zero. All statistical analyses were conducted utilising the IBM SPSS 26.0 and AMOS 24.0 software packages. The level of statistical significance was set at $p < .05$.

Results

This section reports the findings from the scale development process and the structural equation modeling (SEM) analyses. Results are organized as follows: (a) scale development and psychometrics (EFA/CFA), (b) descriptive statistics and correlations, and (c) SEM models testing the mediation of emotion regulation.

Scale Development Findings

In the scale development process, to test the suitability of the factor structure Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were conducted.

Exploratory Factor Analysis (EFA) Results

Table 2 presents the comprehensive results of the exploratory factor analysis.

Table 2. Exploratory Factor Analysis Results

Type of Analysis	Value
Kaiser-Meyer-Olkin (KMO) Test	0.940
Bartlett's Test (Chi-Square)	9098.604
Bartlett's Test (df)	105
Bartlett's Test (Sig.)	0.000
Total Explained Variance (%)	67.60%
Cronbach's Alpha (Reliability Coefficient)	0.965
Number of Items in the Scale	15

Table 2 presents the results of the exploratory factor analysis (EFA) conducted to evaluate the construct validity of the Perceived Parental Blame Scale. An exploratory factor analysis (EFA) supported a unidimensional solution for the Perceived Parental Blame Scale (PPBS). Sampling adequacy was excellent (KMO = .94), and Bartlett's test of sphericity was significant ($\chi^2 = 9098.604$, $df = 105$, $p < .001$), indicating suitability for factor analysis (Kaiser, 1974; Hair et al., 2019). The single-factor solution explained 67.60% of the total variance, exceeding the conventional 60% benchmark (Field, 2013). Internal consistency was excellent (Cronbach's $\alpha = .965$; Nunnally & Bernstein, 1994). Item loadings are reported in Table 3.

Factor Loadings and Reliability Findings

Table 3 shows the loading values and Cronbach's alpha internal consistency coefficients of the items in the scale.

Table 3. Component Matrix and Reliability Statistics

Item	Item wording	Component
es10	My parents would blame me rather than offer solutions when I had a problem.	0.912
es15	When I experienced any failure, my parents placed all the responsibility entirely on me instead of supporting me.	0.876
es7	My parents would criticize me as a generally unsuccessful person based on a single mistake.	0.863
es12	I felt that, during my childhood, they typically held me responsible for problems that occurred.	0.857
es9	In difficult times, my parents preferred to criticize me rather than support me.	0.855
es14	My parents' accusations seriously damaged my self-confidence and self-esteem.	0.850
es11	My parents' constant blaming deeply affected my self-confidence.	0.835
es3	When a problem occurred, my parents would usually attribute all the fault to me.	0.830
es2	My parents continually placed the responsibility for the difficulties I experienced on me.	0.823
es6	My parents would point out my mistakes in front of others and embarrass me.	0.807
es4	Expressions such as "You always do it wrong" were frequently used by my parents.	0.796
es13	Because of my parents' accusatory attitudes, there were times when I felt worthless.	0.787
es5	My parents often labeled me as inadequate and unsuccessful.	0.784
es1	When I made a mistake, my parents would directly criticize me instead of trying to understand me.	0.779
es8	When I did something wrong, my parents would constantly describe me as a bad person.	0.646
Cronbach's α (total scale)		0.965
Number of items		15

Table 3 reports the standardized loadings for the 15 PPBS items. Loadings ranged from .646 to .912, with the majority $\geq .70$, supporting adequate indicator reliability and a single-factor structure consistent with the EFA summary (Table 2) and the CFA fit (Table 4). Overall internal consistency was excellent ($\alpha = .965$). Item content clusters (e.g., blame for problems, global criticism, lack of support) indicate a coherent representation of the perceived parental blame construct without redundancy across stems.

Confirmatory Factor Analysis (CFA) Results

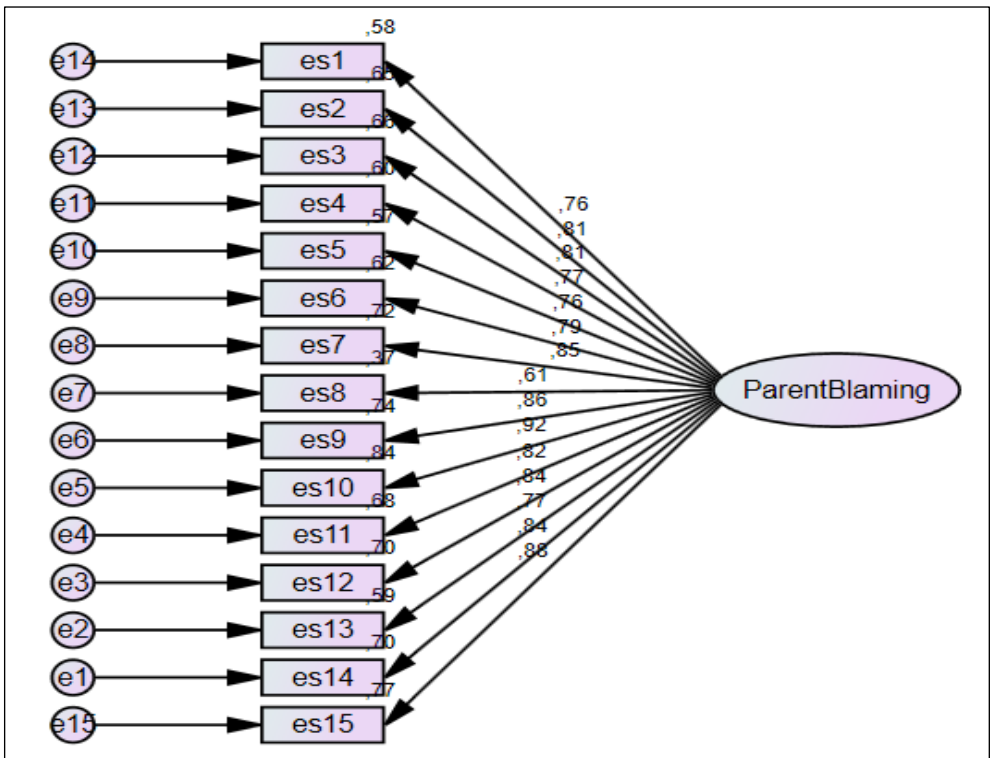


Figure 2. Confirmatory Factor Analysis Result of the Scale

Figure 2 displays the single-factor CFA model for the Perceived Parental Blame Scale (PPBS) with 15 indicators (es1–es15). Standardized factor loadings ranged from .58 to .98; the majority were $\geq .70$, indicating adequate indicator reliability. Two items (es1, es3) showed comparatively lower loadings ($\approx .58$ – $.63$) but were retained for content coverage within the construct and because the overall model fit was satisfactory.

Table 4. Goodness of Fit Values for the Model

Index of Fit	Criteria	Model Result
χ^2/sd	< 3	2.17
GFI (Goodness of Fit Index)	> 0.90	0.94
CFI (Comparative Fit Index)	> 0.95	0.96
TLI (Tucker-Lewis Index)	> 0.95	0.95
RMSEA (Root Mean Square Error of Approximation)	< 0.08	0.06
SRMR (Standardised Root Mean Square Residual)	< 0.08	0.05

Table 4 summarizes model-fit indices, which indicated acceptable-to-good fit: $\chi^2/\text{df} = 2.17$, GFI = .94, CFI = .96, TLI = .95, RMSEA = .06, SRMR = .05. These values meet commonly recommended thresholds for unidimensional measurement models (e.g., Hu & Bentler, 1999; Kline, 2015). No post-hoc correlated error terms were added, preserving model parsimony. Taken together, the CFA results support the unidimensional structure and construct validity of the PPBS.

Structural Equation Modelling (SEM) Findings

Table 5. Combined Correlation and Descriptive Statistics

Variable	1	2	3	4	5	Mean	N	Std. Deviation	Kurtosis	Skewness
Emotional Regulation	1.000	0.441*	0.585*	0.644*	0.599*	42.72	576	10.74	0.276	0.195
Parental Blaming	0.441*	1.000	0.385*	0.338*	0.376*	37.77	576	13.59	-0.615	0.441
Stress	0.585*	0.385*	1.000	0.760*	0.693*	13.50	576	3.64	0.158	0.463
Depression	0.644*	0.338*	0.760*	1.000	0.719*	12.46	576	4.31	0.472	0.937
Anxiety	0.599*	0.376*	0.693*	0.719*	1.000	10.79	576	3.10	0.276	0.870

* $p < .001$

Descriptive statistics and Pearson correlations indicated acceptable distributional properties (skewness/kurtosis within ± 1), thereby supporting the use of Structural Equation Modelling (SEM). Scores on emotion regulation (higher = greater difficulty) were positively associated with stress ($r = .585$), depression ($r = .644$), and anxiety ($r = .599$; all $ps < .001$), consistent with greater regulatory difficulty accompanying higher distress. The present study found a positive correlation between perceived parental blame and stress ($r = .385$), depression ($r = .338$), and anxiety ($r = .376$). All of the p -values were less than .001. The strongest intercorrelations among outcomes were observed between stress and depression ($r = .760$) and between depression and anxiety ($r = .719$; $ps < .001$), reflecting shared variance across distress indicators. In accordance with the wording guidance provided by Reviewer 1, the present study focuses on associations rather than effects; that is, individuals who reported higher levels of perceived parental blame tended to

report higher levels of psychological distress. These bivariate patterns constituted the empirical basis for the subsequent testing of the mediation model with emotion regulation in Structural Equation Models (SEM) analyses. For interpretability, the ER variable in SEM was analyzed as the Emotion-Regulation Difficulties (ERD) composite (higher = greater difficulty) described in Methods (§2.4.2).

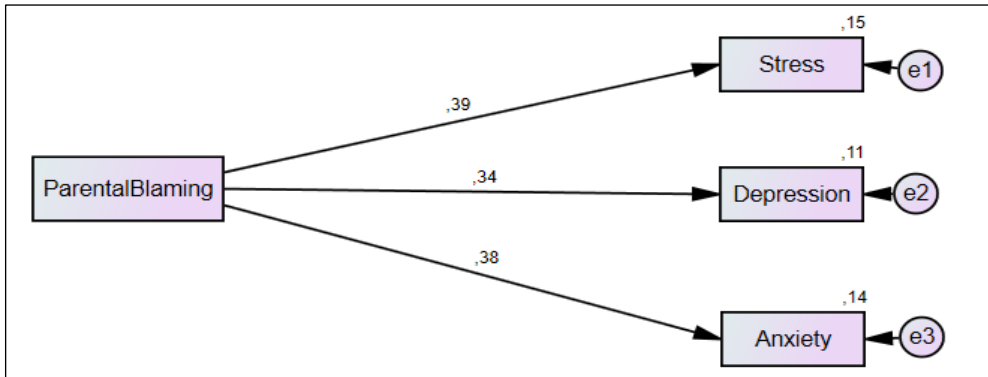


Figure 3. First Model Related to the Hypothesis

As demonstrated in Figure 3, the initial model (first figure) illustrates the direct effects between Parental Blaming and Stress, Depression and Anxiety. In the model, direct effects of. These findings show that parental blame is directly related to psychological distress. Parental Blaming has significant $\beta = .39$ on Stress, $\beta = .34$ on Depression and $\beta = .38$ on Anxiety

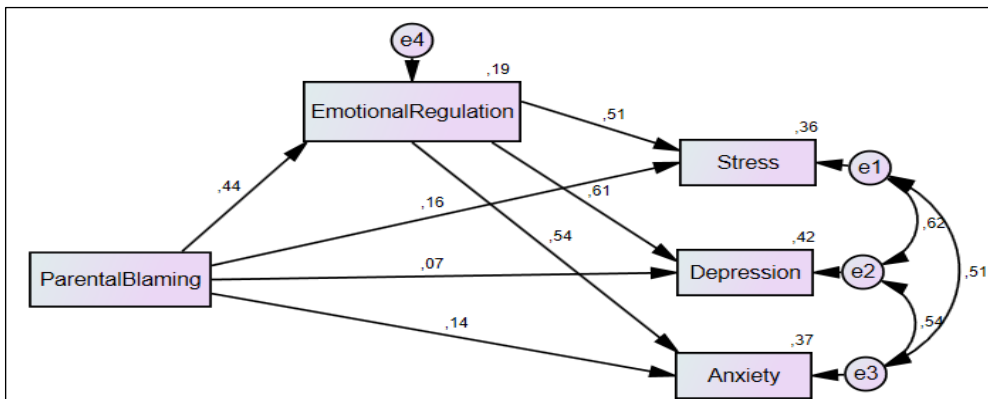


Figure 4. Structural Model for the Mediator Model

As demonstrated in Figure 4, the mediation model delineates emotion-regulation difficulties (higher scores indicating greater difficulty) as the mediator.

The present study found a positive correlation between perceived parental blame and emotion-regulation difficulties, with a significant beta value of .44 and a p-value of .002. Conversely, emotion-regulation difficulties demonstrated a positive association with stress ($\beta = .51$, $p = .005$), depression ($\beta = .54$), and anxiety ($\beta = .37$). Following the incorporation of the mediator, the direct pathways from perceived parental blame were diminished – to stress ($\beta = .16$, $p = .004$), to depression ($\beta = .07$, n.s.), and to anxiety ($\beta = .14$, $p = .006$) – consistent with partial mediation for stress and anxiety and indirect-only mediation for depression within a cross-sectional framework. The coefficients are standardized (β) and represent associations.

Table 6. Standardized Direct and Indirect Effects with Bias-Corrected 95% Bootstrap CIs (5,000 resamples)

Outcome	Path (Predictor → Mediator → Outcome)	Direct β	$p(\text{direct})$	Indirect β	95% CI (LL, UL)	Total β	Mediation pattern
ER difficulties	PPBS → ERD	.441	.002	—	—	.441	—
Stress	PPBS → (—) → Stress	.158	.004	.227	[.177, .273]	.385	Partial
Depression	PPBS → (—) → Depression	.068	.114	.270	[.214, .314]	.338	Indirect-only
Anxiety	PPBS → (—) → Anxiety	.139	.006	.237	[.177, .273]	.376	Partial
Stress	ERD → Stress	.515	.005	—	—	.515	—
Depression	ERD → Depression	.540	—	—	—	.540	—
Anxiety	ERD → Anxiety	.370	—	—	—	.370	—

Note. PPBS = Perceived Parental Blame Scale. ERD = Emotion-regulation difficulties (**higher scores = greater difficulty**). Indirect paths represent PPBS → ERD → outcome. Indirect effects are considered significant when the **bias-corrected 95% CI excludes 0**. All coefficients are standardized (β); two-tailed tests; 5,000 bootstrap resamples.

Brief interpretation. Bias-corrected bootstrapping indicated significant indirect associations from perceived parental blame to stress ($\beta_{\text{ind}} = .227$, 95% CI [.177, .273]), depression ($\beta_{\text{ind}} = .270$, 95% CI [.214, .314]), and anxiety ($\beta_{\text{ind}} = .237$, 95% CI [.177, .273]) via emotion-regulation difficulties. After including the mediator, direct paths from PPBS were attenuated (stress $\beta = .16$, $p = .004$; depression $\beta = .07$, $p = .114$; anxiety $\beta = .14$, $p = .006$), consistent with partial mediation for stress and anxiety and indirect-only mediation for depression within a cross-sectional framework.

Discussion

The present study developed the Perceived Parental Blame Scale (PPBS) and examined its associations with adult psychological health, focusing on the indirect (mediated) association via emotion regulation. The present study will

discuss both the scale development process and the structural equation modelling (SEM) findings in order to illuminate how retrospectively perceived parental blaming relates to current stress, depression, and anxiety. Throughout this study, we have been meticulous in our avoidance of causal language, in accordance with the cross-sectional design (Maxwell & Cole, 2007).

Evaluation of Scale Development Findings

The PPBS was designed to measure the extent to which individuals perceived persistent blaming from their parents during childhood. Results from exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) revealed a single-factor structure with high internal consistency. The KMO value (.940) confirmed sample adequacy, while Bartlett's test ($\chi^2 = 9098.604$, $p < .001$) indicated significant inter-item correlations. The total variance explained (67.60%) strongly reflects the underlying construct (Field, 2013). Furthermore, the Cronbach's alpha coefficient (.965) demonstrates excellent internal reliability (Nunnally & Bernstein, 1994).

These findings confirm that the scale is both valid and reliable. As there is currently no standardized instrument measuring parental blame specifically, this study makes an important contribution to the literature by providing a tool to assess emotionally adverse family dynamics. The PPBS allows researchers and clinicians to evaluate individuals' retrospective perceptions of parental blaming and explore its long-term impact on psychological well-being.

Evaluation of Structural Equation Model Findings

Subsequent to scale development, the SEM (Structural Equation Model) was utilised to investigate the associations between perceived parental blame and adult psychological health, and to evaluate the mediating role of emotion regulation. In the present analyses, the emotion-regulation composite was scored such that higher scores indicate greater difficulties (i.e., emotion-regulation difficulties). The findings indicated that perceived parental blame exhibited direct and positive associations with anxiety ($\beta = .139$, $p = .006$) and stress ($\beta = .158$, $p = .004$), though not with depression ($\beta = .068$, $p = .114$). The incorporation of the mediator resulted in the emergence of indirect pathways, as evidenced by the significant positive associations observed for depression ($\beta_{\text{indirect}} = .270$), stress ($\beta_{\text{indirect}} = .227$), and anxiety ($\beta_{\text{indirect}} = .237$). Furthermore, the findings indicated a positive relationship between emotion-regulation difficulties and stress ($\beta = .515$, $p = .005$), depression ($\beta = .540$), and anxiety ($\beta = .370$), thereby corroborating the bivariate correlations.

Collectively, these results align with a partial mediation pattern for stress and anxiety and an indirect-only (full mediation) pattern for depression within a cross-sectional model. It is imperative to acknowledge that these are associations,

and that inferences regarding temporal precedence or causality necessitate longitudinal data (Maxwell & Cole, 2007).

Comparison with Literature

The PPBS makes a unique contribution by isolating perceived parental blame as a measurable construct, rather than relying solely on broader indicators of parental criticism or bonding. It has been demonstrated in a number of studies that there is a link between critical or accusatory family dynamics and negative psychological outcomes. For example, see Scharf & Goldner (2018), Luyten et al. (2020) and Brenning et al. (2020). For instance, Luyten et al. (2020) reported that persistent criticism in childhood was associated with lower self-efficacy and higher anxiety in adulthood, while Mikulincer and Shaver (2017) described pathways from adverse caregiving to interpersonal vulnerability. The present findings extend this body of literature by operationalising the concept of perceived parental blame and demonstrating that its associations with stress, depression and anxiety are statistically mediated by emotion-regulation difficulties.

Converging evidence also links parental dynamics with emotion (dys)regulation (e.g., Brumariu & Kerns, 2018; Soenens & Vansteenkiste, 2010; Galliher et al., 2017). The current findings on mediation align with these models, suggesting that early relational environments influence regulatory repertoires, which in turn contribute to adult distress. Furthermore, emerging research highlights the importance of positive family processes in supporting well-being. Arslan (2024) posited that positive parenting is associated with adolescents' well-being via prosocial behaviour and meaning in life. In a similar vein, Vallejo-Correa et al. (2025) emphasised the protective role of secure family relationships. Finally, Karagiannaki and Burchardt (2024) highlighted the effects of family structures on developmental outcomes. The present study offers a complementary contribution to the extant literature in this field, demonstrating how perceived blame may potentially compromise emotional development. In addition, it provides a reliable tool for the assessment of this phenomenon.

The observed pattern, which has been demonstrated to be associated with heightened levels of perceived parental blame and subsequent emotion-regulation difficulties, is consistent with theoretical perspectives concerning the aetiology of psychological distress (e.g., Bowlby, 1988; Mikulincer & Shaver, 2016). However, the present study does not make any assertions regarding the directionality of the relationships observed.

The Relationship Between Emotion Regulation and Psychological Health

Emotion regulation has been identified as a pivotal component in the management of stress and the cultivation of psychological resilience (Gross, 2015;

Aldao et al., 2010; Berking & Whitley, 2014). In the present study, the presence of emotion-regulation difficulties was found to significantly mediate the associations between perceived parental blame and psychological symptoms. This finding aligns with prior research (Gratz & Roemer, 2004; Troy et al., 2018), which demonstrated that enhanced regulatory capacity, characterised by reduced difficulties, is associated with reduced levels of stress, anxiety, and depression. The comparatively strong standardized path to stress ($\beta = .515$) underscores the value of targeting regulation processes in contexts where retrospective blame is salient. These results align with developmental models that propose that the socialisation of emotions within the family context is a contributing factor to the development of later coping mechanisms (Morris et al., 2007; Luyten et al., 2020).

Place of the Scale in the Literature and Suggestions for Future Research

The PPBS addresses a measurement gap regarding early relational adversity. Whereas widely utilised tools such as the Inventory of Parent and Peer Attachment (IPPA; Armsden & Greenberg, 1987) assess broader constructs (attachment, support), the PPBS focuses specifically on perceived parental blame. This specificity enables the execution of nuanced tests of how early blame relates to later mental health, thereby inviting both theoretical and applied investigations.

Future research should (a) examine cross-cultural validity, (b) test the measure in diverse demographic contexts (education levels, marital status, and non-clinical vs. clinical samples), and (c) evaluate interventions that target emotion regulation among those reporting high perceived parental blame. Pre-registered longitudinal designs and experimental manipulations would be particularly informative for tracing temporal sequences and for testing potential moderation (e.g., whether high regulatory capacity or supportive relationships buffer the PPBS–distress associations).

Theoretical and Practical Contributions

Theoretically, the PPBS contributes to the advancement of the study of parent–child relational dynamics by offering a well-defined construct and a validated instrument. The SEM findings provide further elucidation on the interrelation between perceived parental blame, emotion-regulation difficulties, and indicators of adult distress.

The practical implications of the findings indicate that the PPBS should be integrated into family therapy and adult psychotherapy intake procedures for the purpose of screening for retrospective blame dynamics that may correspond to current interpersonal schemas. Case formulations have the capacity to explicitly

establish a connection between internalised blame narratives and regulatory patterns (e.g., suppression/avoidance versus cognitive reappraisal), thereby providing a framework for the development of skills-based interventions. Psychoeducational and parent education programmes may emphasise non-blaming communication, validation, and problem-focused support to reduce the likelihood that blame-based exchanges are internalised as enduring self-views.

Limitations and Suggestions for Future Studies

Theoretically, the PPBS contributes to the advancement of the study of parent–child relational dynamics by offering a well-defined construct and a validated instrument. The SEM findings provide further elucidation on the interrelation between perceived parental blame, emotion-regulation difficulties, and indicators of adult distress.

The practical implications of the findings indicate that the PPBS should be integrated into family therapy and adult psychotherapy intake procedures for the purpose of screening for retrospective blame dynamics that may correspond to current interpersonal schemas. Case formulations have the capacity to explicitly establish a connection between internalised blame narratives and regulatory patterns (e.g., suppression/avoidance versus cognitive reappraisal), thereby providing a framework for the development of skills-based interventions. Psychoeducational and parent education programmes may emphasise non-blaming communication, validation, and problem-focused support to reduce the likelihood that blame-based exchanges are internalised as enduring self-views.

Conclusion

The present study emphasises the significance of perceived parental blame for adult psychological distress and identifies emotion-regulation difficulties as a pivotal mediating mechanism connecting these constructs. The PPBS offers a psychometrically robust instrument for the assessment of a hitherto under-examined facet of childhood relational adversity. The findings emphasise the importance of early prevention and skills-based interventions that enhance emotion regulation. The present study contributes to the field by isolating a measurable dimension of family dynamics that relates to adult well-being. This is achieved by means of an evidence-based approach to research and practice. The integration of the PPBS into clinical assessments, educational programmes, and family therapy has the potential to promote psychological resilience and emotional health across the lifespan.

Author's Notes

Data availability statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with Ethical Standards

Ethical Approval: Ethical approval was not required for this study because it involved minimal risk survey research with adult participants and did not include any sensitive personal data. The study was conducted in accordance with the ethical standards of the Declaration of Helsinki.

Informed Consent: All participants were informed about the study and provided written informed consent prior to participation. Before collecting data for this research, permission was obtained from all participants. Approval statements were also received from each person whose data was collected.

Conflict of Interest: The author declares no conflict of interest.

Consent to Publish: The author gives full consent for the publication of this manuscript.

Funding Statement: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions: Nesrullah Okan conceived the study, collected and analyzed the data, and wrote the manuscript.

References

- Aldao, A., Nolen-Hoeksema, S., & Schweizer, S. (2010). Emotion-regulation strategies across psychopathology: A meta-analytic review. *Clinical Psychology Review, 30*(2), 217–237. <https://doi.org/10.1016/j.cpr.2009.11.004>
- Armsden, G. C., & Greenberg, M. T. (1987). The Inventory of Parent and Peer Attachment: Individual differences and their relationship to psychological well-being in adolescence. *Journal of Youth and Adolescence, 16*(5), 427–454. <https://doi.org/10.1007/BF02202939>
- Arsilan, G. (2024). Positive parenting, prosocial behavior, meaning in life, and subjective well-being in adolescents: A parallel and serial mediation analysis. *Child Indicators Research, 17*(5), 1547–1561. <https://doi.org/10.1007/s12187-024-10144-3>
- Babbie, E. (2020). *The Practice of Social Research* (15th ed.). Cengage Learning.
- Berking, M., & Whitley, B. (2014). *Affect regulation training: A practitioner's guide to adaptive emotion regulation*. Springer. <https://doi.org/10.1007/978-1-4939-1022-9>

- Bowlby, J. (1988). *A secure base: Parent-child attachment and healthy human development*. Basic Books.
- Brenning, K., Soenens, B., Braet, C., & Bosmans, G. (2020). The role of parental attachment in adolescent depressive symptoms and self-esteem. *Journal of Child and Family Studies*, 29(5), 1313-1325. <https://doi.org/10.1007/s10826-020-01694-9>
- Brown, G. W., & Harris, T. (2020). Social origins of depression: A study of psychiatric disorder in women. *Routledge*. <https://doi.org/10.4324/9780203787212>
- Brumariu, L. E. (2015). Parent-child attachment and emotion regulation. *New Directions for Child and Adolescent Development*, 2015(148), 31-45. <https://doi.org/10.1002/cad.20098>
- Brumariu, L. E., & Kerns, K. A. (2018). Parent-child attachment and emotion regulation in childhood and adolescence. *Emotion Review*, 10(1), 14-23. <https://doi.org/10.1177/1754073917693358>
- Bryman, A. (2016). *Social Research Methods* (5th ed.). Oxford University Press.
- Chang, Y., Zhang, L., Liu, H., & Yu, G. (2021). The impact of parental blame on mental health: The mediating role of self-esteem and cognitive reappraisal. *Journal of Child and Family Studies*, 30(5), 1123-1138. <https://doi.org/10.1007/s10826-021-01975-8>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319.
- Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Sage Publications.
- Dagan, O., Facompré, C. R., & Bernard, K. (2018). Adult attachment representations and depressive symptoms: A meta-analysis. *Journal of Affective Disorders*, 236, 274-290. <https://doi.org/10.1016/j.jad.2018.04.091>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage publications.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage.
- Gallagher, R. V., Evans, C. M., & Weiser, D. A. (2017). Social-cognitive influences on the link between parenting and adolescent internalizing symptoms. *Journal of Adolescence*, 57, 77-87. <https://doi.org/10.1016/j.adolescence.2017.03.003>
- Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, 26(1), 41-54. <https://doi.org/10.1023/B:JOBA.0000007455.08539.94>
- Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, 26(1), 1-26. <https://doi.org/10.1080/1047840X.2014.940781>
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of*

- Personality and Social Psychology*, 85(2), 348-362. <https://doi.org/10.1037/0022-3514.85.2.348>
- Gross, J. J., & Jazaieri, H. (2014). Emotion, emotion regulation, and psychopathology: An affective science perspective. *Clinical Psychological Science*, 2(4), 387-401. <https://doi.org/10.1177/2167702614536164>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.
- Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 44(2), 227-239. <https://doi.org/10.1348/014466505X29657>
- Hooley, J. M., & Parker, H. A. (2006). Measuring expressed emotion: An evaluation of the shortcuts. *Journal of Family Psychology*, 20(3), 386-396. <https://doi.org/10.1037/0893-3200.20.3.386>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Infurna, F. J., & Luthar, S. S. (2018). Parents' adverse childhood experiences and depressive symptoms in adulthood: Mediating role of psychological resilience. *Development and Psychopathology*, 30(2), 501-516. <https://doi.org/10.1017/S0954579417001010>
- Ingram, R. E., Miranda, J., & Segal, Z. V. (2006). *Cognitive vulnerability to depression*. Guilford Press.
- Joormann, J., & Stanton, C. H. (2016). Examining emotion regulation in depression: A review and future directions. *Behaviour Research and Therapy*, 86, 35-49. <https://doi.org/10.1016/j.brat.2016.07.007>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. <https://doi.org/10.1007/BF02291575>
- Karagiannaki, E., & Burchardt, T. (2024). Living arrangements, intra-household inequality and children's deprivation: Evidence from EU-SILC. *Child Indicators Research*, 17(5), 2319-2359. <https://doi.org/10.1007/s12187-024-10149-y>
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Kline, R. B., Guo, W., & Barrett, L. (2022). Long-term effects of childhood parental blame on adult mental health: A structural equation modeling approach. *Developmental Psychology*, 58(3), 456-470. <https://doi.org/10.1037/dev0001304>
- Knappe, S., Petermann, F., & Hoyer, J. (2019). The role of parental blame in the development of emotional disorders. *Child and Adolescent Psychiatry and Mental Health*, 13(1), 5-17. <https://doi.org/10.1186/s13034-019-0274-9>
- Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. Guilford Press.

- Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335-343. [https://doi.org/10.1016/0005-7967\(94\)00075-U](https://doi.org/10.1016/0005-7967(94)00075-U)
- Luyten, P., Blatt, S. J., Van Houdenhove, B., & Corveleyn, J. (2020). Depression research and treatment: What can we learn from a developmental psychopathology approach? *Clinical Psychology Review*, 81, 101871. <https://doi.org/10.1016/j.cpr.2020.101871>
- Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, 12(1), 23-44. <https://doi.org/10.1037/1082-989X.12.1.23>
- Mikulincer, M., & Shaver, P. R. (2017). *Attachment in adulthood: Structure, dynamics, and change* (2nd ed.). Guilford Press.
- Monroe, S. M., & Simons, A. D. (1991). Diathesis-stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin*, 110(3), 406-425. <https://doi.org/10.1037/0033-2909.110.3.406>
- Morris, A. S., Criss, M. M., Silk, J. S., & Houlberg, B. J. (2017). The impact of parenting on emotion regulation during childhood and adolescence. *Child Development Perspectives*, 11(4), 233-238. <https://doi.org/10.1111/cdep.12238>
- Naragon-Gainey, K. (2020). Affective processes in depression and anxiety: Emotion regulation, anhedonia, and attentional control. *Journal of Affective Disorders*, 275, 372-380. <https://doi.org/10.1016/j.jad.2020.06.021>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- Parker, G., Tupling, H., & Brown, L. B. (1979). A parental bonding instrument. *British Journal of Medical Psychology*, 52(1), 1-10. <https://doi.org/10.1111/j.2044-8341.1979.tb02487.x>
- Renshaw, K. D. (2008). The perceived criticism construct: Measurement issues and future directions. *Clinical Psychology Review*, 28(3), 328-338. <https://doi.org/10.1016/j.cpr.2007.06.002>
- Ryff, C. D. (2014). Psychological well-being revisited: Advances in the science and practice of eudaimonia. *Psychotherapy and Psychosomatics*, 83(1), 10-28. <https://doi.org/10.1159/000353263>
- Sarıçam, H. (2018). The psychometric properties of Turkish version of Depression Anxiety Stress Scale-21 (DASS-21) in health control and clinical samples. *Journal of Cognitive Behavioral Psychotherapy and Research*, 7(1), 19-30.
- Schäfer, J. Ö., Naumann, E., Holmes, E. A., Tuschen-Caffier, B., & Samson, A. C. (2017). Emotion regulation strategies in depressive and anxiety symptoms in youth: A meta-analytic review. *Journal of Youth and Adolescence*, 46(2), 261-276. <https://doi.org/10.1007/s10964-016-0585-0>
- Scharf, M., & Goldner, L. (2018). Long-term effects of parental criticism and overprotection: The role of parenting styles and attachment representations.

- Journal of Family Psychology*, 32(6), 762-773. <https://doi.org/10.1037/fam0000433>
- Schleider, J. L., & Weisz, J. R. (2017). Family influences on depressive symptom trajectories in clinic-referred youth. *Journal of Abnormal Child Psychology*, 45(6), 1243-1255. <https://doi.org/10.1007/s10802-016-0237-9>
- Soenens, B., & Vansteenkiste, M. (2010). A theoretical upgrade of the concept of parental psychological control: Proposing new insights on the basis of self-determination theory. *Developmental Review*, 30(2), 74-99. <https://doi.org/10.1016/j.dr.2009.11.001>
- Southward, M. W., & Cheavens, J. S. (2020). Identifying core mechanisms of change in cognitive therapies: An individual patient data meta-analysis. *Clinical Psychological Science*, 8(4), 627-644. <https://doi.org/10.1177/2167702620917458>
- Steele, H., Steele, M., & Bate, J. (2018). Attachment and trauma in children. *Journal of Traumatic Stress*, 31(4), 589-600. <https://doi.org/10.1002/jts.22307>
- Troy, A. S., Shallcross, A. J., & Mauss, I. B. (2018). A person-by-situation approach to emotion regulation: Cognitive reappraisal can either help or hurt, depending on the context. *Psychological Science*, 24(12), 2505-2514. <https://doi.org/10.1177/0956797618496434>
- Vallejo-Correa, V., Oyanedel, J. C., & Cortés, F. P. (2025). Subjective well-being, parental figures, and quality of family relations: A study of the International Questionnaire on Child Welfare among Chilean 12-year-olds. *Child Indicators Research*. Advance online publication. <https://doi.org/10.1007/s12187-025-10222-0>
- Yiğit, İ., & Guzey Yiğit, M. (2019). Psychometric Properties of Turkish Version of Difficulties in Emotion Regulation Scale-Brief Form (DERS-16). *Current Psychology*, 38, 1503-1511.

EFFECT OF AEROBIC EXERCISE ON RELAPSE TENDENCY IN INDIVIDUALS WITH METHAMPHETAMINE USE DISORDER IN A VIRTUAL SOCIAL ENVIRONMENT: MEDIATED BY CUE-INDUCED CRAVING AND INTEROCEPTIVE SENSIBILITY

*Xiaoxia Liu and Zefan Fan¹**

¹ Department of Physical Education, Tongling University, Tongling 244061, China.

Abstract

Interoceptive sensibility (IS) is a novel therapeutic target in drug relapse prevention. The beneficial mechanisms of exercise on methamphetamine relapse remain unclear. This study assessed the effects of aerobic exercise on cue-induced craving reactivity (including heart rate variability [HRV], methamphetamine-craving, and methamphetamine-using), IS, and relapse tendency in a methamphetamine-related virtual reality environment among individuals with methamphetamine use disorders (MUD). We further examined the role of craving reactivity and IS in aerobic exercise and relapse tendency. Exercise ($n = 36$) and control ($n = 35$) groups received eight weeks of aerobic exercise or health education, respectively (5 times/week, 60 minutes/session). HRV_{cue} , methamphetamine-craving, methamphetamine-using, IS, and relapse tendency were measured at pre-test and post-test on the same day. The exercise group exhibited a larger decrease in craving reactivity, IS, and relapse tendency from the pre-test to the post-test compared to the control group. Craving reactivity and IS changes mediated the relationship between exercise intervention and decreased relapse tendency ($\beta = 1.505\text{--}2.009$, 95% CI: 0.378–3.736). Aerobic exercise improves relapse tendency in individuals with MUD. The underlying mechanisms may include reduced cue-induced craving reactivity and IS.

Keywords: aerobic exercise; cue-induced craving; interoceptive sensibility; METH-related virtual reality environment; relapse tendency

* Correspondence concerning this article should be addressed to Zefan Fan, Department of Physical Education, Tongling University, 1335 Cuihu 4th Road Economic and Technological Development Zone, Tongling 244061, Anhui Province, People's Republic of China.
Telephone: +8613696527047.
E-mail: fl700615@163.com

1. Introduction

Methamphetamine (METH), a highly addictive psychostimulant, is the second most widely used class of illicit drugs worldwide (Chen et al., 2024; United Nations Office on Drugs and Crime, 2024). One main challenge in treating METH addiction is the high rate of relapse during abstinence (Tan et al., 2019; United Nations Office on Drugs and Crime, 2024). When drug users are exposed to drug-related cues, such as drugs, drug-use environments, and paraphernalia, an intense desire for the drug could be induced (Tan et al., 2019; Wang, Liu, & Shen, 2019). This reactivity to drug-related cues, including mental (e.g., craving and attention bias), psychophysiological (e.g., heart rate variability [HRV], skin conductance, and brain activity), and behavioral (e.g., the possibility of subsequent drug-using and drug-seeking) responses, is usually described as craving and results in drug-seeking and drug-taking behaviors, which can cause relapse (Chen et al., 2024; Sayette, 2016; Tan et al., 2019). For methamphetamine use disorder (MUD), craving is one of the most important symptoms and serves to maintain this addictive behavior and relapse (Wang et al., 2019). During abstinence from METH, craving is always elicited by the METH-related cues (Hormes, 2017; Wang et al., 2019). Therefore, reducing METH-related cue-induced craving reactivity may help prevent relapse in individuals with MUD.

Despite amphetamine being the second most common illicit drug abused worldwide (Chen et al., 2024; United Nations Office on Drugs and Crime, 2024), no approved pharmacotherapies are available for amphetamine dependence (AshaRani et al., 2020). Multiple systematic reviews of pharmacotherapies reveal that no medication provides sufficient evidence to support its use in the routine clinical management of amphetamine use disorder (AshaRani et al., 2020; Chan et al., 2019; Lee, Jenner, Harney, & Cameron, 2018). In the absence of effective pharmacotherapy, psychotherapy is considered the first line of treatment, although this approach lacks scientific evidence as a standalone treatment (Donovan & Wells, 2007) due to high relapse rates (AshaRani et al., 2020; McKetin et al., 2012). Nevertheless, a previous study investigated the effectiveness of an 8-week brief intervention compared with an 8-week psychoeducation program and showed that both the brief intervention and psychoeducation are effective in reducing methamphetamine use among individuals with MUD (Srisurapanont, Sombatmai, & Boripuntakul, 2007). Besides psychotherapy, other non-pharmacological interventions such as physical exercise and other programs targeting MUD have been studied in recent years (AshaRani et al., 2020; Liu & Wang, 2021, 2023; Rawson et al., 2015). Given the lack of evidence supporting the use of pharmacotherapies to treat MUD, the efficacy of alternative treatments must be evaluated.

Physical exercise is a promising, medication-free, and cost-effective adjunct treatment for MUD (AshaRani et al., 2020; Lynch, Peterson, Sanchez, Abel, & Smith, 2013). Studies have reported that exercise-based interventions may prevent relapse of MUD not only through a direct pathway but also through indirect pathways, such as the mediating roles of intrinsic restraint and yearning for drugs, as well as by remodeling the reward pathway (Lynch et al., 2013; Morais, Pita, Fontes-Ribeiro, & Pereira, 2018; Ouyang et al., 2020; Song et al., 2023). Traditionally, drug-cue craving reactivity relies mostly on pictures or videos presenting drugs and drug-associated paraphernalia. However, these methods typically elicit modest subjective craving (Avants, Margolin, Kosten, & Cooney, 1995), less effectively inducing a reliable physiological reactivity (Dudish-Poulsen & Hatsukami, 1997; Ooteman, Koeter, Verheul, Schippers, & van den Brink, 2006). Therefore, researchers have developed METH-related virtual reality (VR) social environments, which offer unique advantages due to their full immersion experience and realism (Culbertson et al., 2010; Tan et al., 2019) and are sufficient to elicit more reliable cravings and corresponding physiological reactivity than traditional methods (Wang, Shen, & Wu, 2018; Wang et al., 2019). Exploring the effects of physical exercise on METH-cue craving reactivity in VR cue-exposure environments may have significant practical and clinical implications for evaluating the effectiveness of exercise strategies.

Recently, interoception (i.e., receiving, processing, and integrating body-relevant signals with external stimuli to affect ongoing motivated behavior) has been recognized as another important process that contributes to the degree to which an individual approaches or avoids drug abuse and relapse (Paulus & Stewart, 2014; Reiner & Shaham, 2022). Long-term substance abuse has been found to negatively impact the brain's ability to process interoceptive information and impact the reward system (Paulus, Tapert, & Schulteis, 2009), leading to decreased sensitivity to natural rewards and increased sensitivity to drug stimuli (Herman, 2023; Volkow, Michaelides, & Baler, 2019). A recent analysis provided evidence that individuals with substance use disorder (SUD) show lower interoceptive accuracy and higher interoceptive sensibility (IS) compared to healthy controls (Jakubczyk et al., 2019a). In the context of SUD, higher IS has been associated with clinical correlates of relapse, such as craving, withdrawal, tolerance, and reward processing (Bonaz et al., 2021; Jakubczyk et al., 2019b; Reiner & Shaham, 2022). Therefore, reducing cue-induced or context-induced high IS may represent novel therapeutic targets for the treatment of drug relapse (Herman, 2023; Reiner & Shaham, 2022). Recent evidence suggests that physical exercise may help people with SUD to better decode and interpret bodily-related signals associated with transient states of homeostatic imbalances that usually trigger consumption (Brevers et al., 2024). Nevertheless,

evidence regarding the effectiveness of physical exercise on IS and the role of IS in the relationship between physical exercise and relapse is currently very limited; therefore, further research is needed to examine this (Brevers et al., 2024; Reiner & Shaham, 2022).

This study aimed to compare the effects of physical exercise versus psychoeducation on METH-cue craving reactivity (including HRV, METH-craving, and METH-using), IS, and relapse tendency in a METH-related VR environment among individuals with MUD. We further examined the role of craving reactivity and IS in physical exercise and relapse tendency.

2. Methods

2.1 Participants

In total, 120 men with MUD from local residential illicit drug isolation and treatment centers were recruited for this study. The participants met the Diagnostic and Statistical Manual of Mental Disorders IV-revised diagnostic criteria, as determined via the Mini-International Neuropsychiatric Interview (Liu & Wang, 2021). The inclusion criteria were (1) physician-administered physical examination and collection of medical history with resting 12-lead electrocardiography (ECG; Mortara, Milwaukee, WI, USA) to determine study eligibility; (2) no history of neurological disease or physical disability; (3) no use of medication affecting cardiovascular and cognitive function over the preceding 3 months; (4) no cognitive disorder or medical contraindications to exercise; (5) sufficient capacity to provide informed consent and complete all assessment instruments; and (6) normal color vision, normal or corrected-to-normal visual acuity, and right-handedness. The exclusion criteria included the presence of musculoskeletal conditions and unstable cardiovascular, pulmonary, or other disorders. Finally, 82 participants who met the inclusion and exclusion criteria were randomly divided into two groups (Figure 1). All participants were informed that participation was voluntary, that they were free to withdraw at any time, and that their personal information would be kept confidential. This study was approved by the ethics committee of X Institutional Review Board, and written informed consent was obtained from all participants.

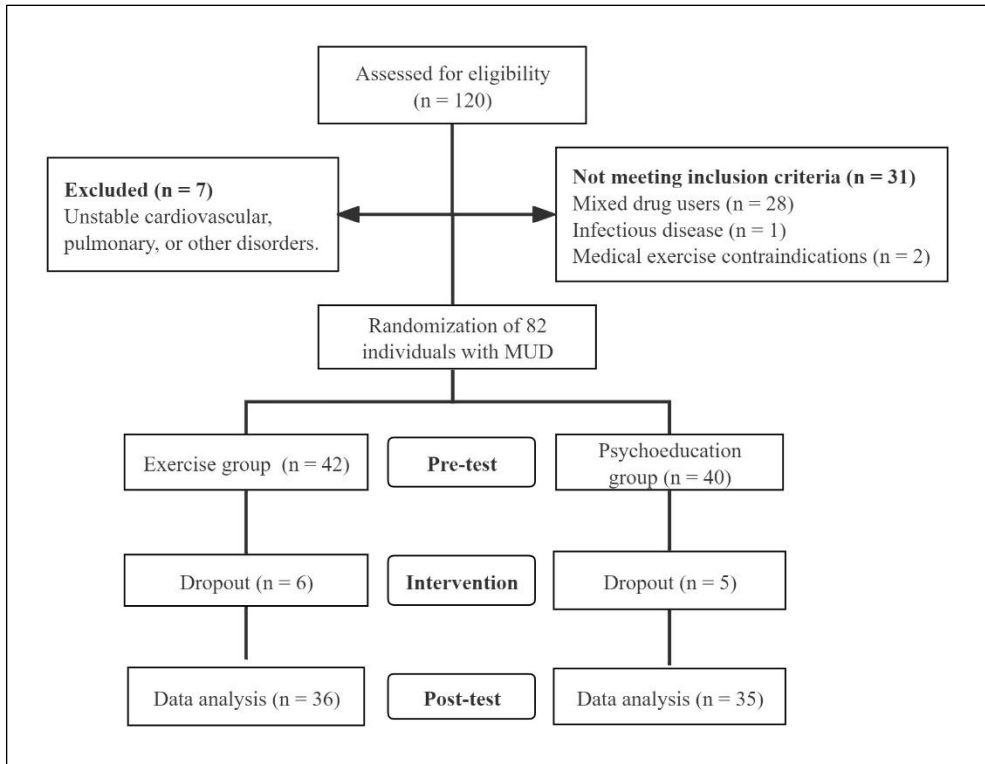


Figure 1. CONSORT diagram presenting the participant flow in the study

2.2 METH-related VR environment

We used the METH-related virtual social environment developed and tested in previous research and promoted in drug rehabilitation centers to elicit subjective psychological and physiological reactivity (Wang et al., 2018, 2019). The video depicts a real-life story of individuals who are using METH together and who invite the observers to take METH. Specifically, the video is divided into six scenes (Figure 2) (Wang et al., 2018; Wang et al., 2019). During the first scene (~60 s), the participants saw a black screen and heard voices luring them to take drugs (sounds likely to invite the participants to take METH, i.e., auditory cues with social interaction). In the second scene (~60 s), the participants viewed a person playing with “ice” and drug paraphernalia (i.e., drugs and drug-associated paraphernalia). During the third scene (~180 s), the participants experienced themselves in a room (with a first-person perspective), in which they could see six people using METH together (i.e., METH-use environment). During the fourth scene (~60 s), the participants viewed close-up scenes highlighting the facial expressions of all six

people using METH (i.e., facial expression information). During the fifth scene (~60 s), the participants viewed close-up scenes highlighting the specific way that METH was being used (i.e., action cues). During the sixth scene (~60 s), the prepared drug and related paraphernalia were presented in a man's hand and moved toward the participants. This scene gave the participants the impression that they were being invited to use METH (i.e., "psychological interaction" with the observers). All the audio tracks were processed using Nuendo to create 5.1 analog sound channels. The VR video was rendered onto a spheroid to create a subject-centric immersive experience, and the viewing angle was pre-determined for specific scenes requiring close-ups (Wang et al., 2018, 2019).

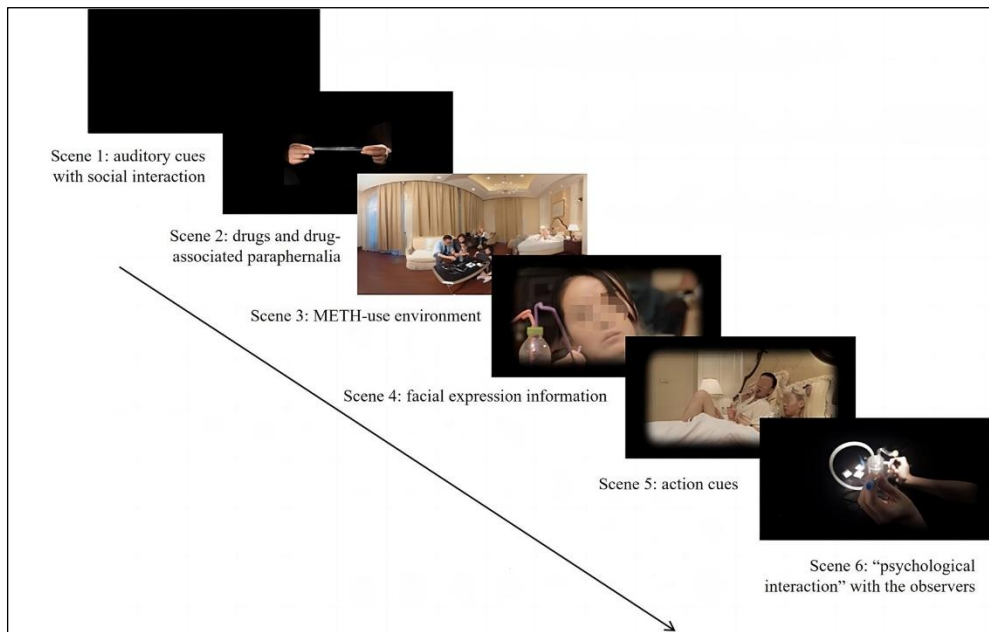


Figure 2. Diagram of the METH-related VR environment

2.3 Procedure and outcome measures

All outcome measures were assessed in a quiet room at a stable temperature (23–25 °C). The participants were asked to avoid strenuous exercise for 48 hours before the test. The participants' demographic characteristics, drug-use history, and HRV_{baseline} were assessed at baseline. A researcher-developed questionnaire was used to assess age, education, and drug-use history (including duration of drug use, frequency, and total amount of METH used). Height, weight, and body mass index were measured using an electronic weight tester in the National Physical Fitness

Monitor (Beijing, China). Before testing the $HRV_{baseline}$, all participants rested comfortably in a chair for at least 5 minutes. They were then introduced to the equipment (i.e., HRV recording device and headphones) and assisted with putting on the equipment. After the participants felt comfortable with all the settings, they were asked to remain seated without speaking or moving. When the test began, the participants were instructed to breathe rhythmically at the same frequency to record an 8-minute $HRV_{baseline}$ (Liu & Wang, 2021). Meanwhile, a heart rate monitor (Polar FS2C, Polar Electro, Oy, Kempele, Finland) was used to measure the resting HR.

HRV_{cue} , METH-craving, METH-using, IS, and relapse tendency, in turn, were measured at pre-test and post-test on the same day. During the formal experiment, all participants were assisted with putting on the equipment (i.e., HRV recording device, VR helmet, and headphones). After the participants felt comfortable with all the settings, they were asked to remain seated without speaking or moving and breathe rhythmically at the same frequency. An 8-minute period of the resting condition continuous HRV signals (i.e., $HRV_{before\ video}$) was recorded. Then, the participants were required to view an 8-minute METH-cue VR video with concurrent recording of HRV signals (i.e., $HRV_{during\ video}$) (Wang et al., 2018; Wang et al., 2019). Immediately after the VR video finished, the participants self-reported METH-craving, METH-using, IS, and relapse tendency in turn. Unlike subjective reports, HRV is an objectively recorded natural physiological response that does not involve learning or practice effects but is highly influenced by the testing environment. Therefore, we calculated HRV_{cue} parameters (i.e., $HRV_{during\ video}$ minus $HRV_{before\ video}$) for statistical analysis.

2.3.1 METH-cue VR-induced HRV reactivity

As per the previously established method, HRV was measured using a wearable HR monitor and a chest belt (RS800CX Polar Electro Oy, Kempele, Finland) with a sampling rate of 1,000 Hz for subsequent analyses of HRV (Eggenberger et al., 2020; Williams et al., 2017). We calculated HRV parameters, including time-domain (i.e., standard deviation of normal-to-normal intervals [SDNN], root mean square differences of the standard deviation [RMSSD], percentage of beats that changed more than 50 ms from the previous beat [pNN50]), and frequency-domain (i.e., low-frequency [LF], high-frequency [HF], and low-frequency to high-frequency ratio [LF/HF]; using autoregressive estimates) (Dolezal et al., 2014; Gamelin, Berthoin, Sayah, Libersa, & Bosquet, 2007; Liu & Wang, 2021).

2.3.2 METH-cue VR-induced subjective METH-craving

A visual analogue scale (VAS) (Mottola, 1993) regarding subjective craving assessed the participants' METH-craving. The participants were asked: "How much do you crave METH/ice right now?" (possible scores ranged from 0–10, "0"

indicating “no craving at all” and “10” indicating “extremely strong craving”). The VAS consisted of an 11-point scale, which has good reliability and validity. A higher score indicates greater METH-craving.

2.3.3 The possibility of METH-using immediately after exposure to VR cues

A VAS regarding the possibility of METH-using was used to assess the participants’ METH-using: “If you had access to METH/ice right now, how likely would you be to use it right away?” (possible scores ranged from 0–10; “0” indicating “certainly not” and “10” indicating “certainly”) (Wang et al., 2019).

2.3.4 IS immediately after exposure to VR cues

The Private Body Consciousness Subscale (PBCS) (Miller, Murphy, & Buss, 1981) from the Body Consciousness Questionnaire consists of 5 items reflecting the tendency to focus on internal body sensations. Specifically, it assesses awareness of internal bodily tensions, dry mouth/throat, heart beating, hunger, and body temperature (e.g., “I am sensitive to internal body tensions,” “I’m very aware of changes in my body temperature”) (Jakubczyk et al., 2019a). Items were rated for the extent to which they are characteristic of the respondent on a 5-point scale, ranging from 0 (extremely uncharacteristic) to 4 (extremely characteristic). Overall scores were computed as the mean of all items so that higher scores indicated greater IS. The PBCS has good reliability and validity in clinical samples (Mehling et al., 2009; Todd et al., 2022). Cronbach’s alpha of the PBCS in this study was 0.79.

2.3.5 Relapse tendency immediately after exposure to VR cues

The Inclination of Relapse Scale of Drug Rehabilitated Addicts (Song et al., 2023; Wang, 2001) was used to assess the participants’ subjective relapse tendency. The scale consists of 18 items, each rated on a 6-point Likert scale. Higher scores indicate a greater tendency to relapse. Cronbach’s alpha of the scale in this study was 0.82.

2.4 Intervention

2.4.1 Physical exercise

The exercise group engaged in a treadmill aerobics program 5 times per week, 60 minutes per session, for 8 weeks. The exercise phase consisted of walking or jogging on a treadmill for 5 minutes (warm-up phase). After the warm-up phase, the participants gradually increased the treadmill speed and slope to reach 60–80% maximum HR (defined as $206.9 - 0.67 \times \text{age}$) within 5 minutes under the guidance of a trainer (Liu & Wang, 2023). The participants then maintained this intensity for 40 minutes. If a participant was unable to complete 40 continuous minutes, rest

periods were introduced until the participant accumulated a total of 40 minutes at this intensity (Dolezal et al., 2014; Liu & Wang, 2021). Finally, the participants performed 10 minutes of cool-down activities. The participants wore an HR monitor (Polar FS2C, Polar Electro, Oy, Kempele, Finland) connected to a system monitoring HR in real-time. At least three well-trained, experienced coaches guided each exercise process.

2.4.2 Psychoeducation

The psychoeducation group received an 8-week structured psychoeducation program (5 times per week, 60 minutes per session for 8 weeks) with three trained and experienced psychological health educators. The courses included education on health screening, healthy diet, stress reduction, healthy relationships, and healthy lifestyles.

2.5 Data processing

In line with previous studies, the Polar Pro-Trainer 5 software (RS800CX Polar Electro Oy, Kempele, Finland) was used to transmit the recorded data to the computer (Eggenberger et al., 2020; Liu & Wang, 2021; Williams et al., 2017). The first 30 recorded seconds were deleted to reduce initial motion artifacts, and the subsequent data were used for analysis. Artifacts were processed in two different ways: (1) interpolation of artifacts using Kubios HRV analysis software (version 2.2, Kubios Oy, Kuopio, Finland; settings: custom=0.2, interpolation rate=4 Hz); and (2) deletion of artifacts using MATLAB software (version R2016b, MathWorks Inc., Natick, MA, USA). R–R intervals <200 ms, >2,000 ms, or >20% interval-to-interval deviation were removed. Finally, using the aforementioned method, artifact examination with MATLAB software yielded a mean of 6.93% ($\pm 3.14\%$) artifacts per R–R interval recording. R–R interval time series with $\geq 80\%$ non-artifact R–R intervals were used for further analyses (Peltola, 2012), while time series below this value (and the associated HRV indices) were replaced with the group mean value (Liu & Wang, 2021).

2.6 Statistical analyses

An a priori power analysis was conducted using G*Power (version 3.1) for a 2 (group: exercise vs. psychoeducation) $\times$ 2 (time: pre-test vs. post-test) repeated-measures ANOVA (within–between interaction). Based on prior studies of exercise interventions suggesting small-to-moderate effects on cue-induced craving reactivity (Liu & Wang, 2023; Peng, Chen, Su, Tao, & Wang, 2021; Shen, Du, Fan, Dai, & Wei, 2023; Wang, Zhu, Zhou, & Chang, 2017), we assumed a conservative effect

size (Cohen's $f = 0.25$), with $\alpha = .05$ and power $(1 - \beta) = .90$. The analysis indicated that a minimum of 23 participants per group was required to detect the Group $\times$ Time interaction. The current study exceeded this requirement (exercise group: $n = 36$; psychoeducation group: $n = 35$).

All data were analyzed using SPSS 20.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were first computed for participant demographic characteristics, drug-use history, and HRV_{baseline}. Data were expressed as means and standard deviations (SDs). The independent samples t -tests or chi-square tests were conducted to compare differences between the two groups in demographic characteristics, drug use history, and HRV_{baseline} at baseline. A repeated-measures ANOVA was used to examine variable changes between the pre- and post-tests. Group was the between-subjects factor, and time point was the within-subjects factor. Further post-hoc multiple comparisons were conducted using the Bonferroni correction.

Pearson correlation was performed to analyze the relationship between changes (post-test – pre-test) for IS, METH-craving, METH-using, HRV_{cue}, and relapse tendency. Furthermore, mediation analysis was conducted using the PROCESS 3.5 macro developed by Hayes (2013). The significant indirect effect was examined using bias-corrected confidence intervals derived from 5,000 bootstrap samples (Shen et al., 2023). An effect was considered significant if its 95% bootstrap confidence interval (CI) did not include zero (Shen et al., 2023). All statistical tests were two-tailed. Statistical significance was indicated as $p < .05$. Effect sizes (η_p^2) were also reported.

3. Results

3.1 Baseline characteristics

No significant differences between groups were found at baseline in demographic characteristics, drug-use history, and HRV_{baseline} (Table 1).

Table 1. Baseline characteristics

Variable	Exercise group ($n=36$)	Psychoeducation group ($n=35$)
Age (years)	35.31 $\pm$ 3.12	34.89 $\pm$ 4.60
Education (years)		
Elementary school	1 (2.78%)	1 (2.86%)
Junior high school	4 (11.11%)	9 (25.71%)

Variable	Exercise group (n=36)	Psychoeducation group (n=35)
Senior high school	27 (75.00%)	23 (65.71%)
College	4 (11.11%)	2 (5.71%)
Height (cm)	171.39 ± 4.99	170.03 ± 5.28
Weight (kg)	72.14 ± 7.57	70.86 ± 6.32
BMI (kg/m ²)	24.55 ± 2.22	24.50 ± 1.80
HR _{rest} (bpm)	73.81 ± 8.42	74.40 ± 7.41
Drug-use history		
Duration of drug use (years)	13.64 ± 3.31	13.81 ± 2.67
Frequency (times/week)	6.03 ± 1.86	5.66 ± 1.75
Total amount of METH used (g)	2394.93 ± 612.78	2381.39 ± 464.99
Withdrawal duration (months)	8.56 ± 3.12	9.03 ± 2.88
HRV _{baseline}		
SDNN	41.98 ± 0.36	42.85 ± 0.27
RMSSD	31.68 ± 0.61	31.57 ± 0.59
pNN50	13.59 ± 0.39	13.85 ± 0.44
nLF	0.62 ± 0.03	0.61 ± 0.02
nHF	0.26 ± 0.03	0.27 ± 0.03
LF/HF ratio	4.00 ± 0.65	3.90 ± 0.57

Note. Values are mean ± standard deviation. BMI, body mass index; bpm, beats per minute; HR_{rest}, resting heart rate.

3.2 Changes in indices before and after intervention

No significant differences between groups were found at the pre-test for all outcome measure scores. Repeated-measures ANOVA indicated significant interactions between time and group in METH-craving [$F_{(1,69)} = 17.204, p < .001, \eta_p^2 = 0.200$], IS [$F_{(1,69)} = 12.935, p = .001, \eta_p^2 = 0.158$], and relapse tendency [$F_{(1,69)} = 8.027, p = .006, \eta_p^2 = 0.104$] and significant main effects of group in METH-craving [$F_{(1,69)} = 12.167, p = .001, \eta_p^2 = 0.150$] and IS [$F_{(1,69)} = 4.866, p = .031, \eta_p^2 = 0.066$], but no significant main effect of group in relapse tendency. The exercise group exhibited significantly lower scores on IS, METH-craving, and relapse tendency than the psychoeducation group in the post-test (Figure 3). For the VAS score on METH-using, there was a marginally significant interaction between time and group [$F_{(1,69)} = 3.893, p = .052, \eta_p^2 = 0.053$], and a significant main effect of time [$F_{(1,69)} = 18.051, p < .001, \eta_p^2 = 0.289$], reflected by a significant decrease from pre- to post-test in both groups. No significant main effect of group was found in METH-using.

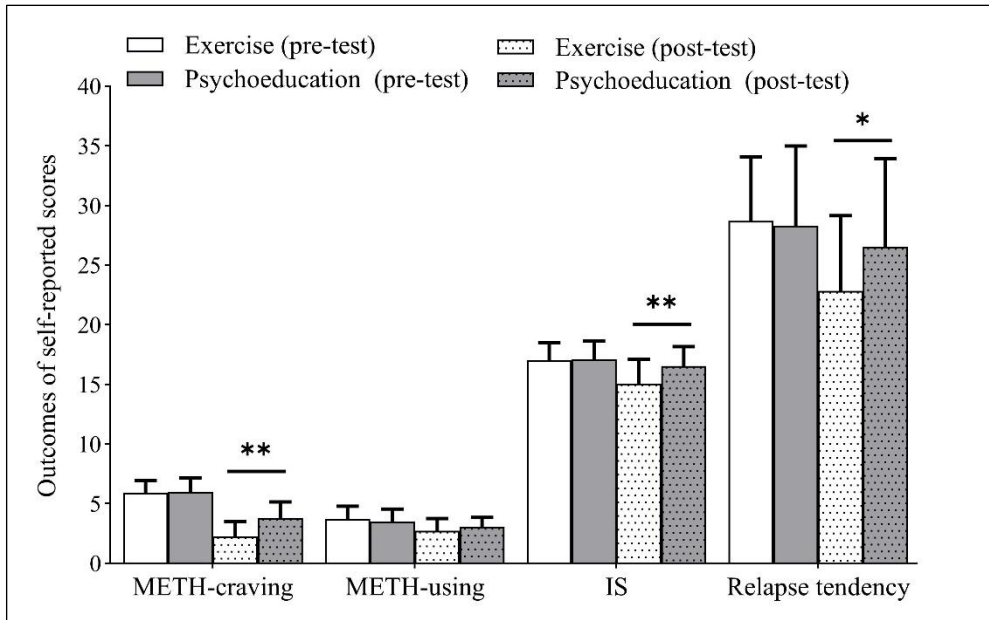


Figure 3. Outcomes of self-reported scores (* indicates $p < .05$, ** indicates $p < .01$)

For HRV_{cue} parameters, there were significant interactions between time and group in SDNN [$F_{(1,69)} = 6.470$, $p = .013$, $\eta_p^2 = 0.086$], RMSSD [$F_{(1,69)} = 11.603$, $p = .001$, $\eta_p^2 = 0.144$], and pNN50 [$F_{(1,69)} = 6.969$, $p = .01$, $\eta_p^2 = 0.092$] and significant main effects of group in SDNN [$F_{(1,69)} = 6.991$, $p = .010$, $\eta_p^2 = 0.092$], RMSSD [$F_{(1,69)} = 17.154$, $p < .001$, $\eta_p^2 = 0.199$], and pNN50 [$F_{(1,69)} = 16.618$, $p < .001$, $\eta_p^2 = 0.194$]. The exercise group exhibited a significantly larger decrease in time-domain HRV_{cue} indices from pre- to post-test compared to the psychoeducation group (Table 2). There were significant main effects of time in nLF [$F_{(1,69)} = 5.507$, $p = .022$, $\eta_p^2 = 0.074$], nHF [$F_{(1,69)} = 19.192$, $p < .001$, $\eta_p^2 = 0.218$], and LF/HF ratios [$F_{(1,69)} = 6.152$, $p = .016$, $\eta_p^2 = 0.082$], reflecting a significant change from pre- to post-test in both groups. No significant main effect of group and no significant interaction between group and time were found in the frequency-domain HRV_{cue} indices.

Table 2. Parameters of HRV in different states at pre-test and post-test for the two groups

Group	Pre-test			HRV _{cue} $P_{between}$	Post-test			HRV _{cue} $P_{between}$
	Before video	During video	Cue-induced		Before video	During video	Cue-induced	
SDNN	Exercise	41.98 ± 0.84	44.96 ± 0.90	2.98 ± 0.19	55.86 ± 0.73	49.86 ± 0.61	-6.00 ± 0.31	< .001
	Psychoeducation	42.86 ± 0.83	45.84 ± 0.76	2.98 ± 0.25	43.38 ± 0.84	39.87 ± 0.73	-3.52 ± 0.24	

Group		Pre-test			HRV _{cue} <i>P</i> _{between}	Post-test			HRV _{cue} <i>P</i> _{between}
		Before video	During video	Cue- induced		Before video	During video	Cue- induced	
RMSSD	Exercise	31.68 ± 0.62	38.32 ± 0.65	6.64 ± 0.10	NS	49.68 ± 0.79	46.34 ± 0.80	-3.34 ± 0.23	< .001
		31.57 ± 0.52	38.43 ± 0.50	6.85 ± 0.18		34.48 ± 0.60	34.61 ± 0.69	0.13 ± 0.26	
	Psychoeducation	13.59 ± 0.39	17.56 ± 0.43	3.97 ± 0.16		34.25 ± 0.61	32.38 ± 0.58	-1.87 ± 0.24	
		13.86 ± 0.45	18.33 ± 0.42	4.47 ± 0.15		18.24 ± 0.91	18.54 ± 0.80	0.29 ± 0.22	
pNN50	Exercise	0.60 ± 0.08	0.56 ± 0.10	-0.04 ± 0.03	NS	0.52 ± 0.10	0.49 ± 0.10	-0.03 ± 0.05	NS
		0.61 ± 0.09	0.56 ± 0.09	-0.05 ± 0.01		0.58 ± 0.09	0.55 ± 0.09	-0.03 ± 0.02	
	Psychoeducation	0.28 ± 0.10	0.30 ± 0.10	0.02 ± 0.03		0.38 ± 0.10	0.38 ± 0.10	0.00 ± 0.05	
		0.27 ± 0.09	0.30 ± 0.09	0.03 ± 0.02		0.32 ± 0.09	0.32 ± 0.09	0.00 ± 0.02	
nLF	Exercise	4.01 ± 0.16	3.83 ± 0.14	-0.17 ± 0.04	NS	3.40 ± 0.17	3.36 ± 0.15	-0.05 ± 0.35	NS
		3.97 ± 0.15	3.75 ± 0.12	-0.22 ± 0.05		3.97 ± 0.10	3.93 ± 0.11	-0.04 ± 0.28	
	Psychoeducation								
nHF	Exercise				NS				NS
	Psychoeducation								
LF/HF ratio	Exercise				NS				NS
	Psychoeducation								

Note. Values are mean ± standard deviation. LF, low-frequency component; LF/HF, low-frequency to high-frequency ratio; nHF, high-frequency component; pNN50, percentage of beats that changed more than 50 ms from the previous beat; RMSSD, root mean square differences of the standard deviation; SDNN, standard deviation of normal-to-normal intervals; NS, not significant.

3.3 Correlation analyses

METH-craving change correlated with changes in SDNN ($r = 0.576$, $p < .001$) and RMSSD ($r = 0.526$, $p = .001$), indicating that greater decreases in subjective drug craving in the METH-related virtual social environment were associated with greater decreases in physiological responses after 8 weeks of aerobic exercise intervention in the exercise group. In addition, changes in IS were positively correlated with changes in METH-craving ($r = 0.453$, $p = .005$), SDNN ($r = 0.439$, $p = .007$), and RMSSD ($r = 0.343$, $p = .041$). The results indicated that greater decreases in IS were associated with larger declines in METH-related cue-induced reactivity. Moreover, greater decreases in relapse tendency correlated with greater declines in IS ($r = 0.539$, $p = .001$), METH-craving ($r = 0.580$, $p < .001$), and time-domain HRV_{cue} reactivity ($0.332 < r < 0.594$, $p < .005$).

3.4 Mediation analyses

Mediation effect analyses showed that changes in METH-craving ($\beta = 2.009$, 95% CI: 0.862–3.736), SDNN ($\beta = 1.505$, 95% CI: 0.378–3.256), and IS ($\beta =$

1.580, 95% CI: 0.380–3.551) mediated the relationship between the physical exercise intervention and changes in relapse tendency.

4. Discussion

This study compared the effects of physical exercise versus psychoeducation on METH-cue reactivity, IS, and relapse tendency in the VR cue-exposure environment among individuals with MUD. We further examined the role of METH-cue craving reactivity and IS in physical exercise and relapse tendency. As predicted, compared with the psychoeducation intervention, the 8-week aerobic exercise intervention produced greater reductions in craving reactivity (i.e., METH-craving and time-domain HRV_{cue}), IS, and relapse tendency from the pre-test to the post-test. METH-cue craving reactivity and IS mediated the relationship between physical exercise and relapse tendency. This suggests that the decreased METH-cue craving reactivity and IS in the VR cue-exposure environment might contribute to the effects of physical exercise on relapse tendency. To our knowledge, this is the first study to provide evidence that IS may be one of the mechanisms underlying the effect of physical exercise on the risk of relapse.

It is well-known that one main challenge in treating MUD is the high rate of relapse during abstinence (Reiner & Shaham, 2022). Cue-induced craving is widely regarded as a primary target in relapse prevention (Wang et al., 2019). The METH-VR model serves as a powerful clinical and research tool for manipulating acute craving in METH abusers (Culbertson et al., 2010). Studies have shown that craving has different forms, including mental, psychophysiological, and behavioral (Chen et al., 2024; Sayette, 2016; Tan et al., 2019). To comprehensively assess the effects of physical exercise on the improvement of acute craving induced by a METH-related VR social environment, we measured different forms of METH-cue craving reactivity. Our results showed that compared with the psychoeducation group, the VAS score for METH-craving in the exercise group significantly decreased after 8 weeks of aerobic exercise intervention. This suggests that physical exercise may effectively attenuate subjective craving in individuals with MUD. The METH-craving results obtained from METH-related VR cue exposure in this study agree with those reported using traditional exposure methods in previous studies (Huang, Zheng, Gao, Hu, & Yuan, 2019; Liu & Wang, 2023; Wang et al., 2017). However, we did not find a significant beneficial effect of physical exercise on the VAS score of METH-using. This may be attributed to the strict restrictions on drug use in the detoxification environment, which make it impossible to obtain drugs during the confinement period, resulting in low self-reported pre- and post-test scores for both groups. Additionally, a floor effect for VAS scores of METH-using may, in part, be

responsible for the relatively small differences between groups, as speculated by previous studies (Wang et al., 2019).

The HRV_{cue} results showed that, compared with participants receiving the psychoeducation intervention, those receiving the aerobic exercise intervention exhibited a larger decrease in time-domain HRV_{cue} indices when exposed to VR cues. This suggests that the exercise group had a lower level of physiological arousal when exposed to METH-related VR cues. These results are consistent with the observations of previous studies in the resting state (Cabral et al., 2018; Dolezal et al., 2014; Liu & Wang, 2021), namely, that aerobic exercise contributes to the homeostasis of the autonomic nervous system in individuals with MUD. Moreover, changes in time-domain HRV_{cue} and METH-craving were positively correlated. As different manifestations of METH-cue reactivity, the reduction in METH-craving and HRV_{cue} components may be the mental and psychophysiological representation of decreased cue-related craving. This further supports that an 8-week aerobic exercise intervention may be a useful strategy in suppressing METH-cue reactivity in a VR cue-exposure environment in individuals with MUD.

In addition, reduced subjective craving and autonomic arousal upon exposure to substance-related cues help individuals with SUD experience fewer subjectively abnormal bodily sensations, which may ultimately reduce urges and the risk of relapse (Sharpe, Tarrier, Schotte, & Spence, 1995). In the present study, our findings revealed significant reductions in relapse tendency after physical exercise, which is consistent with the findings of previous studies (Lynch et al., 2013; Morais et al., 2018; Song et al., 2023). We also found that the change in METH-cue reactivity (i.e., METH-craving and SDNN_{cue}) mediated the relationship between physical exercise and decreased relapse tendency. This suggests that reducing METH-cue craving reactivity might contribute to the effects of physical exercise in preventing relapse in individuals with MUD, which is also consistent with previous research (Song et al., 2023). However, no significant improvement was found in the frequency-domain HRV_{cue} parameters in association with physical exercise. Several physiological studies have shown that many physiological parameters (e.g., respiratory rhythm) can influence frequency-domain HRV indices, and these confounders might lead to large individual differences (Billman, 2013; Billman, Huikuri, Sacha, & Trimmel, 2015; Shaffer & Ginsberg, 2017; Wang et al., 2019). We speculate that this may have masked the beneficial effects of exercise interventions on frequency-domain HRV_{cue}.

A novel finding of this study was the beneficial effect of physical exercise on IS and the significant positive correlation between the altered IS scores and the change in METH-related cue-induced reactivity resulting from the exercise intervention. This is consistent with prior work demonstrating that body-oriented therapy may help decrease craving and improve interoception (Price et al., 2019a, 2019b). More intriguingly, we found that IS mediated the relationship between aerobic exercise intervention and relapse tendency, suggesting that the improved IS

in METH-related cue exposure may be one of the potential pathways underlying the effect of physical exercise on the risk of relapse. Several studies have shown that physical exercise may help people with SUD better decode and interpret bodily-related signals associated with transient homeostatic imbalances that usually trigger consumption (Brevers et al., 2024). For example, the inflow of active inference mechanisms triggered by exercise should allow one to strategically regulate the pace and intensity of effort throughout the exercise session, which may lead to better awareness of stimuli arising within the body and help maintain bodily homeostasis (Brevers et al., 2024; Osypiuk et al., 2020).

Moreover, by training individuals to formulate prospective, momentary, and retrospective judgments toward physical states, exercise would allow individuals with SUD to explore and resolve homeostatic imbalances while adopting healthy behaviors (Brevers et al., 2024). Exercise may also help individuals discriminate between sensations triggered by physical exercise and those triggered by reward-seeking and consumption (Brevers et al., 2024). Vicarious reward processing might be linked to people's sensitivity to internal body states and could facilitate a tendency to act pro-socially (Herman, 2023). This preliminary study supports the notion that individuals with MUD may benefit from interventions focused on redirecting interoceptive input toward competitive natural reinforcers like physical exercise (Jakubczyk et al., 2019a).

The present study is the first to find that physical exercise may decrease relapse tendency in individuals with MUD through improved IS, indicating that IS may be one of the potential mechanisms underlying the reduction in relapse tendency induced by physical exercise. Previous research has revealed that physical activity can affect the relapse tendency of individuals with MUD undergoing compulsory detoxification not only through a direct pathway but also through indirect pathways, such as the mediating roles of intrinsic restraint and yearning for drugs (Lynch et al., 2013; Morais et al., 2018; Ouyang et al., 2020; Song et al., 2023); the current study expanded on these findings.

In addition, regarding the efficacy of physical exercise in improving IS and cue-induced reactivity, previous studies provide some evidence at the neural mechanism level. Research suggests that the insular cortex is the central hub for processing interoception, and insular dysfunction has been reported in several addictive disorders (Bonson et al., 2002; Engelmann et al., 2012; Goudriaan, de Ruiter, van den Brink, Oosterlaan, & Veltman, 2010; Paulus & Stewart, 2014; Prisciandaro, Myrick, Henderson, McRae-Clark, & Brady, 2013). An animal study demonstrating the causal role of the insula in nicotine-induced cue approach behavior (Scott & Hiroi, 2011) supports the central role of this structure in IS and cue reactivity processing. Intense physical exercise has been shown to cause changes in insular activity and potentially decrease the heightened internal response to stimuli related to drug use (Herman, 2023; Paulus, Stewart, & Haase, 2013). Therefore, in this preliminary study, we hypothesize that insular modulation may be the

underlying neural mechanism by which physical exercise improves IS and cue-induced reactivity, leading to a greater reduction in relapse tendency in individuals with MUD. However, we did not provide any evidence regarding the neural mechanisms by which physical exercise improves IS and cue-induced reactivity; instead, we offered directions for future research. More interventional trials are needed to test this hypothesis.

The current study has several strengths. It has provided evidence supporting the notion that physical exercise may be a useful strategy for suppressing VR cue-induced craving reactivity, IS, and relapse tendency in individuals with MUD, with physical exercise showing superior results compared to psychoeducation. In addition to craving reactivity, this study provided new evidence that IS may also be one of the underlying mechanisms by which physical exercise reduces relapse tendency. Second, this preliminary study aimed to understand the interoceptive mechanisms driving drug relapse and may lead to an enhanced understanding of relapse mechanisms, as well as to new treatment strategies for METH addiction. The relationship between interoception and addiction may apply not only to substance dependence but also to behavioral addictions (Herman, 2023). This study also has important implications for the clinical treatment of drug addiction, behavioral addictions, psychiatric disorders, and non-pharmacological therapy. Furthermore, we used one of the most ecologically valid drug-related cue-induced methods, which can more reliably detect subjective cravings and physiological reactivity and may have a greater impact on the participants' perceptions of their own bodily signals.

It should be noted that, although positive results have been found, these findings should be interpreted and applied with caution. As IS assessment is based on subjective self-reported acute responses, further studies are needed to develop new methods for studying how individuals with addiction who receive exercise interventions perceive the process and perceive internal bodily sensations, such as utilizing structural and functional neuroimaging to explore relevant neural mechanisms. Second, interoception consists of interoceptive accuracy, IS, and interoceptive awareness (Garfinkel & Critchley, 2013). Whether interoceptive accuracy and awareness alterations are associated with physical exercise-induced changes in METH-cue reactivity and relapse tendency remains unknown. Third, due to the repetitive nature of measuring acute craving and IS, scales with more items might be optimal and could provide more useful information (Culbertson et al., 2010; Wang et al., 2019). Fourth, the relatively small sample size used in this study may have been insufficient to detect small effects, particularly for frequency-domain HRV indices that are sensitive to physiological influences; therefore, the results should be interpreted with caution. Future studies with larger samples, improved psychoeducation, expanded measurement of key physiological confounders (e.g., respiration), and designs that minimize floor effects in behavioral outcomes are needed to test the robustness of these findings. Finally, there may be gender differences between interoception (Prentice & Murphy, 2022); however, our sample

was composed exclusively of males, which may limit our ability to generalize the results of the current study to females with MUD. Future studies should include both male and female participants to enhance comparability and generalizability. Overall, evidence regarding the relationship between interoception and drug relapse and the effects of exercise interventions on IS remains limited. Therefore, more research aimed at understanding interoceptive mechanisms driving drug relapse is needed, which may lead to a better understanding of the mechanisms of relapse and new treatment strategies for drug addiction.

5. Conclusion

Physical exercise may improve METH-cue craving reactivity, IS, and relapse tendency in association with VR cue exposure in individuals with MUD, yielding greater improvements than psychoeducation. In the current study, reduced relapse tendency attributable to physical exercise was positively correlated with changes in METH-cue craving reactivity and IS. In addition to drug craving, IS may have a mediating effect on the relationship between physical exercise and relapse tendency, indicating that IS may be one of the potential mechanisms underlying the decrease in relapse tendency induced by physical exercise.

Authors' Notes

Acknowledgments: The author would like to thank all participants, exercise specialists, cognitive psychologists, and health educators involved in this study.

Funding: This work was supported by the Tongling University Research Project Program [Grant No. 2025tlxykj05] and the Talent Research Start-up Fund Project of Tongling University [Grant No. 2025tlxyrc147].

Data availability: Deidentified data are available upon reasonable request to the corresponding author.

Ethics approval: The ethics committee of the Tongling University Institutional Review Board approved this study.

Informed consent to participate: Informed consent was obtained from all patients for inclusion in the study.

Conflict of interest: I hereby declare that there are no potential conflicts of interest associated with this publication, and that any financial support has been noted in the Acknowledgment section.

References

- AshaRani, P. V., Hombali, A., Seow, E., Ong, W. J., Tan, J. H., & Subramaniam, M. (2020). Non-pharmacological interventions for methamphetamine use disorder: A systematic review. *Drug and Alcohol Dependence*, 212, Article 108060. <https://doi.org/10.1016/j.drugalcdep.2020.108060>
- Avants, S. K., Margolin, A., Kosten, T. R., & Cooney, N. L. (1995). Differences between responders and nonresponders to cocaine cues in the laboratory. *Addictive Behaviors*, 20(2), 215–224. [https://doi.org/10.1016/0306-4603\(94\)00066-2](https://doi.org/10.1016/0306-4603(94)00066-2)
- Billman, G. E. (2013). The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Frontiers in Physiology*, 4, Article 26. <https://doi.org/10.3389/fphys.2013.00026>
- Billman, G. E., Huikuri, H. V., Sacha, J., & Trimmel, K. (2015). An introduction to heart rate variability: Methodological considerations and clinical applications. *Frontiers in Physiology*, 6, 55. <https://doi.org/10.3389/fphys.2015.00055>
- Bonaz, B., Lane, R. D., Oshinsky, M. L., Kenny, P. J., Sinha, R., Mayer, E. A., & Critchley, H. D. (2021). Diseases, disorders, and comorbidities of interoception. *Trends in Neurosciences*, 44(1), 39–51. <https://doi.org/10.1016/j.tins.2020.09.009>
- Bonson, K. R., Grant, S. J., Contoreggi, C. S., Links, J. M., Metcalfe, J., Weyl, H. L., Kurian, V., Ernst, M., & London, E. D. (2002). Neural systems and cue-induced cocaine craving. *Neuropsychopharmacology*, 26(3), 376–386. [https://doi.org/10.1016/S0893-133X\(01\)00371-2](https://doi.org/10.1016/S0893-133X(01)00371-2)
- Brevers, D., Billieux, J., de Timary, P., Desmedt, O., Maurage, P., Perales, J. C., Suárez-Suárez, S., & Bechara, A. (2024). Physical exercise to redynamize interoception in substance use disorders. *Current Neuropharmacology*, 22(6), 1047–1063. <https://doi.org/10.2174/1570159X21666230314143803>
- Cabral, D. A. R., Tavares, V. D. O., da Costa, K. G., Nascimento, P. H. D., Faro, H. K. C., Elsangedy, H. M., & Fontes, E. B. (2018). The benefits of high intensity exercise on the brain of a drug abuser. *Global Journal of Health Science*, 10(6), 123–135. <https://doi.org/10.5539/gjhs.v10n6p123>
- Chan, B., Freeman, M., Kondo, K., Ayers, C., Montgomery, J., Paynter, R., & Kansagara, D. (2019). Pharmacotherapy for methamphetamine/amphetamine use disorder-a systematic review and meta-analysis. *Addiction*, 114(12), 2122–2136. <https://doi.org/10.1111/add.14755>
- Chen, C.-C., Chung, C.-R., Tsai, M.-C., Wu, E. H.-K., Chiu, P.-R., Tsai, P.-Y., & Yeh, S.-C. (2024). Impaired brain-heart relation in patients with methamphetamine use disorder during VR induction of drug cue reactivity. *IEEE Journal of Translational Engineering in Health and Medicine*, 12, 1–9. <https://doi.org/10.1109/JTEHM.2022.3206333>
- Culbertson, C., Nicolas, S., Zaharovits, I., London, E. D., De La Garza, R. II, Brody, A. L., & Newton, T. F. (2010). Methamphetamine craving induced in an online

- virtual reality environment. *Pharmacology Biochemistry and Behavior*, 96(4), 454–460. <https://doi.org/10.1016/j.pbb.2010.07.005>
- Dolezal, B. A., Chudzynski, J., Dickerson, D., Mooney, L., Rawson, R. A., Garfinkel, A., & Cooper, C. B. (2014). Exercise training improves heart rate variability after methamphetamine dependency. *Medicine & Science in Sports & Exercise*, 46(6), 1057–1066. <https://doi.org/10.1249/MSS.0000000000000201>
- Donovan, D. M., & Wells, E. A. (2007). ‘Tweaking 12-Step’: The potential role of 12-Step self-help group involvement in methamphetamine recovery. *Addiction*, 102(Suppl 1), 121–129. <https://doi.org/10.1111/j.1360-0443.2007.01773.x>
- Dudish-Poulsen, S. A., & Hatsukami, D. K. (1997). Dissociation between subjective and behavioral responses after cocaine stimuli presentations. *Drug and Alcohol Dependence*, 47(1), 1–9. [https://doi.org/10.1016/S0376-8716\(97\)00054-9](https://doi.org/10.1016/S0376-8716(97)00054-9)
- Eggenberger, P., Annaheim, S., Kündig, K. A., Rossi, R. M., Münzer, T., & de Bruin, E. D. (2020). Heart rate variability mainly relates to cognitive executive functions and improves through exergame training in older adults: A secondary analysis of a 6-month randomized controlled trial. *Frontiers in Aging Neuroscience*, 12, Article 197. <https://doi.org/10.3389/fnagi.2020.00197>
- Engelmann, J. M., Versace, F., Robinson, J. D., Minnix, J. A., Lam, C. Y., Cui, Y., Brown, V. L., & Cinciripini, P. M. (2012). Neural substrates of smoking cue reactivity: A meta-analysis of fMRI studies. *NeuroImage*, 60(1), 252–262. <https://doi.org/10.1016/j.neuroimage.2011.12.024>
- Gamelin, F. X., Berthoin, S., Sayah, H., Libersa, C., & Bosquet, L. (2007). Effect of training and detraining on heart rate variability in healthy young men. *International Journal of Sports Medicine*, 28(7), 564–570. <https://doi.org/10.1055/s-2007-964861>
- Garfinkel, S. N., & Critchley, H. D. (2013). Interoception, emotion and brain: New insights link internal physiology to social behaviour. Commentary on: “Anterior insular cortex mediates bodily sensibility and social anxiety” by Terasawa et al. (2012). *Social Cognitive and Affective Neuroscience*, 8(3), 231–234. <https://doi.org/10.1093/scan/nss140>
- Goudriaan, A. E., de Ruiter, M. B., van den Brink, W., Oosterlaan, J., & Veltman, D. J. (2010). Brain activation patterns associated with cue reactivity and craving in abstinent problem gamblers, heavy smokers and healthy controls: An fMRI study. *Addiction Biology*, 15(4), 491–503. <https://doi.org/10.1111/j.1369-1600.2010.00242.x>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Herman, A. M. (2023). Interoception within the context of impulsivity and addiction. *Current Addiction Reports*, 10(2), 97–106. <https://doi.org/10.1007/s40429-023-00482-7>
- Hormes, J. M. (2017). The clinical significance of craving across the addictive behaviors: A review. *Current Addiction Reports*, 4(2), 132–141. <https://doi.org/10.1007/s40429-017-0138-y>

- Huang, J., Zheng, Y., Gao, D., Hu, M., & Yuan, T. (2019). Effects of exercise on depression, anxiety, cognitive control, craving, physical fitness and quality of life in methamphetamine-dependent patients. *Frontiers in Psychiatry*, 10, Article 999. <https://doi.org/10.3389/fpsyt.2019.00999>
- Jakubczyk, A., Skrzyszewski, J., Trucco, E. M., Suszek, H., Zaorska, J., Nowakowska, M., Michalska, A., Wojnar, M., & Kopera, M. (2019). Interoceptive accuracy and interoceptive sensibility in individuals with alcohol use disorder: Different phenomena with different clinical correlations? *Drug and Alcohol Dependence*, 198, 34–38. <https://doi.org/10.1016/j.drugalcdep.2019.01.036>
- Jakubczyk, A., Trucco, E. M., Klimkiewicz, A., Skrzyszewski, J., Suszek, H., Zaorska, J., Nowakowska, M., Michalska, A., Wojnar, M., & Kopera, M. (2019). Association between interoception and emotion regulation in individuals with alcohol use disorder. *Frontiers in Psychiatry*, 10, Article 1028. <https://doi.org/10.3389/fpsyt.2019.01028>
- Lee, N. K., Jenner, L., Harney, A., & Cameron, J. (2018). Pharmacotherapy for amphetamine dependence: A systematic review. *Drug and Alcohol Dependence*, 191, 309–337. <https://doi.org/10.1016/j.drugalcdep.2018.06.038>
- Liu, X.-X., & Wang, S. (2021). Effect of aerobic exercise on executive function in individuals with methamphetamine use disorder: Modulation by the autonomic nervous system. *Psychiatry Research*, 306, Article 114241. <https://doi.org/10.1016/j.psychres.2021.114241>
- Liu, X.-X., & Wang, S. (2023). Effects of aerobic exercise combined with attentional bias training on cognitive function and psychiatric symptoms of individuals with methamphetamine dependency: A randomized controlled trial. *International Journal of Mental Health and Addiction*, 21, 1727–1745. <https://doi.org/10.1007/s11469-021-00686-w>
- Lynch, W. J., Peterson, A. B., Sanchez, V., Abel, J., & Smith, M. A. (2013). Exercise as a novel treatment for drug addiction: A neurobiological and stage-dependent hypothesis. *Neuroscience & Biobehavioral Reviews*, 37(8), 1622–1644. <https://doi.org/10.1016/j.neubiorev.2013.06.011>
- McKetin, R., Najman, J. M., Baker, A. L., Lubman, D. I., Dawe, S., Ali, R., Lee, N. K., Mattick, R. P., & Mamun, A. (2012). Evaluating the impact of community-based treatment options on methamphetamine use: Findings from the Methamphetamine Treatment Evaluation Study (MATES). *Addiction*, 107(11), 1998–2008. [HTTPS://DOI.ORG/10.1111/J.1360-0443.2012.03933.X](https://doi.org/10.1111/J.1360-0443.2012.03933.X)
- Mehling, W. E., Gopisetty, V., Daubenmier, J., Price, C. J., Hecht, F. M., & Stewart, A. (2009). Body awareness: Construct and self-report measures. *PLOS ONE*, 4(5), Article e5614. <https://doi.org/10.1371/journal.pone.0005614>
- Miller, L. C., Murphy, R., & Buss, A. H. (1981). Consciousness of body: Private and public. *Journal of Personality and Social Psychology*, 41(2), 397–406. <https://doi.org/10.1037/0022-3514.41.2.397>
- Morais, A. P. D., Pita, I. R., Fontes-Ribeiro, C. A., & Pereira, F. C. (2018). The neurobiological mechanisms of physical exercise in methamphetamine

- addiction. *CNS Neuroscience & Therapeutics*, 24(2), 85–97. <https://doi.org/10.1111/cns.12788>
- Mottola, C. A. (1993). Measurement strategies: The visual analogue scale. *Decubitus*, 6(5), 56–58.
- Ooteman, W., Koeter, M. W. J., Verheul, R., Schippers, G. M., & van den Brink, W. (2006). Measuring craving: An attempt to connect subjective craving with cue reactivity. *Alcoholism: Clinical and Experimental Research*, 30(1), 57–69. <https://doi.org/10.1111/j.1530-0277.2006.00019.x>
- Osypiuk, K., Ligibel, J., Giobbie-Hurder, A., Vergara-Diaz, G., Bonato, P., Quinn, R., Ng, W., & Wayne, P. M. (2020). Qigong mind-body exercise as a biopsychosocial therapy for persistent post-surgical pain in breast cancer: A pilot study. *Integrative Cancer Therapies*, 19, Article 1534735419893766. <https://doi.org/10.1177/1534735419893766>
- Ouyang, Y., Luo, J., Wang, K., Zhang, T., Zhou, C., & Lu, Y. (2020). Impact of physical activity on relapse tendency of AIDS patients undergoing compulsory detoxification: A chain mediation analysis of intrinsic restraint and yearning for drugs. *Journal of Shanghai University of Sport*, 44(10), 48–57. <https://doi.org/10.16099/j.sus.2020.10.006>
- Paulus, M. P., & Stewart, J. L. (2014). Interoception and drug addiction. *Neuropharmacology*, 76 (Part B), 342–350. <https://doi.org/10.1016/j.neuropharm.2013.07.002>
- Paulus, M. P., Stewart, J. L., & Haase, L. (2013). Treatment approaches for interoceptive dysfunctions in drug addiction. *Frontiers in Psychiatry*, 4, Article 137. <https://doi.org/10.3389/fpsy.2013.00137>
- Paulus, M. P., Tapert, S. F., & Schulteis, G. (2009). The role of interoception and alliesthesia in addiction. *Pharmacology Biochemistry and Behavior*, 94(1), 1–7. <https://doi.org/10.1016/j.pbb.2009.08.005>
- Peltola, M. A. (2012). Role of editing of R-R intervals in the analysis of heart rate variability. *Frontiers in Physiology*, 3, Article 148. <https://doi.org/10.3389/fphys.2012.00148>
- Peng, B., Chen, J., Su, R., Tao, W., & Wang, D. (2021). Effects of resistance exercise on subjective craving and inducing factors of methamphetamine-dependent individuals. *Chinese Journal of Sports Medicine*, 40(1), 46–54. <https://doi.org/10.16038/j.1000-6710.2021.01.009>
- Prentice, F., & Murphy, J. (2022). Sex differences in interoceptive accuracy: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, 132, 497–518. <https://doi.org/10.1016/j.neubiorev.2021.11.030>
- Price, C. J., Thompson, E. A., Crowell, S., & Pike, K. (2019a). Longitudinal effects of interoceptive awareness training through mindful awareness in body-oriented therapy (MABT) as an adjunct to women's substance use disorder treatment: A randomized controlled trial. *Drug and Alcohol Dependence*, 198, 140–149. <https://doi.org/10.1016/j.drugalcdep.2019.02.012>
- Price, C. J., Thompson, E. A., Crowell, S. E., Pike, K., Cheng, S. C., Parent, S., & Hooven, C. (2019b). Immediate effects of interoceptive awareness training

- through mindful awareness in body-oriented therapy (MABT) for women in substance use disorder treatment. *Substance Abuse*, 40(1), 102–115. <https://doi.org/10.1080/08897077.2018.1488335>
- Prisciandaro, J. J., Myrick, H., Henderson, S., McRae-Clark, A. L., & Brady, K. T. (2013). Prospective associations between brain activation to cocaine and no-go cues and cocaine relapse. *Drug and Alcohol Dependence*, 131(1–2), 44–49. <https://doi.org/10.1016/j.drugalcdep.2013.04.008>
- Rawson, R. A., Chudzynski, J., Mooney, L., Gonzales, R., Ang, A., Dickerson, D., Penate, J., Salem, B. A., Dolezal, B., & Cooper, C. B. (2015). Impact of an exercise intervention on methamphetamine use outcomes post-residential treatment care. *Drug and Alcohol Dependence*, 156, 21–28. <https://doi.org/10.1016/j.drugalcdep.2015.08.029>
- Reiner, D. J., & Shaham, Y. (2022). Recent findings on the role of sensory and interoceptive-related circuits in relapse to drug seeking. *Neuropsychopharmacology*, 47(2), 373–374. <https://doi.org/10.1038/s41386-021-01081-8>
- Sayette, M. A. (2016). The role of craving in substance use disorders: Theoretical and methodological issues. *Annual Review of Clinical Psychology*, 12, 407–433. <https://doi.org/10.1146/annurev-clinpsy-021815-093351>
- Scott, D., & Hiroi, N. (2011). Deconstructing craving: Dissociable cortical control of cue reactivity in nicotine addiction. *Biological Psychiatry*, 69(11), 1052–1059. <https://doi.org/10.1016/j.biopsych.2011.01.023>
- Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, 5, Article 258. <https://doi.org/10.3389/fpubh.2017.00258>
- Sharpe, L., Tarrier, N., Schotte, D., & Spence, S. H. (1995). The role of autonomic arousal in problem gambling. *Addiction*, 90(11), 1529–1540. <https://doi.org/10.1046/j.1360-0443.1995.9011152911.x>
- Shen, H., Du, X., Fan, Y., Dai, J., & Wei, G.-X. (2023). Interoceptive sensibility mediates anxiety changes induced by mindfulness-based tai chi chuan movement intervention. *Mindfulness*, 14(7), 1662–1673. <https://doi.org/10.1007/s12671-023-02162-z>
- Song, L., Huang, Z., Zhou, J., Rao, X., Chen, Y., Li, H., & Du, X. (2023). Effects of aerobic exercise combined with mindfulness-based relapse prevention programs on the relapse tendency in drug abusers. *China Sport Science*, 43(1), 69–77. <https://doi.org/10.16469/j.css.202303007>
- Srisurapanont, M., Sombatmai, S., & Boripuntakul, T. (2007). Brief intervention for students with methamphetamine use disorders: A randomized controlled trial. *The American Journal on Addictions*, 16(2), 111–116. <https://doi.org/10.1080/10550490601184431>
- Tan, H., Chen, T., Du, J., Li, R., Jiang, H., Deng, C.-L., Song, W., Xu, D., & Zhao, M. (2019). Drug-related virtual reality cue reactivity is associated with gamma activity in reward and executive control circuit in methamphetamine use

- disorders. *Archives of Medical Research*, 50(8), 509–517. <https://doi.org/10.1016/j.arcmed.2019.09.003>
- Todd, J., Swami, V., Aspell, J. E., Furnham, A., Horne, G., & Stieger, S. (2022). Are some interoceptive sensibility components more central than others? Using item pool visualisation to understand the psychometric representation of interoception. *PLOS ONE*, 17(12), Article e0277894. <https://doi.org/10.1371/journal.pone.0277894>
- United Nations Office on Drugs and Crime. (2024). *World drug report 2024*. United Nations Publications. <https://www.unodc.org/unodc/en/data-and-analysis/world-drug-report-2024.html>
- Volkow, N. D., Michaelides, M., & Baler, R. (2019). The neuroscience of drug reward and addiction. *Physiological Reviews*, 99(4), 2115–2140. <https://doi.org/10.1152/physrev.00014.2018>
- Wang, D., Zhu, T., Zhou, C., & Chang, Y.-K. (2017). Aerobic exercise training ameliorates craving and inhibitory control in methamphetamine dependencies: A randomized controlled trial and event-related potential study. *Psychology of Sport and Exercise*, 30, 82–90. <https://doi.org/10.1016/j.psychsport.2017.02.001>
- Wang, R. (2001). *Psychological health and relapse tendency of male compulsive drug abusers* [Master's thesis, East China Normal University]. https://kns.cnki.net/kcms2/article/abstract?v=WNLjQhMUSxJ-_gmRRojD5gxFAQ_qJWYbotEVw7_vzUB4wZEDALS8nib1JCUXqQ_QadinQkjMdhs5Tm7ig8beu1-p_h0tGNmSG1GBSO6refd-J-c-14sKeXGoqvJvW6izBy4nPXwSZogGYVHSSc5M5TFoxsQZuuf--toiDy_6OlZD27NvZxyH5A0qEv3_L6XXQ&uniplatform=NZKPT&language=CHS
- Wang, Y.-G., Liu, M.-H., & Shen, Z.-H. (2019). A virtual reality counterconditioning procedure to reduce methamphetamine cue-induced craving. *Journal of Psychiatric Research*, 116, 88–94. <https://doi.org/10.1016/j.jpsychires.2019.06.007>
- Wang, Y.-G., Shen, Z.-H., & Wu, X.-C. (2018). Detection of patients with methamphetamine dependence with cue-elicited heart rate variability in a virtual social environment. *Psychiatry Research*, 270, 382–388. <https://doi.org/10.1016/j.psychres.2018.10.009>
- Williams, D. P., Jarczok, M. N., Ellis, R. J., Hillecke, T. K., Thayer, J. F., & Koenig, J. (2017). Two-week test-retest reliability of the Polar® RS800CX™ to record heart rate variability. *Clinical Physiology and Functional Imaging*, 37(6), 776–781. <https://doi.org/10.1111/cpf.12321>

DO VALUES CHANGE AFTER MAJOR LIFE EVENTS? A META-ANALYTIC REVIEW

Doris Pop¹, Oana A. David^{2, 3*}

¹ Babeş-Bolyai University, Evidence-Based Psychological Assessment and Interventions Doctoral School, Babes-Bolyai University of Cluj-Napoca, Cluj-Napoca, Napoca, Romania

² DATA Lab, The International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health, Babes-Bolyai University Cluj-Napoca, Cluj-Napoca, Romania

³ Department of Clinical Psychology and Psychotherapy, Babeş-Bolyai University, No 37 Republicii Street, 400015, Cluj-Napoca, Cluj, Romania

Abstract

Given the growing data accumulated in the research of personal values change, we explored the existence and intensity of values change following major life events, whether individual or cumulative. A second objective of our research was to identify which of the major life events are studied, which study populations and their effects on personal values change.

A total of 35 studies met the criteria to be included in the present meta-analysis with an average sample of 5289 participants. The mean effect size obtained was small with a great heterogeneity and thus subgroup analyses required grouping personal values in their clusters. The results show that traumatic events (e.g., war) and economic events of the region had the largest effect on mean-level changes in conservation cluster values analysis, while the directed interventions brought medium level changes in the self-enhancement and self-transcendence cluster values, followed by economic events and education. Education and time were associated with small level changes in all value clusters.

Keywords: personal value, value change, event

Personal values are part of the factors influencing our behaviors, decisions and perception. They can also guide our behaviours regardless of our mood (Schwartz, 2012; Sagiv et al., 2017; Cieciuch, 2017). On the other perspective, our life contexts can influence and guide our personal values and behaviours (Egri & Ralston, 2004; Fung, 2016). In other words the way we act, the directions in which we choose to develop and invest our resources is correlated with the values we hold.

* Correspondence concerning this article should be addressed to Oana A. David, DATA Lab, The International Institute for the Advanced Studies of Psychotherapy and Applied Mental Health, Babes-Bolyai University Cluj-Napoca, Romania.
E-mail: oanadavid@psychology.ro

The way we perceive, evaluate and respond to stressors is moderated by the values that are important for us (Sherman et al., 2011).

The universal character of values is agreed upon after more than 10 years of studies surveying participants from Europe Asia, Africa, America and Australia (Schwartz, 2012). These mental and abstract characteristics are at the base of the identity. They give us meaning, coherence and are important for our self-image (Schwartz, 2012).

The first paradigm approaching values suggested that values are, as well as personality traits, constant during the lifespan (Schwartz, 1992). More recent studies show however that values do change over the lifetime, but they need a major event in order to raise or lower in the personal rank of each individual (Sortheix, Parker, Lechner, & Schwartz, 2017; Bardi, Buchanan, Goodwin, Slabu, & Robinson, 2014).

Personal experiences contribute to formation of the rank order of values. They not only enhance the chance to develop and maintain personal values, but these events have the potential to inhibit them (Schwartz, 2011). Beyond knowing the motivation behind this research, it is important to define, explain and clarify the concepts studied.

Personal value theory

In the field of personal values, there is an agreement about the content of values, as Schwartz (2011) states, summarizing more than a decade of research. Personal values are cognitive representations of general, abstract life aspirations. These aspirations are considered desirable by the person holding them (Schwartz, 1992). Personal values are abstract mental constructions that are build verbally and textually (Wilson & Dufrene, 2009).

The explaining model most used for presenting personal values is the one proposed and validated by Schwartz (1992). This theoretic model displays the ten universal values on a circle, grouped in four clusters. This clustering is based on the motivational similarities and dissimilarities that are at the bases of values (Schwartz, 2006).

The self-transcendence cluster contains universalism and benevolence. Achievement, power and hedonism form the self-enhancement cluster. The conservation cluster groups tradition, conformity and security. The openness-to-chance cluster includes hedonism, stimulation and self-direction. Some of the values are anxiety based, for example, security, tradition and others are anxiety-free, like stimulation and self-direction. The visual representation of this scholastic categorization is observable in figure 1., representing the four clusters (Schwartz, 2006).

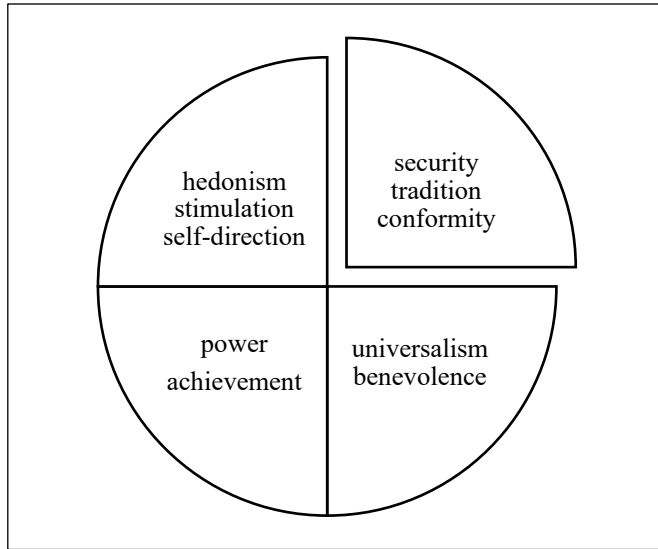


Figure 1. The 4 clusters of Schwartz's personal values model

Major life events and priming

Life events are major subjective and objective disruptive factors of everyday life. Life events can be seen from a stress point of view or from a developmental perspective. In the developmental perspective only transitions are considered life events. This is because transitions disrupt the normal daily life of people, bringing discontinuity in their space. One particular aspect of this type of events is the awareness of the person transiting it (Filipp & Aymanns, 2009).

The major event occurs in the lifespan, bringing chances, and the person walks through it conscious, adapting to the new conditions imposed by it (Filipp & Aymanns, 2009). Some examples of major life events are divorce, a new born baby, the transition from education to work, economic crisis, war.

Do values change during the lifespan?

Experiencing major life events has been associated with value change in individuals (Bardi, Lee, Hofmann-Towfigh, Soutar, 2009). This change can occur in a conditional manner, in accordance with the circular model of values. This means that if the importance of a value increases, the compatible values tend to increase similarly and conflicting values may decrease (Schwartz, 2011).

The present meta-analysis

The current meta-analysis aimed to explore the existence and intensity of value change after major life events. We didn't hypothesize a direction, only an effect, in general, of life events on the hierarchy change of personal values (Boer & Fischer, 2013; Riley, 2017). A second objective of our research was to identify which of the major life events are studied, which study populations are measured and the analysis methods for this type of change that are being used.

Methods

Inclusion criteria

To be included, studies had to fulfill a number of criteria. First, the main outcome of the study had to be personal values. The values had to be measured before and after the major life event or the intervention or for the cohort studies, two times, for one cohort or for two cohorts. Secondly, the predictor or the independent variable had to be a life event or a controlled intervention meant to enhance or decrease the importance of a personal value or more, in this predictor is included passing of time, even if it is not an event. Some studies were excluded because not enough quantitative data was provided, for example, they measured once, or we found some studies where they shown the rank difference score, but did not provide any before or after scores and standard deviations. We only included peer reviewed studies published in English. The complete selection process was documented by the standards from Cochrane, using the PRISMA Flow diagram, as fig. 2 illustrates.

Literature search

We used different methods in order to find as many befitting articles. The first refined literature search was conducted in May 2020, reran the search in 2022, and a third time in 2024, finding new studies each time. We used a search string with variations adapted to the database filters. It had a combination of keywords for: life event, life transition, personal value, value change. For some of the data bases we were forced to shorten the string because of the limits imposed by the search function. The data bases we searched were: Pubmed, PsychInfo, Scopus, Science direct, Proquest, PsychArticles, Web of Science, Wiley. We also included articles that fit the criteria, but were found from other sources. These sources were other searches unrelated to the current study. You can find this summaries in the PRISMA flow figure 2.

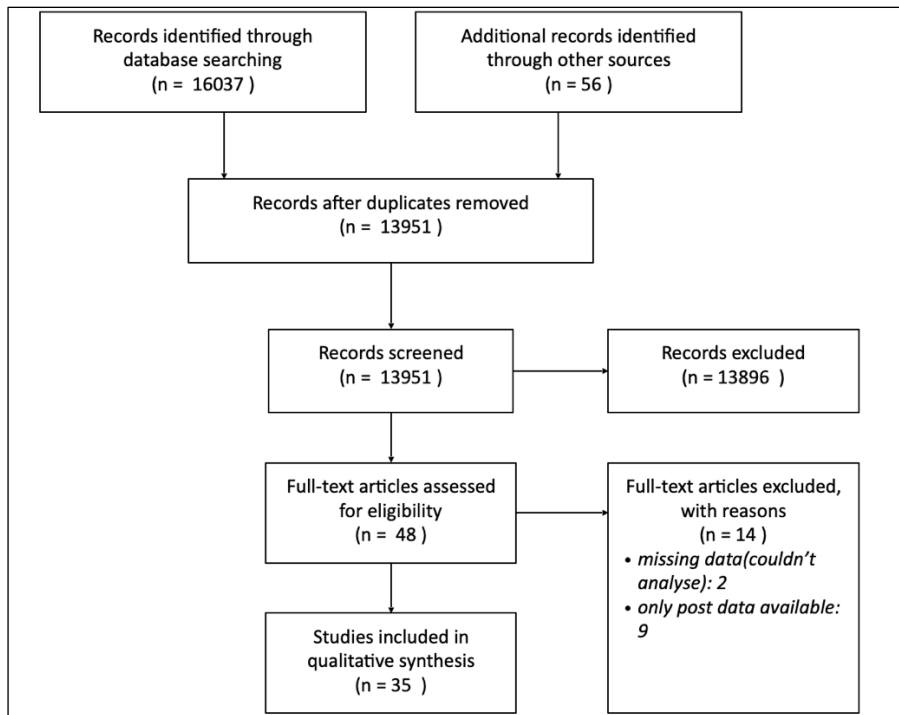


Figure 2. PRISMA Flow

Exclusion reasons

The excluded studies were missing data, on one we could not access the results, only the explanation or the studies reported only a rank difference, no scores and standard deviations for the two times of the measure of personal values. Those had only one time measure, not making a two times comparison so we could not use the data in order to understand the change within that group.

Study variables/data extraction

Personal values. The majority of the studies included in analysis used Schwartz's 10 personal values model, measuring values based on this paradigm. No study used another model, similar to this one and the base for the development of Schwartz's model: Rokeach's system (Rokeach, 1973).

Types of life event. The studies included evaluated the impact of education, emigration, war, terrorist attacks, newborns, shifting from school to the workplace,

interventions, economic crises, socio-political regional changes. We grouped education as a cluster for university courses, police training (of non-police recruits), high school courses and any other type of formal education identified in studies.

Population samples. We did not opt for a specific population, we gathered the articles to explore which populations are represented in these studies. We coded population samples for general population, students, children, adolescences and soldiers. Most studies were designed to assess the difference between before and after the major event, with a few cohort studies which used cross-sectional measures and longitudinal studies.

Populations investigated in these studies are selected from a vast array of regions, from Europe, America, Asia. Each sample is composed, as reported, of people from a single country, exception being the studies that looked at a full continent or a larger region.

Analysis of effect sizes / meta-analytic procedure

Taking into account that the majority of studies used Schwartz's model for assessing personal values, we decided to break the analysis into the four clusters. For each individual cluster we run the analysis to investigate the effect of all events, education, war and terrorism, newborn, socioeconomic change in the region, emigration, intervention and time. Of course, some of this analysis were not valid because only one or none of the studies identified measured that specific cluster in relation with one of the events.

We grouped values outside of Schwartz's model in the main clusters. We mainly did this because the values and their definitions complied with those in the main model. For example, in the self-enhancement cluster we included self-fulfilment and capability along the three classic values, power, achievement and hedonism. Another cluster is conservation, where we added conservationism to the already specified tradition, conformity and security.

The studies selected in this meta-analysis are diverse in the outcomes and factors measured. All thought "personal values" is one concept, measuring these values implies we measure 10 or more continuous variables. We grouped the values in clusters, as research supports the hypothesis that similar values vary the same (Schwartz, 2011). This large display of outcomes and factors requires the use of a random-effects model, assuming there is much heterogeneity in data. We also considered every mean of every value as an outcome, not the mean of the study, if it was presented, because of the complex display of outcomes (Tufanaru, Munn, Stephenson & Aromataris, 2015; Kelley & Kelley, 2012).

The studies included in the analysis of the effect size provided the means and standard deviations for both before and after the event or intervention. Also, most of the studies included has small sample sizes, so we opted for g , over *Cohen's d*. In the case of big samples, these two types of difference analysis offer similar results (Ellis, 2010). The statistical analysis was run using the Comprehensive Meta-Analysis (CMA), using the main features for multiple group means.

Results

Publication bias

One of the important errors in assuming meta-analysis results is considered to be due to publication bias (Sterne & Egger, 2005). We examined the symmetry of the funnel plot for cluster analysis and the asymmetry does not raise concerns (Becker, 2005). We assumed that because some of the samples included in this meta-analysis are relatively large, $N > 40.000$ and some are small, with N greater than 20, the number of unpublished studies needed to reduce the effect size would be large. We also included only peer reviewed studies to insure a greater quality for the results, reducing the risk of enhancing publication bias by selecting only studies from specific research groups known by our team (Ferguson & Brannick, 2012).

Study characteristics

Thorough the search we identified 13951 studies. We ran the search twice, then we added the already selected articles that could match the criteria. All these studies were included in all of the selection phases, from screening to full text trimming. Two individual scientists screened and selected the articles and examined the quality of the studies.

A total of 35 studies met the criteria to be included in the present meta-analysis with an average sample of 5289 participants, with a maximum of 54,931 and a minimum of 36 subjects. We included both small group, within group and cohort samples, with participants from a variety of age groups, from adolescence to +60. The sampled population was from Europe, Asia, America and Pacific Island.

The major events were mostly non-traumatic, with the exception that the war experiences and terrorist attacks, that can be considered traumatic events. The studies that used war as an event did not measure or screen for possible mental health related issues in the selected sample. The most analyzed event was education: higher

education, college, high-school and medium length training (a couple of months) and the least analyzed event was social-economical change: political revolution, economic development.

Mean level changes in values

The combined effect of all events on all values, .328, is significant $p < 0.05$ ($p = 0.013$), with $I^2 = 97.70$, but the heterogeneity is too high to consider a global effect (Hak, van Rhee & Suurmond, 2016). This could mean that in the situation of a major event that changes the life circumstances, for example, a new education cycle, a newborn, a job change, a regional move, could be associated or even considered as a predictor for a change in the preference of personal values.

The subgroup analysis requested grouping personal values in their clusters, and measuring how much they changed for each event or groups of similar events. The four clusters, self-enhancement, openness to change, conservation, self-transcendence were analyzed for: education, time, war, newborn, emigration, short intervention, economic event of the region. Significant cluster results for the mean-level changes on personal values across time and events are presented, in addition to the interpretation, in tables 1,2,3 and 4 and figures 3,4,5 and 6. All results reported in this tables are the significant results from the analysis. All the analysis that used only one study or had no study to analyze were excluded from the reported data.

Self-transcendence cluster

The self-transcendence cluster showed a pattern of small significant increases for economic events ($G = .200$, $SD = .102$), education ($G = .312$, $SD = .057$), time ($G = .372$, $SD = .061$) and the combined events ($G = .321$, $.026$). A medium significant effect on self-transcendence values had the intervention ($G = .507$).

Table 1. Mean-level changes in self-transcendence values

Event	K	Hedges G	SD	Variance	CI		Z	P
					Lower limit	Upper limit		
All events	34	0.321	0.026	0.001	0.271	0.372	12.949	0.000
Economic event	2	0.200	0.102	0.010	0.001	0.399	1.970	0.049
Education	11	0.312	0.057	0.003	0.201	0.424	5.470	0.000
Intervention	6	0.507	0.117	0.014	0.277	0.736	4.330	0.000
Time	8	0.372	0.061	0.004	0.252	0.491	6.088	0.000

Note: K = number of samples; SD = standard error; CI = 95% random effects confidence interval

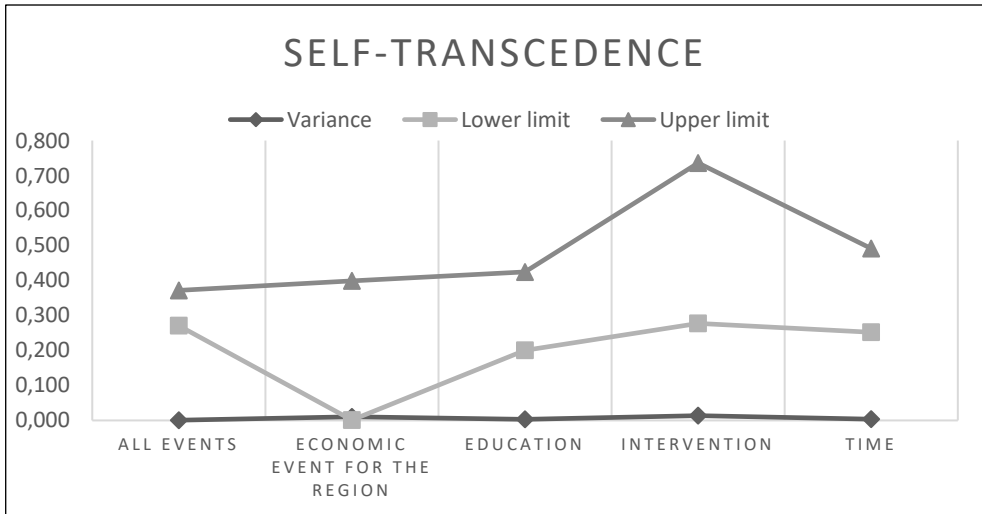


Figure 3. Mean-level changes in self-transcendence values

Conservation cluster

Economic events of the region had the biggest effect on mean-level change in conservation cluster analysis for adult population ($G = 0.790$, $SD = .138$). The difference identified for war and terrorist attack event was conducted on both child and adult samples, with a big effect size ($G = .936$) with a confidence interval that supports the significance of the result ($CI = .622; 1.249$), but a standard deviation of 5.847. We assumed the value of this standard deviation is due to the small samples and the different age groups that were combined. Time had a medium effect on this cluster ($G = .456$, $SD = .064$) whereas all events ($G = .399$, $SD = .032$), education ($G = .322$, $SD = .051$) and emigration ($G = .222$, $SD = .075$) had small significant effect.

Table 2. Mean-level changes in conservation values

Event	K	Hedges G	SD	Variance	CI		Z	P
					Lower limit	Upper limit		
All events	25	0.399	0.032	0.001	0.337	0.461	12.659	0.000
Economic event	2	0.790	0.138	0.019	0.520	1.060	5.730	0.000
Education	9	0.322	0.051	0.003	0.222	0.422	6.313	0.000
Emigration	4	0.222	0.075	0.006	0.075	0.368	2.965	0.003
Time	7	0.456	0.064	0.004	0.331	0.581	7.161	0.000
War	3	0.936	5.847	0.026	0.622	1.249	5.847	0.000

Note: K = number of samples; SD = standard error; CI = 95% random effects confidence interval

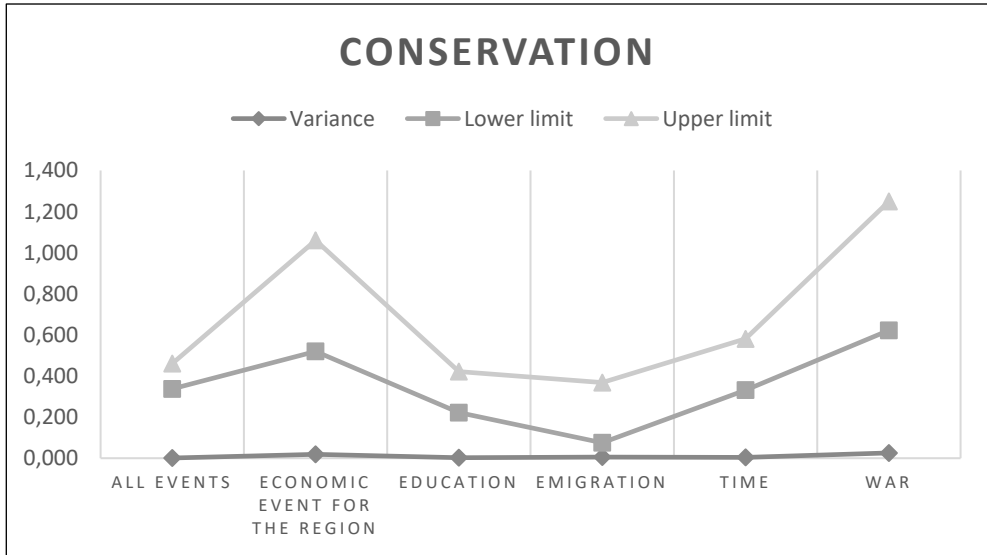


Figure 4. Mean-level changes in conservation values

Self-enhancement cluster

For self-enhancement cluster the effects of intervention stand out ($G = .503$, $SD = .173$). In the uncontrolled conditions, the effects were small: all events combined ($G = .334$, $SD = .025$), economic events for the region ($G = .373$, $SD = .103$), education ($G = .285$, $SD = .044$), emigration ($G = .203$, $SD = .070$), and time ($G = .284$, $SD = .046$).

Table 3. Mean-level changes in self-enhancement values

Event	K	Hedges G	SD	Variance	CI		Z	P
					Lower limit	Upper limit		
All events	26	0.334	0.025	0.001	0.284	0.384	13.085	0.000
Economic event	2	0.373	0.103	0.011	0.171	0.576	3.614	0.000
Education	11	0.285	0.044	0.002	0.200	0.371	6.552	0.000
Emigration	3	0.203	0.070	0.005	0.065	0.341	2.886	0.004
Intervention	3	0.503	0.173	0.030	0.164	0.842	2.905	0.004
Time	7	0.284	0.046	0.002	0.193	0.375	6.114	0.000

Note: K = number of samples; SD = standard error; CI = 95% random effects confidence interval

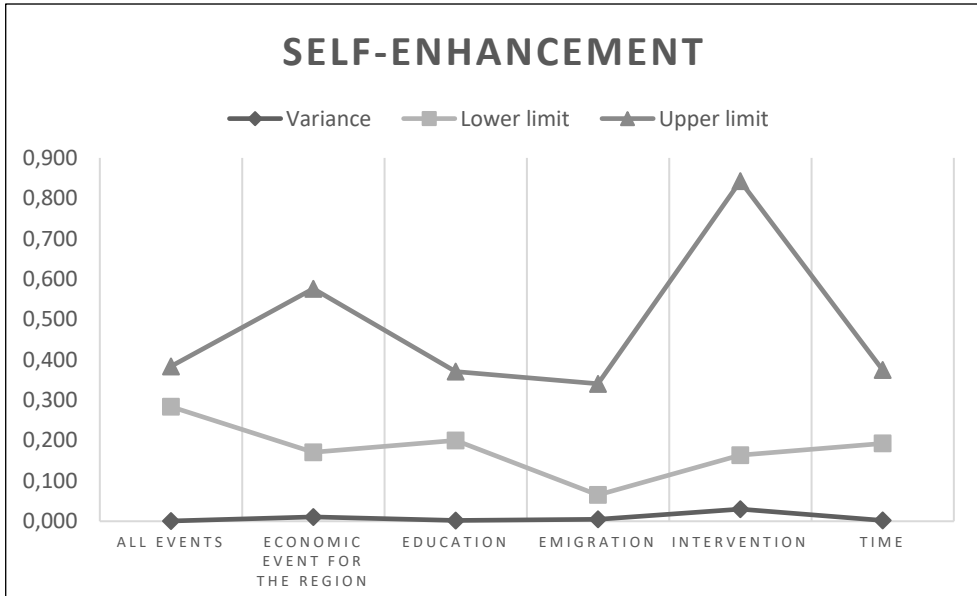


Figure 5. Mean-level changes in self-enhancement values

Openness to change cluster

The results for openness to change cluster provide small significant effect sizes, with time having a slightly larger effect ($G = .362$, $SD = .072$), followed by all events combined ($G = .298$, $SD = .026$), economic events for the region ($G = .225$, $SD = .069$), education ($G = .225$, $SD = .042$) and emigration ($G = .205$, $SD = .101$).

Table 4. Mean-level changes in openness to change values

Event	K	Hedges G	SD	Variance	CI		Z	P
					Lower limit	Upper limit		
All events	21	0.298	0.026	0.001	0.247	0.350	11.320	0.000
Economic event	2	0.225	0.069	0.005	0.091	0.360	3.280	0.001
Education	10	0.225	0.042	0.002	0.143	0.306	5.400	0.000
Emigration	3	0.205	0.101	0.010	0.007	0.402	2.028	0.043
Time	6	0.362	0.072	0.005	0.222	0.502	5.055	0.000

Note: K = number of samples; SD = standard error; CI = 95% random effects confidence interval;

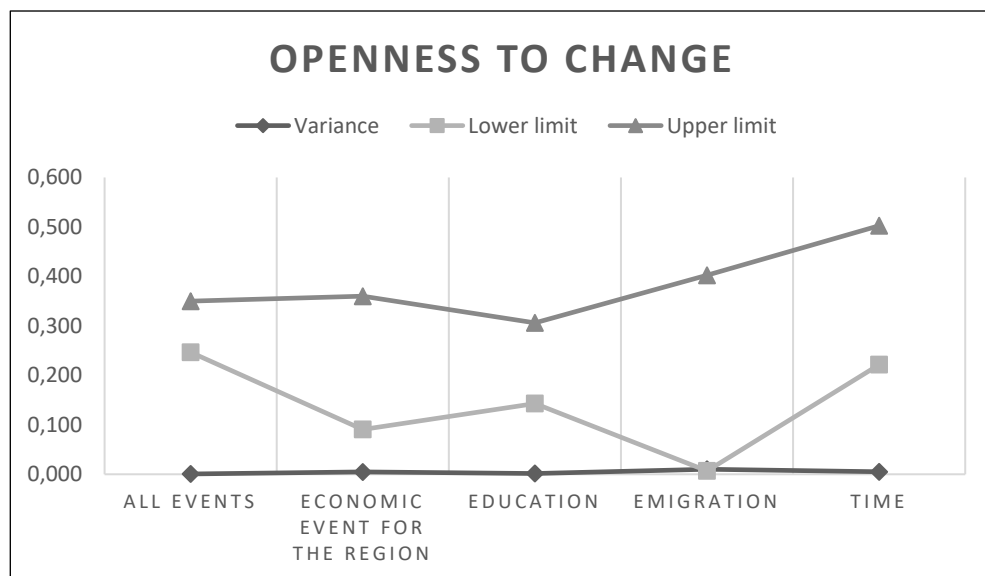


Figure 6. Mean-level changes in openness to change values

Discussion

Our results show that major life events can indeed be a predictor for the change in importance of one or more type of values and their mobility on the personal values prioritising hierarchy. We have found a significant small effect size for the overall changes in personal values following major life events and a large heterogeneity for the studies included. Since the studies look at different populations being affected by different events at different times, we continued to analyze changes in values subgroups in relation to subgroups of major life events. When looking at values cluster subgroups, we have found that different kinds of events impact differently on the clusters of values. Traumatic events (e.g., war) and economic events of the region had the largest effect on mean-level changes in conservation cluster values importance, while the intervention brought medium level changes in the self-enhancement and self-transcendence cluster values, followed by economic events and education.

Interventions type of events had a medium effect on self-transcendence and self-enhancement values. This can partly be explained by the common factor of the clusters, the anxiety-free base for these values, because those people tend to be more open to change and less scared to do things differently (Paunonen & Ashton, 2001; Paunonen, 2003). This could mean that interventions should be aimed at enhancing

these two clusters by priming or by long term interventions, like counseling and coaching. We must take into account that the interventions are implying a controlled event, hence their larger effect.

War and terrorist attack tend to have a larger effect on the importance of values, but there are a few studies that use samples from population that went through these types of events. This is the reason for having a significant results only on conservation values mean-level change, despite it being the largest in the whole meta-analysis ($G = .936$). Such a traumatic life event seems to affect people in becoming more appreciative and fond of security, tradition since they had experienced their lack and gained understanding of their importance for a good or valued life.

Time and education had a small effect on the mean-level change in all values, in all clusters, with a larger effect size of time for openness to change cluster values. Time was conceptualized as growing in age or the passing of time, in different studies. This effect could be used to explain part of the total effect from other events that are long-term, like education, emigration, political events in the region or economic development. As the passing of time is a common partly explaining predictor, we could propose that from now on the studies could take into account this factor when pondering about the effects of major life events on the value hierarchy.

We expected small effect sizes taken into consideration the number of samples and the number of subjects in some of the samples and the results support this assumption. Given the fact that values, like personality traits, are part of the identity and pattern of human behaviour, it is expected to have smaller effect sizes from all the events (Parks, Feldman & Bardi, 2015), except those which are having the traumatic potential, as war and terrorist attack. This finding implies that group level of studying personal values is valid and should be looked into more, especially as a trend for smaller or more niche populations.

While interpreting the results, we might consider that these studies are pioneering the path of personal values change. Taking into account the time effect discussed before, we can assume there are other factors that can influence these results. Starting with the individual characteristics of sampled subjects, the randomization of the selection process, most studies use convenience samples of people going through the events. Partly, the variability of the chance can be explained by the use of different scales or the variation in the time period between the initial measure and the measure post event.

Limits

Results from previous studies show that values measured with Portrait Value Questionnaire and Schwartz Value Survey yield different relations with personality (Parks-Leduc, Feldman, Bardi, 2015). This means that the current results can be

biased by the use of those two different scales and the combined analysis of their measures.

Some of the samples used in calculating the combined effect sizes are small and this is a source of error. There is a possibility that using G instead of Cohen's D can lead to higher error associated with small samples. (Lin, 2018)

There are several factors that could have contributed to this effects, within is the time effect. As results support, time had an effect on personal values so we have to be careful when interpreting the resulted data because in all other events effect could be a factor of time. Moreover, each event has a limited number of reseach teams evaluating it from the perspective of personal value change.

Future results

The field can expand in the direction of testing these changes in personal values by controlling the environment and the event. This means that using guided imagination, films and virtual reality, we can verify some of the mechanisms and directions of change in a variety of populations. More priming intervention studies are needed in order to conclude that the values do change in this priming process. Follow up measures are also needed because is would be valuable to know for how long do values stay in the new order before they regress to the initial hierarchy.

More leaterature and research is need for verifying replicability of the results in the same and different participants. We consider this to be fundamental with the argument that it postulated the current state of the research, with identifying the gaps: not enough replicating studies for each event or value type or participant, some proof of change with no event being the main predictor for the change, as the studies where the passing of time was the sole measurement, the need to verify other parameters which could be useful in understanding the change, as are mindset change, how influencable those people are, their personality, their trait emotions, if they have any, the level of self -awareness. We also consider that being a well-being and life span area of human functioning, we could be interested in asking how this changes correlate with improved or reduced well-being, life satisfaction and cognitive dissonance.

Conclusions

Our results demonstrate that value hierarchy changes on small groups and niche population samples, not only in the cohort analysis, conducted on houndreds of thousands of people at once. This supports the necessity of developing new studies that identify specifically how values change for specific events and people can be beneficial in understanding an individualized change, not just a generic one and what mechanisms sustain the chance. This research field can have profound implications

in explaining mental health fluctuations in the general population. More research is needed that measure the relation between how much personal values change, the severity of mental health problems and how affected is the life satisfaction of the person. We consider that this deepened understanding in the field of personal values research could potentially improve our understanding of both changes in mental functioning and identity.

Authors' Notes

Aknowlegment: The research was not funded, we have no conflicts of interest, we aknowlegde the help we had from contributors not named as authors in the study Tomoioagă Cristina Teodora and Simona Miron for verifying the articles and the data extraction.

References

- Arieli, S., Grant, A. M., & Sagiv, L. (2014). Convincing Yourself to Care About Others: An Intervention for Enhancing Benevolence Values. *J Pers*, 82(1), 15-24. doi:10.1111/jopy.12029
- Bardi, A., Buchanan, K. E., Goodwin, R., Slabu, L., & Robinson, M. (2014). Value stability and change during self-chosen life transitions: self-selection versus socialization effects. *J Pers Soc Psychol*, 106(1), 131-147.
- Bardi, A., & Goodwin, R. (2011). The Dual Route to Value Change: Individual Processes and Cultural Moderators. *Journal of Cross-Cultural Psychology*, 42(2), 271-287. doi:10.1177/0022022110396916
- Bardi, A., Lee, J. A., Hofmann-Towfigh, N., & Soutar, G. (2009). The structure of intraindividual value change. *J Pers Soc Psychol*, 97(5), 913-929. doi:10.1037/a0016617
- Baxter, J., Buchler, S., Perales, F., & Western, M. (2014). A Life-Changing Event: First Births and Men's and Women's Attitudes to Mothering and Gender Divisions of Labor. *Social Forces*, 93(3), 989-1014. doi:10.1093/sf/sou103
- Becker, B. J. (2005). Failsafe N or file-drawer number. *Publication bias in meta-analysis: Prevention, assessment and adjustments*, 111-125.
- Boer, D., & Fischer, R. (2013). How and when do personal values guide our attitudes and sociality? Explaining cross-cultural variability in attitude–value linkages. *Psychological bulletin*, 139(5), 1113.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2022). *Comprehensive Meta-Analysis (Version 4)* [Computer software]. Biostat, Englewood, NJ. <https://meta-analysis.com/pages/faq>

- Cieciuch, J. (2017). Exploring the complicated relationship between values and behaviour. *Values and behavior: Taking a cross cultural perspective*, 237-247.
- Dæhlen, M. (2005). Change in job values during education. *Journal of Education and Work*, 18(4), 385-400. doi:10.1080/13639080500327774
- Daniel, E., Fortuna, K., Thrun, S. K., Cioban, S., & Knafo, A. (2013). Brief report: early adolescents' value development at war time. *J Adolesc*, 36(4), 651-655. doi:10.1016/j.adolescence.2013.03.009
- Doring, A. K., & Hillbrink, A. (2015). Brief report: Into the wild? How a film can change adolescents' values. *J Adolesc*, 40, 78-82. doi:10.1016/j.adolescence.2015.01.006
- Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455-463.
- Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge university press.
- Egri, C. P., & Ralston, D. A. (2004). Generation cohorts and personal values: A comparison of China and the United States. *Organization science*, 15(2), 210-220.
- Fassaert, T., De Wit, M. A., Tuinebreijer, W. C., Knipscheer, J. W., Verhoeff, A. P., Beekman, A. T., & Dekker, J. (2011). Acculturation and psychological distress among non- Western Muslim migrants--a population-based survey. *Int J Soc Psychiatry*, 57(2), 132-143. doi:10.1177/0020764009103647
- Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: prevalence, methods for identifying and controlling, and implications for the use of meta analyses. *Psychological methods*, 17(1), 120.
- Filipp, S., & Aymanns, P. (2009). Critical life events and life crises: Coping with the dark sides of life.
- Fung, H. H., Ho, Y. W., Zhang, R., Zhang, X., Noels, K. A., & Tam, K. P. (2016). Age differences in personal values: Universal or cultural specific?. *Psychology and aging*, 31(3), 274.
- Fukasawa, M., Watanabe, K., Nishi, D., & Kawakami, N. (2020). Longitudinal association between adolescent work values and mental health and well-being in adulthood: a 23- year prospective cohort study. *Scientific Reports*, 10(1). doi:ARTN 13547 10.1038/s41598-020-70507-y
- Gavreliuc, A. (2012). Continuity and change of values and attitudes in generational cohorts of the post-communist Romania. *Cognitie, Creier, Comportament/Cognition, Brain, Behavior*, 16(2).
- Hak, T., van Rhee, H., & Suurmond, R. (2016). How to interpret results of meta-analysis. *Available at SSRN 3241367*.
- Helkama, K. (2010). Moral Reasoning and Values in Medical School: A longitudinal study in Finland. doi:10.1080/00313830308589
- Isaksson, K. (2004). Work values accross the unemployed.

- Johnson, M. K. (2001). Change in job values during the transition to adulthood. *Work and Occupations*, 28(3), 315-345. doi:10.1177/0730888401028003004
- Katz-Wise, S. L., Priess, H. A., & Hyde, J. S. (2010). Gender-role attitudes and behavior across the transition to parenthood. *Dev Psychol*, 46(1), 18-28.
- Kelley, G. A., & Kelley, K. S. (2012). Statistical models for meta-analysis: A brief tutorial. *World journal of methodology*, 2(4), 27-32.
- Kromrey, J. D., & Rendina-Gobioff, G. (2006). On knowing what we do not know: An empirical comparison of methods to detect publication bias in meta-analysis. *Educational and Psychological Measurement*, 66(3), 357-373.
- Lin, L., Chu, H., Murad, M. H., Hong, C., Qu, Z., Cole, S. R., & Chen, Y. (2018). Empirical comparison of publication bias tests in meta-analysis. *Journal of general internal medicine*, 33(8), 1260-1267.
- Michael Borenstein et al. (2009), *Introduction to Meta-Analysis*, Chichester (UK): Wiley.
- Milfont, T. L., Milojev, P., & Sibley, C. G. (2016). Values Stability and Change in Adulthood: A 3-Year Longitudinal Study of Rank-Order Stability and Mean-Level Differences. *Pers Soc Psychol Bull*, 42(5), 572-588. doi:10.1177/0146167216639245
- Myrsky, L., Juujarvi, S., & Pessa, K. (2013). Change in values and moral reasoning during higher education. *European Journal of Developmental Psychology*, 10(2), 269-284. doi:10.1080/17405629.2012.757217
- Nilsson, A., Denison, E., & Lindberg, P. (2011). Life values as predictors of pain, disability and sick leave among Swedish registered nurses: a longitudinal study. *BMC Nurs*, 10, 17. doi:10.1186/1472-6955-10-17
- Parks-Leduc, L., Feldman, G., & Bardi, A. (2015). Personality Traits and Personal Values: A Meta-Analysis. *Personality and Social Psychology Review*, 19(1), 3-29.
- Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of personality and social psychology*, 81(3), 524.
- Racko, G., Strauss, K., & Burchell, B. (2017). Economics Education and Value Change: The Role of Program-Normative Homogeneity and Peer Influence. *Academy of Management Learning & Education*, 16(3), 373-392. doi:10.5465/amle.2014.0280
- Paunonen, S. V. (2003). Big Five factors of personality and replicated predictions of behavior. *Journal of personality and social psychology*, 84(2), 411.
- Ralston, D. A., Egri, C. P., Karam, C. M., Li, Y., & Fu, P. P. (2017). Changes in work values across the regions of China. *Asia Pacific Journal of Management*, 35(1), 145-179. doi:10.1007/s10490-017-9519-y.
- Riley, R. D., Jackson, D., Salanti, G., Burke, D. L., Price, M., Kirkham, J., & White, I. R. (2017). Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *bmj*, 358.
- Rokeach, M. (1973). *The nature of human values*. Free press

- Robinson, O. C. (2013). Values and adult age: findings from two cohorts of the European Social Survey. *Eur J Ageing*, 10(1), 11-23.
- Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). Publication bias in meta-analysis. *Publication bias in meta-analysis: Prevention, assessment and adjustments*, 1-7.
- Rudnev, M. (2014). Value Adaptation Among Intra-European Migrants: Role of Country of Birth and Country of Residence. *Journal of Cross-Cultural Psychology*, 45(10), 1626- 1642. doi:10.1177/0022022114548482
- Sagiv, L., Roccas, S., Cieciuch, J., & Schwartz, S. H. (2017). Personal values in human life. *Nature human behaviour*, 1(9), 630-639.
- Sagiv, L., & Schwartz, S. H. (2022). Personal values across cultures. *Annual review of psychology*, 73(1), 517-546.
- Sheldon, K. M. (2005). Positive value change during college: Normative trends and individual differences. *Journal of Research in Personality*, 39(2), 209-223. doi:10.1016/j.jrp.2004.02.002
- Sherman, D. K., Uskul, A. K., & Updegraff, J. A. (2011). The role of the self in responses to health communications: A cultural perspective. *Self and Identity*, 10(3), 284-294.
- Sortheix, F. M., Parker, P. D., Lechner, C. M., & Schwartz, S. H. (2017). Changes in Young Europeans' Values During the Global Financial Crisis. *Social Psychological and Personality Science*, 10(1), 15-25.
- Schwartz, S. H. (2012). An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture*, 2(1). doi:10.9707/2307-0919.1116
- Schwartz, S. H. (2011). Studying Values: Personal Adventure, Future Directions. *Journal of Cross-Cultural Psychology*, 42(2), 307-319. doi:10.1177/0022022110396925
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in experimental social psychology* (Vol. 25, pp. 1-65). Academic Press.
- Schwartz, S. H. (2006). Basic human values: An overview.
- Sortheix, F. M., Parker, P. D., Lechner, C. M., & Schwartz, S. H. (2017). Changes in Young Europeans' Values During the Global Financial Crisis. *Social Psychological and Personality Science*, 10(1), 15-25. doi:10.1177/1948550617732610
- Sterne, J. A., & Egger, M. (2005). Regression methods to detect publication and other bias in meta-analysis. *Publication bias in meta-analysis: Prevention, assessment and adjustments*, 99-110.
- Sundberg, R. (2016). Value Stability and Change in an ISAF Contingent. *J Pers*, 84(1), 91- 101. doi:10.1111/jopy.12142
- Tan, S. J., & Tambyah, S. K. (2016). Shifting Values and Life Satisfaction: A Sequential Cross-Sectional Study of the Influence of Values on Subjective Wellbeing in Singapore. *Social Indicators Research*, 127(3), 1391-1416. doi:10.1007/s11205-015- 1015-5

- Tufanaru, C., Munn, Z., Stephenson, M., & Aromataris, E. (2015). Fixed or random effects meta-analysis? Common methodological issues in systematic reviews of effectiveness. *International journal of evidence-based healthcare*, 13(3), 196-207.
- Tulviste, T., & Tamm, A. (2014). Brief report: value priorities of early adolescents. *J Adolesc*, 37(5), 525-529. doi:10.1016/j.adolescence.2014.04.006
- Vecchione, M., Schwartz, S., Alessandri, G., Doring, A. K., Castellani, V., & Caprara, M. G. (2016). Stability and change of basic personal values in early adulthood: An 8-year longitudinal study. *Journal of Research in Personality*, 63, 111-122. doi:10.1016/j.jrp.2016.06.002
- Verkasalo, M., Goodwin, R., & Bezmenova, I. (2006). Values following a major terrorist incident: Finnish adolescent and student values before and after September 11, 2001. *Journal of Applied Social Psychology*, 36(1), 144-160.
- Ye, S. Q., & Ng, T. K. (2019). Value change in response to cultural priming: The role of cultural identity and the impact on subjective well-being. *International Journal of Intercultural Relations*, 70, 89-103. doi:10.1016/j.ijintrel.2019.03.003

EFFECTIVENESS OF GROUP-BASED COGNITIVE BEHAVIORAL THERAPY IN REDUCING DEPRESSION AND ANXIETY AMONG ADOLESCENTS IN RESIDENTIAL INSTITUTIONS: A RANDOMIZED CONTROLLED TRIAL

Ding Wai Eng¹ and Firdaus Mukhtar^{2*}

¹ Ding Psychology Centre

² Department of Psychiatry, Faculty of Medicine & Health Sciences, Universiti Putra Malaysia

Abstract

Research on cognitive behaviour therapy (CBT) for depressed and anxious orphaned adolescents in Malaysian sheltered homes is limited. This randomised controlled trial evaluated the effectiveness of group CBT, known as the SAHABAT programme, compared to a waitlist control. One hundred thirty-nine orphaned adolescents from nine sheltered homes participated: SAHABAT (4 homes, $n = 71$) and waitlist control (5 homes, $n = 68$), selected through clustered sampling. The SAHABAT group attended eight sessions over a month, while the control group received no intervention. Psychological outcomes were assessed at six-time points over six months using five self-report questionnaires. Generalised linear multi-modelling analysis showed that the SAHABAT programme had mild to moderate effects in reducing depression ($\eta^2 = .05$), anxiety ($\eta^2 = .05$), negative automatic thoughts ($\eta^2 = .05$), and anger ($\eta^2 = .07$). It also significantly improved self-esteem ($\eta^2 = .02$). Depression symptoms reduced from mild to minimal, anxiety from moderate to mild, negative automatic thoughts from high to low, and anger from mildly elevated to normal, while self-esteem increased. The SAHABAT programme positively impacted mental health outcomes over six months post-intervention. The findings highlight SAHABAT's positive impact and the need for enhanced mental health policies for vulnerable adolescents in Malaysia.

Keywords: Anxiety, depression, group cognitive behavioural therapy, negative automatic thoughts, orphaned adolescents, self-esteem

* Correspondence concerning this article should be addressed to Firdaus Mukhtar, Department of Psychiatry, Faculty of Medicine & Health Sciences, Universiti Putra Malaysia, UPM Serdang, Malaysia.

E-mail: drfirdaus@upm.edu.my

Official statistics on orphans and orphanages in Malaysia remain limited, resulting in insufficient information regarding the mental health and well-being of orphaned children. Recent estimates indicate that approximately 64,000 children reside in institutional care, encompassing both registered and unregistered government and private orphanages (OrphanCare Foundation, 2020). These institutions frequently house a heterogeneous population comprising orphans and non-orphans from diverse family circumstances (Registrar of Societies Malaysia, 2013).

Studies indicate that more than 80% of children and adolescents residing in Malaysian orphanages or foster care facilities experience significant mental health issues, a stark contrast to the approximately 20% prevalence observed in the general adolescent population (Jayakumar, 2023). Orphans in institutional care frequently demonstrate elevated levels of emotional distress, including depression, anxiety, and diminished self-esteem (Mohammadzadeh et al., 2018). In Malaysia, reported prevalence rates of depression and anxiety among institutionalised youth range from 70.9% to 85.2% and 80.1% to 82.3%, respectively, with 64.1% experiencing comorbid symptoms (Mukhtar et al., 2021; Mohammadzadeh et al., 2018). Higher levels of depression and anxiety are strongly associated with increased negative automatic thoughts (Mukhtar et al., 2021), while 84.7% of institutionalised adolescents report experiencing stress and 70% exhibit low self-esteem (Mohammadzadeh et al., 2018). Collectively, these findings underscore the substantial emotional burden faced by this population and the urgent need for effective, targeted mental health interventions for vulnerable institutionalised youth.

For decades, cognitive behavioural therapy (CBT) has been widely recognised as an effective intervention for adolescent depression and anxiety (James et al., 2015; Keles & Idsøe, 2018; Kreuze et al., 2018; Zhou et al., 2015) as well as anger-related difficulties (Dejid, Delgerjav, & Tsogzolmaa, 2022). CBT has been shown to effectively reduce negative automatic thoughts in adolescents, with cognitive restructuring mediating improvements in depressive and anxiety symptoms (Topper et al., 2023). Emerging evidence further demonstrates that CBT reduces ruminative thoughts (Mirzaian et al., 2023) and self-harming cognitions (Elahabadi, 2025). Mediation findings additionally indicate that decrease in negative automatic thoughts partially account for broader reductions in depression and anxiety, underscoring the central importance of targeting maladaptive cognitions in adolescent CBT (Florea et al., 2024).

Randomised control trials (RCTs) evaluating CBT have consistently reinforced its efficacy. Meta-analysis evidence demonstrates that CBT significantly reduces subclinical depressive symptoms in youths both immediately post-treatment and at subsequent follow-up assessments (Gearing et al., 2013; James et al., 2015; Keles & Idsøe, 2018; Rasing et al., 2017; Weisz et al., 2006; Zhou et al., 2015). In Malaysia, Saw et al. (2020) reported that the STAR school-based CBT programme

produced significant reductions in adolescent depressive symptoms compared with a waitlist control. However, most Malaysian research examining CBT for adolescent depression and anxiety is limited to case studies (Chee et al., 2023) and lacks comparison groups (Zakaria et al., 2021). Only a small number of RCTs targeting anxiety among Malaysian adolescents exist, and those that do primarily focus on non-CBT interventions (Ab Ghaffar et al., 2019; Mohammadzadeh et al., 2020).

Globally, CBT programmes such as Coping Cat and FRIENDS have met the criteria for empirically supported treatment and have demonstrated effectiveness in both preventing and treating anxiety disorders and depression among school-aged children and adolescents (Cooper & Jacobs, 2011; Hollon & Beck, 2013). Their efficacy has also been documented across diverse and underserved contexts, including economically disadvantaged urban communities (Cooley-Strickland et al., 2011), children and adolescents from low socio-economic backgrounds in Puerto Rico (Rodgers & Dunsmuir, 2015), and orphaned youth in Mexico orphanages (Gallegos-Guajardo et al., 2015).

Similarly, a universal school-based CBT package in Sri Lanka yielded significant post-intervention increases in self-esteem and reductions in depressive and anxiety symptoms, although these effects were less stable at follow-up (Perera et al., 2020). Additionally, an RCT incorporating problem-solving and assertiveness training demonstrated improvements not only in self-esteem but also in overall mental health, problem-solving abilities, and communication skills among adolescent girls (Sharif et al., 2021). Collectively, these findings indicate that CBT-based interventions are effective in enhancing adolescent self-esteem while simultaneously improving broader emotional and functional outcomes.

Furthermore, CBT is widely recognised as an effective intervention for anger in adolescent (Matthys & Schutter, 2021), incorporating skills such as emotion recognition, relaxation, and problem-solving to reduce aggressive responses. Although traditional CBT yields small to moderate effects, emerging evidence indicates that exposure-based and mechanistically informed approaches show greater promise, particularly by targeting underlying processes such as frustration tolerance and emotion regulation (Naim et al., 2023). School-based anger management programmes similarly report improvements in communication and problem-solving skills, although their effectiveness varies depending on delivery format and the degree of caregiver involvement (Anjanappa et al., 2023). Overall, CBT remains a first-line psychosocial treatment for adolescent anger, with growing support for adapted interventions design to enhance long-term outcomes (Richard et al., 2022).

Despite CBT being widely recognised as the gold standard in psychotherapy, with substantial evidence demonstrating its superiority or equivalence to alternative treatments (David et al., 2018), rigorous outcome research in Malaysia, particularly RCTs involving adolescents and orphaned youth remains limited. Existing local

studies are scarce, often relying on randomized control designs, and rarely assess multiple internalising and emotional outcomes simultaneously. Furthermore, there is a notable lack of published RCTs in Malaysia evaluating the effectiveness of group-based CBT programmes, and no study to date has examined the SAHABAT programme (Mukhtar, 2014) using procedures consistent with CONSORT 2010 guidelines.

Purpose of the Study

This study addresses these gaps by conducting a CONSORT-aligned RCT comparing group-based CBT (SAHABAT programme) with a waitlist control to evaluate its effectiveness in reducing depression, anxiety, anger, and negative automatic thoughts, and in enhancing self-esteem among Malaysian orphaned adolescents residing in non-government-run sheltered homes. SAHABAT, a Malay term meaning “friends,” reflects the programme’s emphasis on supportive, peer-based learning. By generating robust RCT evidence, this study has the potential to inform and strengthen mental health interventions for this vulnerable population.

Method

Pilot Study

A preliminary pilot study was conducted to evaluate the feasibility, acceptability, and potential effectiveness of the SAHABAT programme, as well as its suitability for replication in similar settings. Forty orphaned adolescents were initially screened for symptoms of depression and anxiety using the Beck Depression Inventory–Malay version (BDI-Malay; Beck, 1961; Mukhtar & Oei, 2008) and the Beck Anxiety Inventory–Malay version (BAI-Malay; Beck et al., 1988; Mukhtar & Zulkefly, 2011). Nineteen adolescents aged 13–17 years, residing in a Malay-operated, non-governmental sheltered home in Selangor and meeting criteria for depression (BDI-Malay ≥ 10) and anxiety (BAI-Malay ≥ 8), were recruited to participate. Participants engaged in eight bi-weekly group sessions of the SAHABAT programme over one month. Findings indicated that the SAHABAT programme was feasible, well-received by participants, culturally adaptable, and effective in alleviating depressive and anxious symptoms, supporting its potential for replication in similar contexts.

Primary Study

Sampling

A randomized controlled trial (RCT) was conducted. The sampling frame, comprising both orphans and non-orphans residing in Malay-operated, non-governmental sheltered homes, was compiled using records from the Registrar of Societies Malaysia, the Department of Social Welfare Malaysia, and the Department of Islamic Selangor. Selangor was selected as the study site due to its accessibility to a high number of registered sheltered homes that can potentially draw a larger sample size and the relatively high prevalence of adolescent depression and anxiety compared to other Malaysian states (Institute for Public Health, 2022). Government-run sheltered homes were excluded to ensure consistency in governance and facilities, and minimize potential sources of bias. Moreover, government-run sheltered homes were not included because there are very few of them in Malaysia. Most residential facilities for children and adolescents are run by non-governmental organizations, making nongovernmental-run homes more accessible and representative for this study. Cultural and linguistic considerations were prioritized, focusing on Malay-operated, non-governmental homes to ensure contextual relevance and reduce language barriers.

Administrators and participants were briefed on the purpose, duration, and procedures of the study. Informed consent was obtained from home administrators, who served as legal guardians for the adolescents.

The inclusion and exclusion criteria for the RCT of this study were determined by the study authors, both of whom are experienced clinical psychologists. Sheltered homes were eligible for inclusion if they were Malay-operated, non-governmental facilities providing accommodation, food, financial assistance, medical access, and national-school education to underprivileged children and adolescents, including both orphans and non-orphans. Exclusion criteria comprised government-run sheltered homes, non-Malay-operated sheltered homes, religious schools, homeschooling centres, and management offices that do not provide residential care. Facilities were also excluded if they housed fewer than three adolescents, accommodated adults or older persons, or served populations with physical or intellectual disabilities. Kindergartens, childcare centres, nursing homes, and community service centres were similarly excluded.

A multi-stage cluster sampling strategy was utilised. Seventeen Malay-operated, non-governmental sheltered homes initially met the eligibility criteria; however, eight withdrew prior to recruitment, resulting in nine homes being retained as the final sampling units. By coincidence, each of these homes was located in a different district of Selangor, forming nine distinct clusters. The Accuracy in Parameter Estimation (AIPE) method was applied to enhance precision in population

parameter estimation, ensuring adequate sample sizes with appropriately narrow confidence intervals.

This study employed a two-arm, parallel, open, single-blinded RCT with a pre-test and post-test experimental design. The objective was to compare the effectiveness of the SAHABAT programme with that of a waitlist control condition in reducing depression, anxiety, negative automatic thoughts, and anger, as well as in enhancing self-esteem. Single-blinded simple lottery randomisation was employed to allocate the participating sheltered homes to either the intervention or waitlist control group.

Participants

Eligible participants were invited to voluntarily participate, with explicit clarification of the study objectives, expectations, and their right to withdraw at any time. Participants were eligible if they were orphans aged 13–17 years exhibiting depressive and/or anxiety symptoms ($\text{BDI-Malay} \geq 10$; $\text{BAI-Malay} \geq 8$), had resided in the shelter for at least one day, and demonstrated proficiency in the Malay language. Exclusion criteria included being a non- orphan, having a learning disorder or organic brain disorder, current use of psychotropic medication, engagement in ongoing psychotherapy, or absence during the recruitment period. Non-orphans were excluded because the project's funding was specifically allocated for interventions targeting the well-being of orphaned adolescents. Participants aged 13–17 years were recruited to guarantee adequate cognitive and linguistic competence for participation in cognitive-behavioral therapy (CBT) (Halder & Mahato, 2019). While factors such as parental loss and resilience are known to influence emotional distress, practical constraints limited the comprehensive assessment of these variables.

Figure 1 presents the CONSORT flow diagram. Initially, 17 Malay-operated, non- governmental run sheltered homes ($N = 632$ adolescents) were randomised, resulting in four homes ($n = 71$) being assigned to the intervention group and five homes ($n = 68$) to the waitlist control group. Of the 632 potential participants, only 243 adolescents from the nine consenting homes were eligible for further screening. Among these, 139 orphaned adolescents met the inclusion criteria and were retained in the final sample, with 71 allocated to the intervention group and 68 to the waitlist control group. The slight imbalance in group sizes resulted from the cluster-level randomisation procedure, whereby all adolescents within each randomised home were assigned to the same study arm.

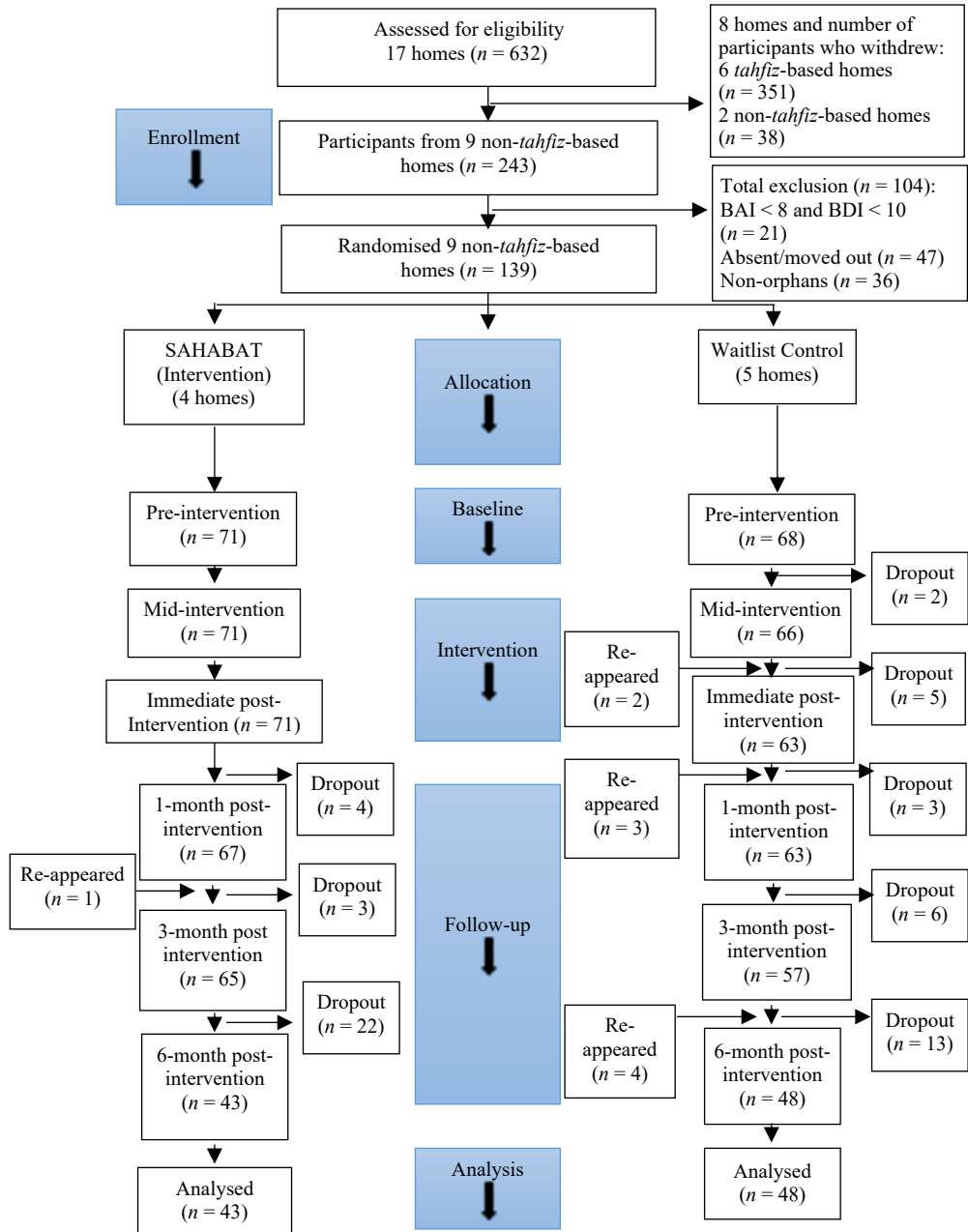


Figure 1. CONSORT Flow Chart

Following completion of the RCT, the SAHABAT programme and its accompanying materials were offered to participants in the waitlist control group;

however, attrition prevented some individuals from participating in the post-trial intervention. Subsequently, the intervention was also disseminated to non-orphaned adolescents residing in non-Malay- operated sheltered homes.

Intervention

The SAHABAT manual, developed in Malay, consists of eight structured core sessions: Session 1 focused on ice-breaking activities, personal perceptions, and self- evaluation; Session 2 introduces relaxation strategies, self-esteem building, and anger management; Session 3 addresses the recognition of physiological symptoms, cognitive distortions, emotions, and behaviours; Session 4 emphasises strategies for modifying cognitive distortions to facilitate emotional and behavioural change; Session 5 covers problem-solving skills; Session 6 guides participants in identify their support system; Session 7 focuses on cultivating inner peace through positive self-talk; and Session 8 involves a review and consolidation of learning, culminating in a celebration of participants' progress.

Measures

The BDI-Malay (Beck et al., 1961; Mukhtar & Oei, 2008) and the BAI-Malay (Beck et al., 1988; Mukhtar et al., 2011) both demonstrated strong construct validity and high internal consistency in Malay samples ($\alpha = .90$). The Automatic Thoughts Questionnaire (ATQ-Malay) (Hollon & Kendall, 1980; Oei & Mukhtar, 2008) showed good factorial validity with Cronbach's alphas ranging from .83 to .93. The Beck Youth Inventories–Anger Scale (BANI-Y Malay) (Beck et al., 2005; Ding et al., 2021) demonstrated excellent reliability ($\alpha = .92$) and convergent validity. The Rosenberg Self-Esteem Scale (RSES-Malay) (Rosenberg, 1965; Mohd Jamil, 2006) also showed acceptable psychometric properties in Malay adolescents, with alphas around .80.

Data Collection

At each measurement interval of pre-intervention (T1), mid-intervention (T2), immediate post-intervention (T3), 1-month follow-up (T4), 3-month follow-up (T5), and 6- month follow-up (T6), participants completed a battery of validated Malay-language versions of standardised questionnaires. The waitlist control group did not receive the intervention across the six measurement points.

Data Analyses

Data analyses comprised descriptive statistics, independent samples t-tests, and assumption testing for normality. A generalized linear mixed model (GLMM) was subsequently implemented in SPSS to enhance the robustness of the inferential

analyses by accounting for inter-individual variability nested within sheltered homes. The GLMM examined the effects of the SAHABAT programme, relative to the waitlist control condition, on depression, anxiety, negative automatic thoughts, self-esteem, and anger across the six measurement points. Sociodemographic covariates, baseline scores, intervention group, time, and the group x time interaction term were included as fixed effects, whereas sheltered homes were modelled as random effects to account for cluster-level dependency. Model fit was evaluated using Akaike's Information Criterion (AIC), and Bonferroni-adjusted significance thresholds were applied ($p \leq .05$). Effect sizes were calculated using eta-squared (η^2).

Results

Table 1 presents the sociodemographic characteristics of the orphaned adolescents in the RCT. Nearly all participants were Malay (99.9%) and all identified as Muslim (100%). The mean age of the sample was 14.81 years ($SD = 1.36$). Younger adolescents (13–15 years) constituted 67.6% ($n = 94$), while older adolescents (16–17 years) comprised 32.4% ($n = 45$). The sample included 63 males (45.3%) and 76 females (54.7%). A majority of participants were single orphans 91.4% ($n = 127$), with 8.6% ($n = 12$) identified as double orphans. Additionally, 54% ($n = 75$) reported having previously attended school-based counselling, whereas 46% ($n = 64$) had never engaged in counselling services.

A comparison between the SAHABAT intervention group and the waitlist control group at pre-intervention (T1) indicated no significant differences in age category, gender, or orphan status. Religious affiliation, ethnicity, and type of education were not statistically analysed due to near-homogeneity of the sample, as almost all participants were Malay Muslims enrolled in public schools. Although a significant difference emerged in prior counselling attendance, most adolescents had participated only once or twice, with fewer than three sessions reported overall.

Table 1. Sociodemographic characteristics of orphaned adolescents at pre-intervention for SAHABAT and waitlist control ($n = 139$)

Socio-demographic characteristics	Total		SAHABAT (<i>n</i> = 71)		Waitlist (<i>n</i> = 68)		Total	Statistical test	<i>p</i>
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>N</i>		
Age (Mean = 14.81, <i>SD</i> = 1.37)									
Age group (years)									
13–15	94	67.6	44	46.8	50	53.2	94	$\chi^2 = 1.62$	.20
16–17	45	32.4	27	60.0	18	40.0	45		
Gender									
Male	76	54.7	43	56.6	33	43.4	76	$\chi^2 = 1.57$	.21

Socio-demographic characteristics	Total		SAHABAT (<i>n</i> = 71)		Waitlist (<i>n</i> = 68)		Total	Statistical test	<i>p</i>
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>N</i>		
Female	63	45.3	28	44.4	35	55.6	63		
Type of orphan									
Single	127	91.4	66	52.0	61	48.0	127	$\chi^2 = .15$	.70
Double	12	8.6	5	41.7	7	58.3	12		
HoCA									
Attended	75	54.0	45	60.0	30	40.0	75	$\chi^2 = 4.44$	.04*
Never attended	64	46.0	26	40.6	38	59.4	64		

Note: Significant at **p* < .05 (2-tailed); HOCA = History of counselling attendance

Table 2 presents baseline comparisons between the SAHABAT intervention and waitlist control groups. Independent samples t-tests indicated no significant differences in pre-intervention (T1) scores for depression ($t = -.15, p = .88$), anxiety ($t = 1.90, p = .06$), negative automatic thoughts ($t = 1.88, p = .06$), self-esteem ($t = .78, p = .44$), or anger ($t = 1.91, p = .06$), confirming equivalence between groups prior to the intervention.

Table 2. Depression, anxiety, negative automatic thoughts, self-esteem, and anger of orphaned adolescents at pre-intervention for SAHABAT and waitlist control (*n* = 139)

Measure	SAHABAT Mean (<i>SD</i>) <i>n</i> = 71	Waitlist Mean (<i>SD</i>) <i>n</i> = 68	Statistical test	<i>p</i>
BDI-Malay	17.73 (9.90)	17.97 (8.92)	$t(137) = -.15$	.88
BAI-Malay	23.44 (10.62)	20.16 (9.66)	$t(137) = 1.90$	.06
ATQ-Malay	39.28 (10.22)	35.68 (10.26)	$t(137) = 1.88$	.06
RSES-Malay	31.86 (3.71)	31.37 (3.72)	$t(137) = .78$	.44
BANI-Y Malay	22.65 (9.24)	19.29 (11.43)	$t(137) = 1.91$	.06

Note: Significant at **p* < .05 (2-tailed).

Table 3 presents the mean depression scores and standard errors for the SAHABAT and waitlist control groups across six measurement points.

Table 3. Means and standard errors of depression, anxiety, negative automatic thoughts, anger, and self-esteem for SAHABAT and waitlist control across time (*n* = 139)

Time	Group	Mean	SE	95% CI	
				LL	UL
BDI-Malay					
T1 Pre-intervention	SAHABAT	18.40	1.54	15.38	21.42
	Waitlist control	18.91	1.64	15.70	22.13
T2 Mid-intervention	SAHABAT	13.24	1.54	10.22	16.26
	Waitlist control	17.03	1.64	13.80	20.25
T3 Post-intervention	SAHABAT	8.60	1.54	5.58	11.62

Time	Group	Mean	SE	95% CI	
				LL	UL
T4 1-month post-intervention	Waitlist control	16.35	1.65	13.11	19.60
	SAHABAT	9.78	1.55	6.74	12.81
T5 3-month post-intervention	Waitlist control	15.37	1.66	12.11	18.63
	SAHABAT	6.56	1.55	3.51	9.61
T6 6-month post-intervention	Waitlist control	14.43	1.69	11.11	17.74
	SAHABAT	6.56	1.67	3.28	9.83
	Waitlist control	14.79	1.71	11.43	18.15
BAI-Malay					
T1 Pre-intervention	SAHABAT	23.73	1.56	20.67	26.79
	Waitlist control	21.83	1.51	18.87	24.79
T2 Mid-intervention	SAHABAT	18.45	1.56	15.39	21.50
	Waitlist control	22.45	1.52	19.47	25.43
T3 Post-intervention	SAHABAT	14.50	1.56	11.44	17.56
	Waitlist control	21.09	1.54	18.07	24.11
T4 1-month post-intervention	SAHABAT	14.72	1.57	11.64	17.81
	Waitlist control	20.80	1.55	17.76	23.54
T5 3-month post-intervention	SAHABAT	10.80	1.58	7.69	13.90
	Waitlist control	19.72	1.59	16.59	22.85
T6 6-month post-intervention	SAHABAT	12.28	1.78	8.79	15.78
	Waitlist control	20.30	1.65	17.07	23.54
ATQ-Malay					
T1 Pre-intervention	SAHABAT	40.25	4.91	30.61	49.88
	Waitlist control	36.39	3.96	28.61	44.16
T2 Mid-intervention	SAHABAT	33.28	4.91	23.64	42.91
	Waitlist control	36.19	3.96	28.41	43.97
T3 Post-intervention	SAHABAT	30.15	4.91	20.51	39.79
	Waitlist control	36.01	3.97	28.22	43.81
T4 1-month post-intervention	SAHABAT	30.82	4.91	21.18	40.47
	Waitlist control	36.87	3.98	29.06	44.69
T5 3-month post-intervention	SAHABAT	30.82	4.91	21.18	40.47
	Waitlist control	36.87	3.98	29.06	44.69
T6 6-month post-intervention	SAHABAT	28.77	4.97	19.02	38.53
	Waitlist control	34.93	4.01	27.07	42.80
BANI-Y Malay					
T1 Pre-intervention	SAHABAT	21.44	2.26	17.01	25.87
	Waitlist control	20.86	1.42	18.08	23.64
T2 Mid-intervention	SAHABAT	15.24	2.26	10.81	19.67
	Waitlist control	20.44	1.43	17.63	23.25
T3 Post-intervention	SAHABAT	13.02	2.26	8.58	17.45
	Waitlist control	20.25	1.44	17.42	23.08
T4 1-month post-intervention	SAHABAT	12.67	2.26	8.23	17.12
	Waitlist control	21.47	1.46	18.61	24.33
T5 3-month post-intervention	SAHABAT	10.14	2.27	5.69	14.60

Time	Group	Mean	SE	95% CI	
				LL	UL
T6 6-month post-intervention	Waitlist control	21.69	1.49	18.76	24.62
	SAHABAT	10.11	2.36	5.48	14.75
	Waitlist control	20.09	1.53	17.08	23.09
RSES-Malay					
T1 Pre-intervention	SAHABAT	32.02	.64	30.77	33.27
	Waitlist control	31.70	.82	30.09	33.31
T2 Mid-intervention	SAHABAT	34.10	.64	32.86	35.35
	Waitlist control	31.87	.82	30.25	33.49
T3 Post-intervention	SAHABAT	34.67	.64	33.42	35.91
	Waitlist control	32.30	.83	30.66	33.93
T4 1-month post-intervention	SAHABAT	34.59	.64	33.32	35.85
	Waitlist control	31.91	.83	30.27	33.54
T5 3-month post-intervention	SAHABAT	36.23	.65	34.95	37.51
	Waitlist control	32.25	.85	30.58	33.92
T6 6-month post-intervention	SAHABAT	36.55	.75	35.09	38.01
	Waitlist control	32.91	.88	31.19	34.63

The results of the generalized linear mixed model (GLMM) analyses, presented in Table 4, indicate that the SAHABAT programme produced significant improvements across multiple psychological outcomes. A significant group $\times$ time interaction was observed for depression [$F(5, 729) = 6.53, p < .001, \eta^2 = .05$; small effect], anxiety [$F(5, 723) = 7.64, p < .001, \eta^2 = .05$; small effect], negative automatic thoughts [$F(5, 734) = 6.89, p < .001, \eta^2 = .05$; small effect], and anger [$F(5, 728) = 10.66, p < .001, \eta^2 = .07$; moderate effect]. These results indicate that the SAHABAT group demonstrated significantly greater reductions in these outcomes over time (T1–T6) compared to the waitlist control group.

Between-subject analyses further confirmed significant group differences for depression [$F(1, 729) = 8.48, p < .001, \eta^2 = .13$; moderate effect], anxiety [$F(1, 723) = 9.77, p < .001, \eta^2 = .07$; moderate effect], negative automatic thought [$F(1, 734) = .34, p < .001, \eta^2 = .00$, small effect], and anger [$F(1, 728) = 9.27, p < .01, \eta^2 = .06$; moderate effect]. These findings highlight the efficacy of the SAHABAT programme relative to the control.

Within-subject analyses revealed significant reductions over time (T1–T6) in depression [$F(5, 729) = 27.12, p < .001, \eta^2 = .17$, large effect], anxiety [$F(5, 723) = 15.02, p < .001, \eta^2 = .10$; moderate effect], negative automatic thoughts [$F(5, 734) = 7.88, p < .001, \eta^2 = .05$, small effect], and anger [$F(5, 728) = 9.93, p < .001, \eta^2 = .07$; moderate effect].

For self-esteem, a significant but small group $\times$ time interaction was observed ($F(5, 730) = 3.30, p < .001, \eta^2 = .02$), indicating that the SAHABAT programme significantly led to greater improvements over time compared to the waitlist control in enhancing self-esteem. Between-subject analyses demonstrated a

moderate effect [$F(1, 730) = 10.65, p < .001, \eta^2 = .07$], while within-subject analyses showed a small but significant effect of time [$F(5, 730) = 7.41, p < .001, \eta^2 = .05$]. Notably, only participants in the SAHABAT group exhibited a significant increase in self-esteem over time (T1-T6).

Table 4. Effectiveness of the SAHABAT and waitlist control on depression across time ($n = 139$)

	Grand mean	SE	95% CI		F	df ₁	df ₂	p	η^2
			LL	UL					
BDI-Malay									
Fixed effect model	13.00	.45	12.12	13.86	47.63	20	729	$p < .001^{**}$	.26
Random effect model	13.33	1.03	11.31	15.36	42.08	20	729	$p < .001^{**}$	.23
Group					8.48	1	729	$p < .01^*$	.13
Time					27.12	5	729	$p < .001^{**}$	.17
Group $\times$ Time					6.53	5	729	$p < .001^{**}$	.05
BAI-Malay									
Fixed effect model	17.92	.55	16.85	19.00	30.12	20	723	$p < .001^*$	.18
Random effect model	18.39	0.94	16.53	20.24	26.15	20	723	$p < .001^*$	.16
Group					9.77	1	723	$p < .001^*$	.07
Time					15.02	5	723	$p < .001^*$	.10
Group $\times$ Time					7.64	5	723	$p < .001^*$	.05
ATQ-Malay									
Fixed effect model	33.40	.59	32.24	34.56	23.57	20	734	$p < .001^*$	.15
Random effect model	34.28	3.10	28.19	40.37	21.94	20	734	$p < .001^*$	.13
Group					.34	1	734	$p < .001^*$	.00
Time					7.88	5	734	$p < .001^*$	.05
Group $\times$ Time					6.89	5	734	$p < .001^*$	.05
BANI-Y Malay									
Fixed effect model	17.91	.50	16.93	18.90	23.87	20	728	$p < .001^{**}$	.15
Random effect model	17.28	1.25	14.84	19.73	19.87	20	728	$p < .001^{**}$	.13
Group					9.27	1	728	$p < .01^*$	.06
Time					9.93	5	728	$p < .001^{**}$	.07
Group $\times$ Time					10.66	5	728	$p < .001^{**}$	.07
RSES-Malay									
Fixed effect model	33.54	.27	33.02	34.07	16.97	20	730	$p < .001^{**}$	.11

	Grand mean	SE	95% CI		F	df ₁	df ₂	p	η^2
			LL	UL					
Random effect model	33.42	.45	32.55	34.30	12.93	20	730	$p < .001^{**}$	.09
Group					10.65	1	730	$p < .01^*$	.07
Time					7.41	5	730	$p < .001^{**}$	.05
Group x Time					3.30	5	730	$p < .01^*$	.02

Note: Significant at $^{**}p < .001$ (2-tailed), $^*p < .01$ (2-tailed).

Figures 2–5 depict the magnitude of symptom reduction within the SAHABAT group from pre-intervention to post-intervention. Mean depression scores decreased from the mild range ($M = 18.40$) to minimal levels ($M = 6.56$), while mean anxiety scores declined from moderate ($M = 23.73$) to mild levels ($M = 12.28$). Negative automatic thoughts were reduced from high ($M = 40.25$) to low ($M = 28.77$), and anger scores decreased from mildly elevated ($M = 21.44$, T -score = 54–57) to the average range ($M = 10.11$, T -score = 42–47).

Figure 6 illustrates a corresponding increase in self-esteem, with mean scores rising from $M = 32.02$ to $M = 36.55$, indicating a significant enhancement that SAHABAT programme in fostering positive self-worth.

Overall, these findings demonstrate that the SAHABAT programme effectively promotes psychological well-being by reducing depression, anxiety, negative automatic thoughts, and anger, while simultaneously enhancing self-esteem. The observed improvements further underscore the intervention's superiority over the waitlist control.

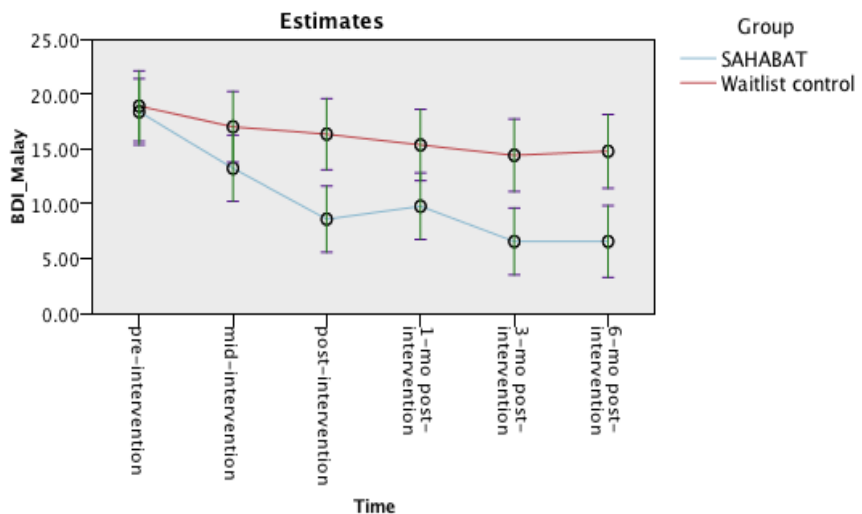


Figure 2. Group and time interaction effect of the SAHABAT and waitlist control on depression across time ($n = 139$)

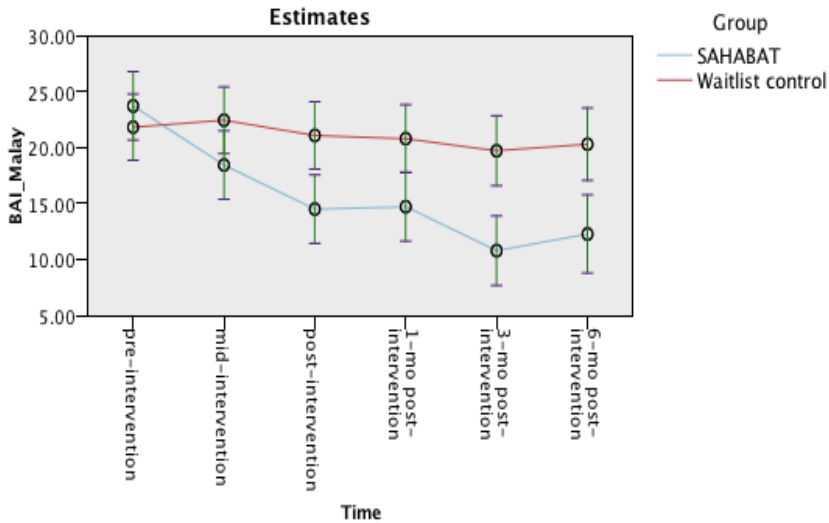


Figure 3. Group and time interaction effect of the SAHABAT and waitlist control on anxiety across time (n = 139)

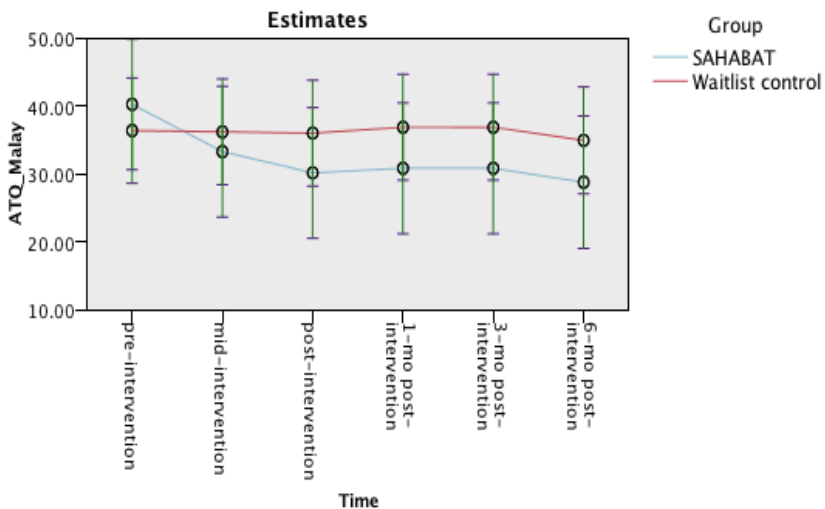


Figure 4. Group and time interaction effect of the SAHABAT and waitlist control on negative automatic thoughts across time (n=139)

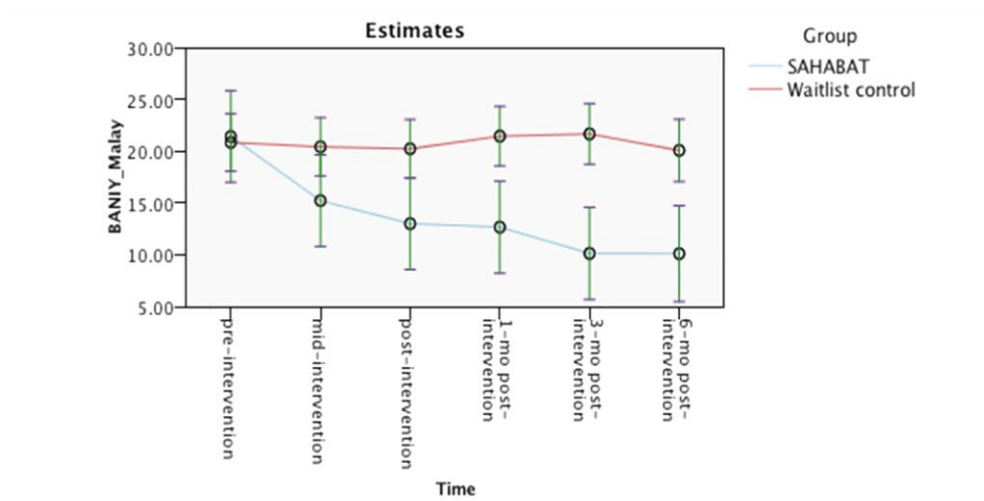


Figure 5. Group and time interaction effect of the SAHABAT and waitlist control on anger across time (n=139)

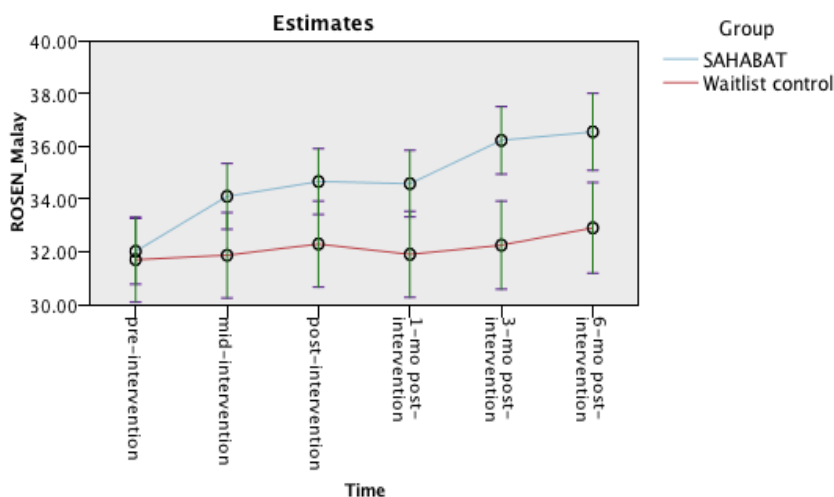


Figure 6. Group and time interaction effect of the SAHABAT and waitlist control on self-esteem across time (n=139)

Discussion

The SAHABAT programme demonstrated a lasting positive impact on the mental health of orphaned adolescents, with benefits sustained for at least six months post- intervention. Compared to the waitlist control, participants in the SAHABAT group exhibited significant reductions in depression, anxiety, and negative automatic thoughts, with small effect sizes, alongside moderate reductions in anger. Self-esteem also improved, albeit with a small effect size, underscoring the intervention's long-term effectiveness.

Consistent with prior research, Keles and Idsøe (2018) noted that while group-based CBT interventions initially yield strong post-intervention effects, these effects may diminish over time. In the present study, significant reductions in depression, anxiety, negative automatic thoughts, and anger, as well as increases in self-esteem, were observed immediately following the intervention (T3, 8th session). These improvements were likely initiated by cognitive restructuring during the mid-intervention phase (T2; 3rd–4th sessions), followed by behavioral strategies implemented between T2 and T3 (5th–6th sessions), which effectively managed symptoms. Reinforcement of cognitive components during the final sessions (T3; 7th–8th sessions) further strengthened the intervention's efficacy, echoing Tang et al.'s (2005) observation that substantial cognitive shifts typically precede symptom reductions after approximately four CBT sessions.

The small effect sizes observed in this study may be partially attributable to the adherence to stringent CONSORT-aligned research protocols and intention-to-treat analyses, which tend to yield conservative estimates of intervention effects (Keles & Idsøe, 2018). Other factors influencing reduced effect sizes include the absence of clinical populations treated with active comparison groups, participant attrition, and limited follow-up assessments (Keles & Idsøe, 2018; Zhou et al., 2015). Moreover, cross-cultural adaptation may play a role in moderating treatment efficacy, as CBT studies conducted outside North America often report smaller effect sizes (Weisz et al., 2006).

Compared to Saw et al. (2020), in which a CBT intervention for secondary school students produced moderate-to-large effect sizes for depression and negative automatic thoughts, the SAHABAT programme yielded smaller gains. This discrepancy may reflect the heightened psychological vulnerability of orphaned adolescents residing in sheltered homes. Nonetheless, the SAHABAT programme demonstrated transdiagnostic efficacy, effectively addressing primary concerns (depression and anxiety) and secondary outcomes (negative automatic thoughts, anger, and self-esteem), consistent with evidence supporting transdiagnostic CBT across diverse populations and formats (Cuijpers et al., 2023; Reinholt & Krogh, 2014; Riise et al., 2021; Zakaria et al., 2021).

RCTs consistently support CBT over passive controls in treating various child and adolescent mental health disorders, including depression and anxiety (Hamed et al., 2020; Keles & Idsøe 2018; Zhou et al., 2015, Wergeland et al., 2021), as well as secondary targets such as negative automatic thoughts (Flores et al., 2024; Mirzaian et al., 2023; Saw et al., 2020; Topper et al., 2023; Elahabadi, 2025), self-esteem (Waite et al., 2012; Wergeland et al., 2021), and anger (Dejid, Delgerjav, & Tsogzolmaa, 2022).

The enduring benefits of SAHABAT align with evidence highlighting the role of booster sessions in consolidating CBT outcomes, especially comparing to passive controls (Gearing et al., 2013; Vittengl et al., 2019). While this study did not employ DSM-5 diagnostic criteria, participants were recruited based on symptom severity, reflecting the programme's dual function as both treatment and prevention. Small effect sizes in CBT-based prevention programs for depression and anxiety have been reported in prior studies (Rasing et al., 2017; Werner-Seidler et al., 2017).

At the 1-month post-intervention assessment in this study, participants exhibited fluctuations in scores, with increases in depression, anxiety, negative automatic thoughts, and anger, accompanied by a decline in self-esteem. These variations were likely due to participant absences during data collection, as many spent the year-end school holidays with relatives and returned to the sheltered homes only after school resumed, approximately six weeks later. Despite these short-term inconsistencies, subsequent assessments at 3 and 6 months post-intervention demonstrated steady decreases in depression, anxiety, and negative automatic thoughts, along with improvements in self-esteem, indicating the sustained effectiveness of the SAHABAT programme.

Cultural and familial factors may constrain the application of CBT to behavioral rather than cognitive domains (Halder & Mahato, 2019); nevertheless, the SAHABAT programme proved effective within the complex familial and Malay cultural context of orphaned adolescents. Notably, the short-term, 8-session intervention was sufficient to achieve meaningful improvements, consistent with prior reports of brief CBT (≤ 10 sessions) effectively treating depression, anxiety, negative automatic thoughts, low self-esteem, and anger (Beidas et al., 2013; Saw et al., 2020; Straub et al., 2014; Sukhodolsky et al., 2016; Waite et al., 2012). Despite challenges inherent to group CBT such as maintaining confidentiality, addressing diverse clinical needs, and logistical coordination, the intervention remains cost-effective and feasible within sheltered homes.

A major strength of this study is its strict adherence to CONSORT standards for quantitative research, which ensured methodological rigor, objectivity, and validity in assessing variables and causal relationships. The study design was carefully structured to address the research questions while minimizing extraneous influences, thereby enabling valid generalizations. The Malay versions of the questionnaires demonstrated strong reliability and validity, supporting the robustness of the measurements. Multi-stage cluster random sampling was employed

to select sheltered homes and allocate participants into the intervention and waitlist control groups, with clearly defined inclusion and exclusion criteria. Additionally, the use of Accuracy in Parameter Estimation (AIPE) enhanced the precision of population estimates, and generalized linear mixed models (GLMM) accounted for variability across homes. The study was conducted in a naturalistic setting, with the SAHABAT programme delivered within participants' sheltered homes, thereby enhancing ecological validity within the RCT framework. Finally, the intervention was delivered by trained clinical psychologists, ensuring fidelity and credibility. Collectively, these methodological strengths provide robust evidence regarding the effectiveness of CBT in alleviating emotional distress among orphaned adolescents in Malaysia.

This study has several limitations. The generalizability of the findings is constrained, as the sample was drawn exclusively from a single state in Malaysia due to logistical and financial constraints. Additionally, the study focused solely on registered non-governmental sheltered homes, excluding unregistered facilities, which may limit the representativeness of the findings for the broader population of sheltered homes in Selangor. Another notable limitation is the absence of active comparison groups. While some studies report that CBT is more effective than active controls (James et al., 2015), others suggest that the inclusion of active control conditions can inflate the perceived efficacy of psychotherapies (Zhou et al., 2015). Although CBT has consistently demonstrated effectiveness in reducing depression and anxiety relative to waitlist controls, its superiority over other psychotherapeutic approaches remains inconclusive (Tolin, 2010). Despite careful sample size calculation, a larger sample would have increased the statistical power and robustness of the analyses. Finally, the follow-up period was limited to six months post-intervention, which may be insufficient to evaluate long-term outcomes, and attrition further constrained the ability to conduct extended follow-up assessments.

Conclusion

The SAHABAT programme represents a culturally adapted, evidence-based CBT intervention that effectively reduces emotional distress and enhances psychological well-being among orphaned adolescents. Its transdiagnostic properties, feasibility, and sustained effects highlight its potential as both a preventive and therapeutic intervention within sheltered home contexts in Malaysia. To the best of the authors' knowledge, this study constitutes the first RCT evaluating group CBT for adolescents with anxiety, low self-esteem, and anger in Malaysia, and the second for adolescent depression. The findings align with existing literature, address theoretical and research gaps, and support the biopsychosocial model and the applicability of CBT in Malaysian populations.

The eight-session SAHABAT programme demonstrated sustained improvements in depression, anxiety, negative automatic thoughts, self-esteem, and anger for at least six months post-intervention. Its user-friendly and cost-effective format, alongside its superiority over waitlist control, underscores its potential for wider implementation in schools and community settings, provided challenges such as the cost of training the therapists, therapist fees, logistical coordination, and participant engagement are addressed.

School counsellors and mental health agencies are encouraged to consider integrating SAHABAT into community-based programmes. Additionally, given the prevalence of mental health issues in non-government-run sheltered homes and the shortage of mental health professionals in Malaysia, urgent measures are needed to expand government-run facilities, enhance support for non-governmental homes, and strengthen foster care systems. The present study contributes to the growing literature on transdiagnostic CBT for underprivileged adolescents in institutional care and provides valuable insights into the mental health of orphaned adolescents in Malay-operated, non-governmental shelters.

Future research should focus on establishing clinical thresholds for mental health assessments among participants to enhance diagnostic accuracy, ideally combining structured diagnostic interviews with validated self-report measures. High-quality RCTs comparing CBT with pharmacological treatments and other psychotherapeutic approaches, including booster sessions, are warranted. Studies with larger sample sizes, extended follow-up periods, and diverse populations such as non-Malay sheltered homes, and diverse settings including schools and community centres, will improve the generalizability of findings. Additionally, comparative studies examining manually delivered CBT versus online CBT, as well as third-wave CBT approaches such as mindfulness-based interventions, are recommended, given the growing evidence supporting online CBT for alleviating depression and anxiety among adolescents.

Collectively, the findings of this study underscore the promise of the SAHABAT programme as a scalable, culturally relevant intervention to promote psychological well-being and resilience among orphaned adolescents in Malaysia.

Authors' Notes

Wai Eng Ding  <https://orcid.org/0000-0001-6434-3406>

Firdaus Mukhtar  <https://orcid.org/0000-0002-6573-7658>

Wai Eng Ding is the founder and clinical director of Ding Psychology Centre, Malaysia.

Firdaus Mukhtar is a Professor at the Department of Medicine & Health Sciences, Universiti Putra Malaysia.

We have no known conflict of interest to disclose.

References

- Ab Ghaffar, S. F., Mohd Sidik, S., Ibrahim, N., Awang, H., & Rampal, L. (2019). Effect of a school-based anxiety prevention program. *International Journal of Environmental Research and Public Health*, 16(24), 4913. <https://doi.org/10.3390/ijerph16244913>
- Anjanappa, S., Govindan, R., Munivenkatappa, M., & Bhaskarapillai, B. (2023). Effectiveness of anger management program on anger level, problem-solving skills, communication skills, and adjustment among school-going adolescents. *Journal of Education and Health Promotion*, 12, Article 90. https://doi.org/10.4103/jehp.jehp_1216_22
- Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). *An inventory for measuring clinical anxiety: Psychometric properties*. *Journal of Consulting and Clinical Psychology*, 56(6), 893–897.
- 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>
- Beidas, R. S., Benjamin, C. L., Puleo, C. M., Edmunds, J. M., & Kendall, P. C. (2013). Flexible applications of the Coping Cat program for anxious youth. *Cognitive and Behavioral Practice*, 20(3), 304–312. <https://doi.org/10.1016/j.cbpra.2012.08.003>
- Chee, H. N., Othman, A., Umar, H., Mohamed Mustafar, M. F. (2023). Assimilating problematic experiences through cognitive behaviour therapy in adolescence with social anxiety disorder: A case study. *Malaysian Journal of Pediatrics and Child Health*, 29(3), 28–32. <https://doi.org/10.51407/mjpc.v29i3.276>
- Cooley-Strickland, M. R., Griffin, R. S., Darney, D., Otte, K., & Ko, J. (2011). Urban African American youth exposed to community violence: A school-based anxiety preventive intervention efficacy study. *Journal of Prevention and Intervention in the Community*, 39(2), 149–166. <https://doi.org/10.1080/10852352.2011.556573>
- Cooper, P., & Jacobs, B. (2011). *An international review of the literature of evidence of best practice models and outcomes in the education of children with emotional disturbance/behavioural difficulties*. National Council for Special Education. https://ncse.ie/wp-content/uploads/2016/08/Research_Report_7_EBD.pdf
- Cuijpers, P., Miguel, C., Harrer, M., Plessen, C. Y., & Karyotaki, E. (2023). Transdiagnostic treatment of depression and anxiety: A meta-analysis. *Psychological Medicine*, 53(14), 6535–6546. <https://doi.org/10.1017/S0033291722003841>
- David, D., Cristea, I., & Hofmann, S. G. (2018). Why cognitive behavioral therapy is the current gold standard of psychotherapy. *Frontiers in Psychiatry*, 9, Article 4. <https://doi.org/10.3389/fpsy.2018.00004>

- Dejid, E., Delgerjav, M., & Tsogzolmaa, D. (2022). Cognitive behavioral therapy-based anger management program among adolescents. *International Journal of Social Science and Humanities Research*, 2(3), 1–12. <https://doi.org/10.12345/ijsshr.v2i3.123>
- Ding, W. E., Tan, K. A., Fam, J. Y., Mukhtar, F., Lye, M. S., Awang, H., & Munawar, K. (2021). Psychometric evaluation of the Malay version of the multidimensional Beck Youth Inventories-2 (BYI-2 Malay): Measuring self-concept, anxiety, depression, anger, and disruptive behaviour among adolescents in sheltered homes. *Pertanika Journal of Social Sciences and Humanities*, 29(1), Article 15. <https://doi.org/10.47836/pjssh.29.1.15>
- Elahabadi, F. F. (2025). Effectiveness of CBT on self-harming thoughts, impulsivity, and suppressed anger in adolescents with depression syndrome. *Journal of Adolescent and Youth Psychological Studies*, 6(2), 90–100. <https://doi.org/10.61838/kman.jayps.6.2.11>
- Florea, I. S., Botis, S., Călinoiu, A., Chereji, S. V., Szentagotai-Tătar, A., & Opre, A. (2024). The mediator role of negative and positive automatic thoughts in the relationship between attachment and internalizing problems in adolescents. *Children*, 11(4), 447. <https://doi.org/10.3390/children11040447>
- Gallegos-Guajardo, J., Ruvalcaba-Romero, N. A., Langley, A., & Villegas-Guinea, D. (2015). Selective prevention for anxiety and resilience promotion: Outcomes of an anxiety prevention and resilience program with girls at risk. *Pensando Psicología*, 11(18), 11–23. <https://doi.org/10.16925/pe.v11i18.1001>
- Gearing, R. E., Schwalbe, C. S. J., Lee, R. H., & Hoagwood, K. E. (2013). The effectiveness of booster sessions in CBT treatment for child and adolescent mood and anxiety disorders. *Depression and Anxiety*, 30(9), 800–808. <https://doi.org/10.1002/da.22118>
- Halder, S., & Mahato, A. K. (2019). Cognitive behavior therapy for children and adolescents: Challenges and gaps in practice. *Indian Journal of Psychological Medicine*, 41(3), 279–283. https://doi.org/10.4103/IJPSYM.IJPSYM_470_18
- Hamedi, V., Hamid, N., Beshlideh, K., Marashi, S. A., & Hashemi Sheikh Shabani, S. E. (2020). Effectiveness of conventional cognitive-behavioral therapy and its computerized version on reduction in pain intensity, depression, anger, and anxiety in children with cancer: A randomized, controlled trial. *Iranian Journal of Psychiatry and Behavioral Sciences*, 14(4), e83110. <https://doi.org/10.5812/ijpbs.83110>
- Hollon, S. D., & Beck, A. T. (2013). Cognitive and cognitive-behavioral therapies. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 393–446). Hoboken, NJ: Wiley.
- Hollon, S. D., & Kendall, P. C. (1980). *Cognitive self-statements in depression: Development of an automatic thoughts questionnaire*. *Cognitive Therapy and Research*, 4, 383–393.
- Institute for Public Health. (2022). *National Health and Morbidity Survey (NHMS) 2022: Adolescent Health Survey*. Ministry of Health Malaysia. https://iku.gov.my/images/nhms-2022/Report_Malaysia_nhms_ahs_2022.pdf

- Jayakumar, J. J. D. (2023). *Urgent Need for Protection and Welfare of Institutionalised Children in Malaysia*. EMIR Research. <https://www.emirresearch.com/urgent-need-of-protection-and-welfare-of-institutionalised-children-in-malaysia/>
- James, A. C., James, G., Cowdrey, F. A., Soler, A., & Choke, A. (2015). Cognitive behavioural therapy for anxiety disorders in children and adolescents. *Cochrane Database of Systematic Reviews*, (2), CD004690. <https://doi.org/10.1002/14651858.CD004690.pub4>
- Keles, S., & Idsøe, T. (2018). A meta-analysis of group cognitive behavioral therapy (CBT) interventions for adolescents with depression. *Journal of Adolescence*, 67(May), 129–139. <https://doi.org/10.1016/j.adolescence.2018.05.011>
- Kreuze, L. J., Pijnenborg, G. H. M., de Jonge, Y. B., & Nauta, M. H. (2018). Cognitive-behavior therapy for children and adolescents with anxiety disorders: A meta-analysis of secondary outcomes. *Journal of Anxiety Disorders*, 60, 43–57. <https://doi.org/10.1016/j.janxdis.2018.10.005>
- Matthys, W., & Schutter, D. J. L. G. (2021). Increasing effectiveness of cognitive behavioral therapy for conduct problems in children and adolescents: What can we learn from neuroimaging studies? *Clinical Child and Family Psychology Review*, 24, 484–499. <https://doi.org/10.1007/s10567-021-00346-4>
- Mirzaian, N., Vafaeinejad, Z., Zamanpour, Z., Shabannejad, A., & Abbasian, A. (2023). The effectiveness of cognitive-behavioral therapy (CBT) on rumination in self-harming adolescents without suicidal intent. *Journal of Adolescent and Youth Psychological Studies*, 4(3), 119–125. <https://doi.org/10.61838/kman.jayps.4.3.11>
- Mohammadzadeh, M., Awang, H., Ismail, S., & Shahr, H. K. (2020). Improving coping mechanisms of Malaysian adolescents living in orphanages through a life skills education program: A multicentre randomized controlled trial. *Asian Journal of Psychiatry*, 48, 101892. <https://doi.org/10.1016/j.ajp.2019.101892>
- Mohammadzadeh, M., Awang, H., Kadir Shahar, H., & Ismail, S. (2018). Emotional health and self-esteem among adolescents in Malaysian orphanages. *Community Mental Health Journal*, 54(1), 117–125. <https://doi.org/10.1007/s10597-017-0128-5>
- Mohd Jamil, M. Y. (2006). Validity and reliability study of Rosenberg Self-Esteem Scale in Seremban school children. *Malaysian Journal of Psychiatry*, 15(2), 35–39.
- Mukhtar, F. (2014). *SAHABAT: A group cognitive-behavioral therapy program for depressed and anxious adolescents in Malaysia* [Unpublished manuscript]. Universiti Putra Malaysia.
- Mukhtar, F., Ding, W. E., Lye, M. S., Awang, H., & Munawar, K. (2021). Prevalence and associated factors of depression and anxiety of adolescents residing in Malay-operated non-government-run sheltered homes in Selangor, Malaysia. *Pertanika Journal of Social Sciences and Humanities*, 29(1), Article 16. <https://doi.org/10.47836/pjssh.29.1.16>

- Mukhtar, F., & Oei, T. P. S. (2008). Exploratory and confirmatory factor validation and psychometric properties of the Beck Depression Inventory for Malays (BDI-Malay) in Malaysia. *Malaysian Journal of Psychiatry*, 17(1), 51–64.
- Mukhtar, F., & Zulkefly, N. S. (2011). The Beck Anxiety Inventory for Malays (BAI-Malay): A preliminary study on psychometric properties. *Malaysian Journal of Medicine and Health Sciences*, 7(1), 75–81.
- Naim, R.; Dombek, K.; German, R. E.; Haller, S. P.; Kircanski, K.; and Brotman, M. A. (2023). An exposure-based cognitive-behavioral therapy for youth with severe irritability: Feasibility and preliminary efficacy. *Journal of Clinical Child & Adolescent Psychology*. <https://doi.org/10.1080/15374416.2023.2264385>
- Oei, T. P. S., & Mukhtar, F. (2008). Exploratory and confirmatory factor validation and psychometric properties of the Automatic Thoughts Questionnaire for Malays (ATQ- Malay) in Malaysia. *Hong Kong Journal of Psychiatry*, 18(3), 92–100.
- OrphanCare Foundation. (2020). *The problem of institutionalisation: Every day is one too many*. <https://orphancare.org.my/the-problem-of-institutionalisation-everyday-is-one-too-many/>
- Perera, H., Jeewandara, K. C., Seneviratne, S., & Guruge, C. (2020). Effectiveness of a school-based cognitive–behavioral intervention program for symptoms of depression, anxiety, and low self-esteem in adolescents: A randomized controlled trial in Sri Lanka. *Child and Adolescent Psychiatry and Mental Health*, 14(1), 30. <https://doi.org/10.1186/s13034-020-00339-3>
- Rasing, S. P. A., Creemers, D. H. M., Janssens, J. M. A. M., & Scholte, R. H. J. (2017). Depression and anxiety prevention based on cognitive behavioral therapy for at-risk adolescents: A meta-analytic review. *Frontiers in Psychology*, 8, Article 1066. <https://doi.org/10.3389/fpsyg.2017.01066>
- Richard, Y., Tazi, N., Frydecka, D., Hamid, M. S., & Moustafa, A. A. (2022). A systematic review of neural, cognitive, and clinical studies of anger and aggression. *Current Psychology*. Advance online publication. <https://doi.org/10.1007/s12144-022-03143-6>
- Reinholt, N., & Krogh, J. (2014). Efficacy of transdiagnostic cognitive behaviour therapy for anxiety disorders: A systematic review and meta-analysis of published outcome studies. *Cognitive Behaviour Therapy*, 43(3), 171–184. <https://doi.org/10.1080/16506073.2014.897367>
- Riise, E. N., Wergeland, G. J. H., Njardvik, U., & Öst, L.-G. (2021). Cognitive behavior therapy for externalizing disorders in children and adolescents in routine clinical care: A systematic review and meta-analysis. *Clinical Psychology Review*, 83, 101954. <https://doi.org/10.1016/j.cpr.2020.101954>
- Rodgers, A., & Dunsmuir, S. (2015). A controlled evaluation of the “FRIENDS for life” emotional resiliency programme on overall anxiety levels, anxiety subtype levels, and school adjustment. *Child and Adolescent Mental Health*, 20(1), 13–19. <https://doi.org/10.1111/camh.12030>

- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Saw, J. A., Tam, C. L., Thanzami, V., & Bonn, G. (2020). Contextualized school-based cognitive behavioral therapy (CBT) intervention for Malaysian secondary school students. *Frontiers in Psychiatry*, 11, 565896. <https://doi.org/10.3389/fpsy.2020.565896>
- Sharif, F., Hatami, M., Saberi, Z., & Vossoughi, M. (2021). The effect of cognitive-behavioral group therapy on self-esteem and mental health of female adolescents: A randomized controlled trial. *BMC Psychology*, 9(1), 54. <https://doi.org/10.1186/s40359-021-00565-8>
- Straub, J., Sproeber, N., Plener, P. L., Fegert, J. M., Bonenberger, M., & Koelch, M. G. (2014). A brief cognitive-behavioural group therapy programme for the treatment of depression in adolescent outpatients: A pilot study. *Child and Adolescent Psychiatry and Mental Health*, 8(9), 1–10. <https://doi.org/10.1186/1753-2000-8-9>
- Sukhodolsky, D. G., Smith, S. D., McCauley, S. A., Ibrahim, K., & Piasecka, J. B. (2016). Behavioral interventions for anger, irritability, and aggression in children and adolescents. *Journal of Child and Adolescent Psychopharmacology*, 26(1), 58-64. <https://doi.org/10.1089/cap.2015.0120>
- Tang, T. Z., Beberman, R., DeRubeis, R. J., & Pham, T. (2005). Cognitive changes, critical sessions, and sudden gains in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, 73(1), 168-172. <https://doi.org/10.1037/0022-006X.73.1.168>
- Tolin, D. F. (2010). Is cognitive-behavioral therapy more effective than other therapies? A meta-analytic review. *Clinical Psychology Review*, 30(6), 710–720. <https://doi.org/10.1016/j.cpr.2010.05.003>
- Topper, M., Emmelkamp, P. M. G., & Bögels, S. M. (2023). Mechanisms of change in cognitive behavioral therapy for adolescent depression: The role of negative automatic thoughts. *Behaviour Research and Therapy*, 168, 104390. <https://doi.org/10.1016/j.brat.2023.104390>
- Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: A comparative meta-analysis of cognitive- behavioral therapy's effects. *Journal of Consulting and Clinical Psychology*, 75(3), 475–488. <https://doi.org/10.1037/0022-006X.75.3.475>
- Waite, P., McManus, F., & Shafran, R. (2012). Cognitive behavior therapy for low self-esteem: A preliminary randomized controlled trial in a primary care setting. *Journal of Behavior Therapy and Experimental Psychiatry*, 43(4), 1049–1057. <https://doi.org/10.1016/j.jbtep.2012.04.006>
- Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin*, 132(1), 132–149. <https://doi.org/10.1037/0033-2909.132.1.132>
- Wergeland, G. J. H., Fjermestad, K. W., Marin, C. E., Haugland, B. S. M., Bjaastad, J. F., Oeding, K., Bjelland, I., Havik, O. E., Heiervang, E. R., & Silverman, W. K. (2021). Change in quality of life and self-esteem in a randomized controlled

- trial targeting anxious and depressive symptoms in school children. *Journal of Anxiety Disorders*, 78, 102345. <https://doi.org/10.1016/j.janxdis.2021.102345>
- Werner-Seidler, A., Perry, Y., Cascar, A. L., Newby, J. M., & Christensen, H. (2017). School- based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. *Clinical Psychology Review*, 51, 30–47. <https://doi.org/10.1016/j.cpr.2016.10.005>
- Zakaria, S., Sasagawa, S., & Essau, C. A. (2021). Exploring the impact of a transdiagnostic cognitive behavioural therapy-based intervention on a group of Malaysian adolescents with problematic drug use and emotional problems. *Addictive Behaviours Reports*, 14, 100381. <https://doi.org/10.1016/j.abrep.2021.100381>
- Zhou, X., Hetrick, S. E., Cuijpers, P., Qin, B., Barth, J., Whittington, C. J., Cohen, D., Del Giovane, C., Liu, Y., Michael, K. D., Zhang, Y., & Xie, P. (2015). Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. *World Psychiatry*, 14(2), 207-222. <https://doi.org/10.1002/wps.20217>

IMPROVING STUDENTS' SELF-ESTEEM WITH AN ONLINE GROUP INTERVENTION BASED ON THE FENNEL MODEL: AN OPEN-LABEL, UNCONTROLLED FEASIBILITY TRIAL

Luana M. Alexa¹, Andrei Rusu¹*

¹ Department of Psychology, West University of Timișoara, Romania

Abstract

Increasing self-esteem through an online intervention may be essential for students' psychological well-being and academic performance. Based on Melanie Fennell's cognitive model (1997, 2016), this study investigates the feasibility of an online intervention program aimed at improving students' self-esteem. After completing online screening, 33 students ($M_{age} = 22$ years, $SD_{age} = 6.75$; 70% females) with low or medium levels of self-esteem participated in this open-label uncontrolled pilot trial. The primary outcome measure was the Rosenberg Self-Esteem Scale (RSE). As part of the intervention program, we utilized psychoeducation and cognitive-behavioral techniques. Furthermore, due to the pandemic context in which the intervention's feasibility was tested, we delivered the intervention online, in working groups. Following this five-week program, participants exhibited higher levels of explicit self-esteem ($d = .74$) and lower levels of depression and anxiety as secondary outcomes ($d = .46$ for depression and $d = .48$ for anxiety). Despite inherent limitations, these results provide initial evidence supporting the feasibility of an online working group intervention aimed at improving students' self-esteem.

Keywords: Self-esteem, Internet-delivered intervention, Anxiety, Depression, Working groups

In the context of personal and professional development, as well as overall life satisfaction, numerous previous studies have emphasized the importance of maintaining a healthy level of self-esteem (Arslan, 2019; Coffey & Warren, 2020; Hawi & Samaha, 2017; Orth & Robins, 2022). This concept has been extensively

* Correspondence concerning this article should be addressed to Luana Alexa, West University of Timișoara, Department of Psychology, Bd. Vasile Pârvan no. 4, Timișoara, 300223, Timiș county, Romania.

Phone number: +40735722069

E-mail: alexa.luana@e-uvt.ro

studied over the years due to its influence on various aspects of life, including academic achievement (Giofrè et al., 2017). Furthermore, Rossi et al. (2020) highlighted how self-esteem can act as a buffer against mental health challenges, such as fear and loneliness. Despite the consistent emphasis on the significance of self-esteem, there is reduced evidence supporting online interventions aimed at enhancing students' self-esteem, particularly in digital learning environments. However, considering both short and longer-term implications, improving students' self-esteem should constantly be on the headline of the scholarly focus.

In the short term, the levels of self-esteem among students play a pivotal role in their academic success (Di Giunta et al., 2013; Keltikangas-Järvinen, 1992; Marsh et al., 2006; Pullmann & Allik, 2008). This aspect is of high actuality, since unfortunately, the percentage of individuals who successfully complete their higher education remains alarmingly low. According to Eurostat (2020), in Romania, only 25% of the 25-34-year-old population has achieved a high level of educational attainment, which falls far short of the EU-level target of 40.5%. Looking at the medium term, self-esteem significantly impacts career decision-making. Research by Javed and Tariq (2016) has shown that higher self-esteem is associated with fewer difficulties in making career decisions. In the long term, multiple previous studies have underscored the impact of self-esteem on mental health (Hiçdurmaz & Karahan, 2017; Sowislo & Orth, 2013; Triana et al., 2019). Additionally, students with higher scores on mental health assessments tend to perform less well academically across various measures (Jeffries & Salzer, 2020). In today's world, it is more important than ever to examine students' self-esteem due to the increased influence of social media and its impact on personal self-concept, especially since the COVID-19 pandemic (Polyakova et al., 2024). Jan et al. (2017) noted that spending just one hour on Facebook daily results in a 5.574-point decrease in an individual's self-esteem score.

Regarding existing interventions for boosting self-esteem, previous research has demonstrated their effectiveness. These interventions include mindfulness-based approaches (Randal et al., 2015), physical activity interventions (Liu et al., 2015), support groups (Adamson et al., 2019), and cognitive-behavioral therapy (CBT) interventions (Kolubinski et al., 2018; Morton et al., 2012; Swartzman et al., 2021). However, the unique challenges posed by the COVID-19 pandemic, such as social isolation and remote work environments, require a more tailored approach. Additionally, pandemic-related effects, such as increased online activity, heightened anxiety, and depressive symptoms (Hu et al., 2022), need specific attention. Therefore, while previous interventions have proven effective, the recent context demands a more customized approach.

In our study, we adopted the framework provided by the Fennell (1997, 2016) model for boosting self-esteem (see Appendix). This model was selected because it targets the cognitive processes that underlie low self-esteem, addressing the mechanisms that sustain negative self-beliefs rather than focusing solely on

clinical symptoms. This process-oriented focus makes the model particularly relevant for our non-clinical population, which exhibits subclinical low self-esteem, depression and anxiety rather than full-blown psychiatric conditions. Notably, applying the Fennell model in a non-clinical population represents a novel extension of its use, as most previous research has focused on clinical samples, and enables preventive as well as enhancement-focused interventions. In addition, having access to the worksheets and exercises from the author's self-help guide was an added advantage, as they facilitated the practical application of the model and offered participants concrete tools for self-reflection and practicing the proposed strategies. While this theoretical approach's efficacy was supported by a systematic review in 2018 (Kolubinski et al., 2018), the evidence was primarily derived from clinical populations.

The Fennell model posits that individuals form negative beliefs and assumptions about themselves and others based on early experiences, which Fennell termed "the bottom line." To protect their psychological well-being, people develop "rules for living" (e.g., "I must be perfect" or "I must be loved by everyone"). However, certain trigger situations challenge or break these rules and activate the "bottom line" (e.g., receiving a low grade or facing rejection). This activation leads to automatic negative beliefs, emotions (e.g., depression, anxiety), and unhelpful behaviors (e.g., avoidance or withdrawal). These emotions and behaviors, in turn, reinforce the negative thoughts, creating a vicious cycle. While this model forms the basis for CBT interventions aimed at improving self-esteem, its efficacy remains limited in many studies (Kolubinski et al., 2018; Taylor & Montgomery, 2007). This limitation may stem from the fact that many CBT interventions primarily target reducing clinical symptoms (e.g., anxiety and depression), with self-esteem improvement considered a secondary outcome (Crisp et al., 2014). Consequently, most research samples consist of clinical patients.

Chandra et al. (2019) emphasized the lack of methods to boost students' self-confidence and demonstrated the efficacy of self-instruction and cognitive restructuring techniques in group support settings. A systematic review by Suh et al. (2019) found no significant differences in outcomes between online and face-to-face delivery modalities in decreasing perfectionism dimensions, depression, and anxiety. Additionally, the shift to online teaching reduced opportunities for first-year students to establish strong relationships with peers and teachers. A recent meta-analysis by Harris and Orth (2020) confirmed the importance of social relationships in influencing self-esteem, leading us to opt for delivering our intervention in working groups.

Given the dearth of interventions designed specifically for students and considering the unique circumstances brought about by the COVID-19 pandemic, we developed an intervention aimed at enhancing students' self-esteem, which can be implemented today thanks to the accessibility provided by the online format. Moreover, Peyton et al. (2022) conclude that the pandemic does not pose a

fundamental threat to the generalizability of results from online experiments to other time periods. The five intervention modules are based on Melanie Fennell's cognitive behavioral techniques and exercises proposed in her self-help guide for improving self-esteem (Fennell, 2016), and were adapted for the online group format.

To assess the feasibility of this intervention, we conducted an open-label uncontrolled trial to measure participants' self-esteem before and after a five-week period during which they participated in one of our online working groups led by a clinical psychologist.

Therefore, we started our study with the following hypothesis:

H1: After the intervention, participants will report a significantly higher level of self-esteem, both (H1a) explicit and (H1b) implicit, compared to the initial assessment.

H2: After the intervention, participants will report significantly lower levels of depressive symptoms, compared to the initial assessment.

H3: After the intervention, participants will report significantly lower levels of anxiety symptoms, compared to the initial assessment.

H4: After the intervention, participants will report a significantly lower level of academic burnout, compared to the initial assessment.

H5: After the intervention, participants will report a significantly higher level of academic self-esteem, compared to the initial assessment.

Method

Study design

Through a one-arm, open-label, uncontrolled study, we assessed the feasibility of an online working group intervention aimed at enhancing students' self-esteem. The primary outcome targeted in the intervention was self-esteem, measured both explicitly and implicitly. Additionally, data on depressive symptoms, anxiety symptoms, academic burnout, and academic self-esteem were collected from students at the beginning and the end of the intervention.

Intervention programme

The group-based intervention was delivered by a clinical psychologist affiliated with the university, ensuring that the program was implemented by a professional with expertise. The five intervention modules were based on Melanie Fennell's cognitive-behavioral techniques and exercises proposed in her self-help guide for improving self-esteem (Fennell, 2016) and were adapted for the online group format. For each module, a detailed session plan was prepared, including exercises, timing, and discussion topics, to guide implementation and ensure

consistency across sessions. Session plans and related worksheets are available in Romanian upon request.

In the initial modules, our emphasis was on psychoeducation regarding the mechanisms involved in forming and maintaining low self-esteem. We introduced Fennell's model (Fennell, 1997; 2016) and provided college-specific examples. For instance, when discussing triggering situations, we referenced scenarios such as peer comparison, receiving low grades, or experiencing romantic rejection. Inspired by the Socratic Questioning method (Beck, 2011), in the online meetings we used open-ended questions to explore students' thoughts and self-value condition. As part of their homework between sessions, students were assigned specific worksheets to help them better understand their own vicious cycles of maintaining low self-esteem.

The concluding modules concentrated on combating self-criticism and fostering self-compassion, primarily through the utilization of the double standard technique (Leahy, 2003). Additionally, to address self-negative thoughts, students were instructed to recognize the negative beliefs underlying their conditions of worth, reassess them, and present new evidence supporting more adaptive beliefs (cognitive restructuring) using specific worksheets. All of these materials have been translated into Romanian from Melanie Fennell's guide (2016) and can be made available upon request.

A more detailed overview of the intervention is provided in the Appendix.

Participants and procedure

Following Billingham et al. (2013) recommendations, we estimated an initial sample size of 25 for adequate statistical power ($1-\beta = 0.80$) in performed paired samples t-tests for a moderate effect size ($d = 0.50$).

An email sent to all first-year students through the institutional address by the Career Counseling and Guidance Centre of the West University of Timișoara invited them to enroll in an online self-esteem-boosting program, by completing a Google forms with a force response setting activated in order to prevent missing data, resulting in 119 participants. After the registration period, the next step was to select participants for the online intervention. Following the data analysis of the initial screening, 65 students were invited to join one of the working groups, while the other 54 were invited to individual counseling sessions.

Participants' selection was based on the following eligibility criteria: low or moderate levels of self-esteem (i.e., scores <35 on the Rosenberg Self-Esteem Scale, RSES; Rosenberg, 1965) and low to moderate levels of depression and anxiety (i.e., scores <15 on the Patient Health Questionnaire 9, PHQ9; Kroenke et al., 2001, and scores <15 on the GAD-7; Spitzer et al., 2006). Students who did not meet these inclusion criteria (those with high self-esteem scores (ceiling effect) or clinical levels of depression or anxiety) were excluded from the group intervention. For ethical reasons, excluded students were not left without support. They were referred to

individual counseling services, where they could receive appropriate psychological assistance tailored to their needs. This measure ensured that all students had access to adequate care, even if they were not eligible for participation in the feasibility study.

Although initially, there were 65 eligible students, 32 dropped out of the program immediately after screening (they did not log in to any session), resulting in a final sample of 33 first-year students ($M_{age} = 21.61$, $SD_{age} = 6.75$, 70% females), distributed among 4 working groups, with a maximum of 9 participants in each. Figure 1 illustrates the flow of participants through the study.

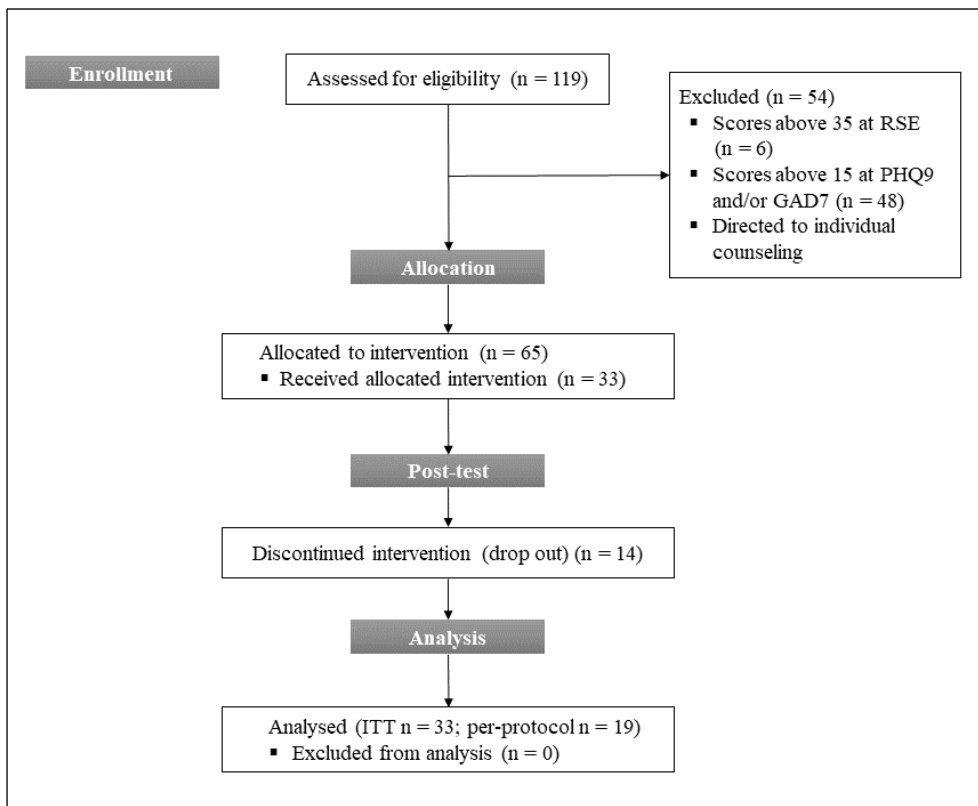


Figure 1. The flow of participants through the study

There was a weekly meeting with each working group for 1 hour and a half using the Zoom platform. The same psychologist coordinated all working groups to ensure delivery integrity for the intervention. At the end of the five weeks, 19 (42% dropout rate) participants completed the final set of questionnaires (post-test). The intervention program is described in Appendix.

Measures

Previous studies that used Fennell's model to tailor their intervention measured three primary psychological outcomes: self-esteem, depression, and anxiety (Brown et al., 2004; Rigby & Waite, 2006; Horrell et al., 2014; Pack & Condren, 2014; Beattie & Beattie, 2018; Berg et al., 2022).

Given the specificity of our sample, we also considered the three above-mentioned psychological outcomes, as well as academic self-esteem and academic burnout. Moreover, to our knowledge, no previous study that relies on Fennell's model considered implicit self-esteem as an outcome of the program. Given all these, we chose to include in our study the following scales:

Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965) was used to measure explicit self-esteem. This scale was culturally adapted on the Romanian population, being part of the Clinical Assessment System instrument (David, 2020), and it is widely used on Romanian samples with optimal psychometric properties. The instrument contains ten items to which participants respond on a scale from 1 (total disagreement) to 4 (total agreement). A higher score represents a higher level of self-esteem. For this study, the scale had good internal consistency, Cronbach's $\alpha = .80$ pre-test and $.86$ post-test.

Name Letter Test (Nuttin, 1987) was used to measure implicit self-esteem. The test involves calculating the difference between the assessed attractiveness of a person's initials and the attractiveness of the same letters evaluated by others whose names do not contain those letters (Albers et al., 2009). The NLT had proper psychometric indices in previous studies conducted on Romanian undergraduate students, including an investigation in relation to rationality (Sava et al., 2011). On the sample of this study, the internal consistency was excellent, with Cronbach's $\alpha = .95$ pre-test and $.97$ post-test.

General Anxiety Disorder Scale - 7 (GAD7; Spitzer et al., 2006) was used to measure anxiety symptoms. The scale was validated for the Romanian population (Cotiga et al., 2023), as also successfully used in previous studies testing internet interventions for the local population (e.g., Isbășoiu et al., 2021). The instrument contains seven items to which participants respond on a scale from 0 (not at all) to 3 (nearly every day). A higher score represents an increased level of anxiety symptoms. On the sample of this study, the scale had acceptable internal consistency, with Cronbach's $\alpha = .72$ pretest and $.77$ post-test.

Patient Health Questionnaire Scale - 9 (PHQ9; Kroenke et al., 2001) was used to measure the severity of depressive symptoms. The scale was previously adapted to Romanian Type 2 Diabetes patients (Lupascu et al., 2019) and was also successfully used in previous studies on internet interventions with Romanian participants (e.g., Isbășoiu et al., 2021). The instrument contains nine items to which participants respond on a scale from 0 (not at all) to 3 (nearly every day), each

corresponding to a diagnostic criterion from DSM-IV. A higher score represents an increased level of depressive symptoms. On the sample of this study, the scale had poor internal consistency, Cronbach's $\alpha = .36$ pre-test and $.61$ post-test.

Maslach Burnout Inventory-Student Survey (MBI-SS; Schaufeli et al., 2002) is a 16-item scale for measuring academic burnout. The instrument was previously used on Romanian undergraduate students, exhibiting proper psychometric properties (Paloş et al., 2019). Participants respond on a scale 7 points Likert scale, from 0 (never) to 6 (every day), measuring three dimensions of academic burnout, namely exhaustion ($\alpha = .76$ pre-test and $.90$ post-test for this study), cynicism ($\alpha = .64$ pre-test and $.88$ post-test for this study) and reduced efficacy ($\alpha = .78$ pre-test and $.89$ post-test for this study).

Bachman's School Ability Self-Concept Index (Bachman, 1970) was used to measure specific (academic) self-esteem and consists of participants' answers to 3 questions. For the first two („How do you rate yourself in school ability compared with your peers?” and “How intelligent do you think you are compared to others your age?”), participants answered on a 6-step Likert scale, where 1 = far below average and 6 = far above average. For the 3rd question („Compared to others your age, how important is it to you to be able to use your intelligence?”), participants answered on a 5-step Likert scale, where 1 = much less important than average and 5 = much more important than average. A similar way of measuring specific self-esteem has been used by Rosenberg et al., (1995). Since this instrument was not already translated into Romanian, we used the standard back-translation approach to obtain the Romanian version of its items. In the present study, we had to eliminate item 3 to increase the scale's internal consistency (from $.50$ to $.94$). Thus, for the remaining two items, the scale had good/excellent internal consistency, Cronbach's $\alpha = .79$ pre-test and $.94$ post-test.

Descriptive therapeutic factors. Given that the intervention was delivered in working groups, we considered it relevant to have an overview of particular factors that may improve personal indices in addition to the actual exercises of the intervention. Yalom and Leszcz (2005) point out the presence of 11 primary factors that facilitate the therapeutic process, namely: instillation of hope, universality, imparting information, altruism, the corrective recapitulation of the primary family group, development of socializing techniques, imitative behavior, interpersonal learning, group cohesiveness, catharsis, and existential factors. However, assuming the working group approach, not the psychotherapeutic one, the presence of some of these 11 factors could not be possible (e.g., corrective recapitulation of the primary family group). Thus, starting from a Q-sort with 60 items used to highlight the most relevant therapeutic factors used by the authors mentioned above (Yalom & Leszcz, 2005), we selected six items pertinent to the present study, each highlighting a factor (in the original research, each factor corresponds to 5 items, and the factor development of socialization techniques is divided into 2, resulting in 12 categories for Q-sort). These were *group cohesiveness* (belonging to and being accepted by a

group), *universality* (learning I am not the only one with my type of problem; “We are all in the same boat”), *guidance* (group members suggesting or advising something for me to do), *catharsis* (getting things off my chest), *self-understanding* (find out some causes and sources of my problems), *inoculating hope* (see how others solve problems similar to mine). In addition, we added three more items, each corresponding to some of the most used methods of intervention in the therapeutic process: *psychoeducation* (theoretical aspects presented during the meetings), *homework* (work materials between meetings), and *therapeutic relationship* (attitude of the group coordinator). For each of these nine factors, students had to rate from 1 to 5 the degree to which that factor contributed to the perceived increase in self-esteem (where 1 = did not contribute at all, and 5 = contributed a lot). Since this is not a validated instrument, we developed and used it only for a descriptive purpose.

System Usability Scale (SUS; Brooke, 1996) is a 10-item scale that provides an overview of subjective usability evaluations, used in this case to evaluate the Zoom platform, on a 5-step Likert scale (where 1 = strongly disagree and 5 = strongly agree). On the sample of this study, the scale shows good internal consistency, with Cronbach’s $\alpha = .89$. It was previously used in a feasibility study for an internet intervention delivered to Romanian participants (Bodnaru et al., 2023). For this scale, the raw scores have to be standardized as follows: for odd items, 1 has to be extracted from the raw score, and for even ones, the raw score has to be extracted from number 5. The total score, multiplied by 2, has to be reported to 100 (Lewis, 2018). Values between 0 and 64 show that the platform used is unacceptable to users, values between 65 and 84 reflect the acceptance of the system by users, and values between 85 and 100 mean that users perceive the system/platform used as excellent (McLellan et al., 2012).

Finally, we asked our participants to rate on a 5-point scale their satisfaction with the program, the quality of information, and the pace of the intervention, where 1 stands for very unhappy/poor/slow and 5 for very satisfied/good/long.

Data analysis

The statistical analysis included data from 33 participants involved in at least one meeting. Given the large number of withdrawn participants (14 out of 33), we used both *per-protocol* (PP) and *intent-to-treat* (ITT) analyses. The ITT principle means that all participants are included in the final analysis, regardless of their path to the intervention. The missing data of those who withdrew are supplemented based on the *last observation carried forward* (LOCF) principle, assuming that outcomes from withdrawn participants did not change from the last observation until the end of treatment. One-tailed paired sample t-tests were used to analyze pre- to post-intervention changes for all outcomes.

Ethics and approval

The University Ethics Committee approved the study (57758/10.11.2020) before its initiation. The study was pre-registered in the ClinicalTrials.gov PRS (NCT04668287).

Results*Dropout*

A total of 33 participants registered at least once in a working group session. Of these, 19 completed the final set of questionnaires, resulting in a 42% dropping rate. We compared the baseline data of those who finished the intervention with those who dropped out. The only significant difference between the two groups was for gender. In the completers group, 53% were females, compared with the dropouts, where 93% were females (see Table 1, for all baseline comparisons). No additional data or feedback was obtained from the 14 participants who did not complete the final set of questionnaires.

Table 1. Sample characteristics

Variable	Total ($N = 33$)	Comparison between completers ($N = 19$) and dropouts ($N = 14$)
Age $M(SD)$	21.61 (6.75)	$t(16) = -1.050, p = .155$
Gender (%)	70 % females	$\chi^2(1, 33) = 6.17, p = .013$
Explicit self-esteem $M(SD)$	25.69 (4.30)	$t(31) = -.505, p = .308$
Implicit self-esteem $M(SD)$	0.30 (2.11)	$t(31) = -.137, p = .446$
Depression $M(SD)$	08.03 (2.89)	$t(31) = .899, p = .187$
Anxiety $M(SD)$	08.12 (3.69)	$t(31) = .253, p = .401$
Exhaustion $M(SD)$	10.24 (5.56)	$t(31) = 1.038, p = .153$
Cynicism $M(SD)$	09.09 (5.02)	$t(31) = .645, p = .262$
Inefficiency $M(SD)$	10.87 (6.73)	$t(31) = .849, p = .201$
Academic self-esteem $M(SD)$	10.66 (2.34)	$t(31) = .269, p = .395$

Outcomes

Following the intervention, participants reported a significantly higher level of explicit self-esteem (H1a) $t(18) = -3.14, p = .003$, with a medium to large effect size ($d = .74$). Although there were no significant results for implicit self-esteem (H1b): $t(18) = -1.56, p = .067, d = .29$ (the result is marginally significant). Thus, Hypothesis 1 was partially supported by the data.

Hypothesis 2 was supported by the results. Following the intervention, participants reported significantly lower levels of depressive symptoms: $t(18) = 2.20, p = 0.02$, with a small to moderate effect size ($d = .46$). Similarly, Hypothesis

3 was also supported by the statistical analysis. Following the intervention, participants reported significantly lower levels of anxiety symptoms, the results being as follows: $t(18) = 2.15$, $p = .022$ with a small to moderate effect size ($d = .48$).

Hypothesis 4 was not supported by the data, the level of burnout felt by participants was not reduced by the course of the intervention. Non-significant results were obtained on all three dimensions of burnout (exhaustion: $t(18) = .186$, $p = .427$, $d = .04$; cynicism: $t(18) = -.774$, $p = .224$, $d = -.16$ and inefficiency: $t(18) = .346$, $p = .367$, $d = .07$).

Results did not support Hypothesis 5 either. After the intervention, students did not report higher levels of academic self-esteem ($t(18) = -0.741$, $p = .234$, $d = .16$).

For subjective evaluations of Zoom platform usability, a score of 74.86 was obtained ($N = 19$, $M = 74.86$, $SD = 22.66$) using the formula presented in the methodology. This means the participants evaluated the Zoom platform that was used as acceptable and satisfactory.

Table 2. Paired sample T-test results (one-tailed) from *PP* analysis

	Pre-test		Post-test		<i>t</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Explicit self-esteem	25.36	4.89	28.31	5.17	-3.14	.003	.74
Implicit self-esteem	0.26	2.29	1.33	1.35	-1.56	.067	.29
Depression	08.42	3.20	06.36	3.96	2.20	.020	.46
Anxiety	08.26	3.60	06.10	3.84	2.15	.022	.48
Exhaustion	11.10	6.22	10.89	7.01	0.18	.427	.04
Cynicism	09.57	5.28	10.73	7.94	-0.77	.224	-.16
Inefficiency	11.73	7.17	11.10	8.39	0.34	.367	.07
Academic self-esteem	11.00	2.72	11.57	2.52	-0.74	.234	.16

$N = 19$

Table 3. Paired sample T-test results (one-tailed) from *ITT* analysis

	Pre-test		Post-test		<i>t</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Explicit self-esteem	25.69	4.30	27.39	4.60	-2.86	.003	.51
Implicit self-esteem	0.30	2.11	0.92	1.66	-1.54	.066	.30
Depression	08.03	2.89	06.84	3.40	2.11	.021	.29
Anxiety	08.12	3.69	06.87	3.93	2.07	.023	.35
Exhaustion	10.24	5.56	10.12	6.05	0.18	.426	.03
Cynicism	09.09	5.02	09.75	6.78	-0.78	.220	-.12
Inefficiency	10.87	6.73	10.51	7.45	0.34	.364	.05
Academic self-esteem	10.66	2.34	11.00	2.27	-0.74	.231	.12

$N = 33$

For each therapeutic factor, the average score was calculated. An average close to the maximum value (5) means that students have perceived this factor as contributing to increased self-esteem. We list them in order of their perceived importance: therapeutic relationship ($M = 4.79$, $SD = 0.14$), self-understanding ($M = 4.68$, $SD = 0.17$), group cohesiveness ($M = 4.3$, $SD = 0.23$), inoculating hope ($M = 4.37$, $SD = 0.23$), catharsis ($M = 4.32$, $SD = 0.24$), universality ($M = 4.32$, $SD = 0.24$), psychoeducation ($M = 4.16$, $SD = 0.19$), homework ($M = 3.84$, $SD = 0.20$), and guidance ($M = 2.89$, $SD = 0.32$).

Table 4. Descriptive statistics for participants' ratings of the therapeutic factors

	<i>M</i>	<i>SD</i>
Therapeutic relationship	4.79	0.14
Self-understanding	4.68	0.17
Group cohesiveness	4.37	0.23
Inoculating hope	4.37	0.23
Catharsis	4.32	0.24
Universality	4.32	0.24
Psychoeducation	4.16	0.19
Homework	3.84	0.20
Guidance	2.89	0.32

As for general opinions about the intervention, 63% of students declared very satisfied, and 37% were satisfied with the program. 74% evaluated the quality of information as very good and 26% good. The pace was perceived as problematic: 70% declared the program short/too short, while only 30% perceived the tempo as adequate.

Discussion

This study aimed to investigate the feasibility of an online intervention designed to enhance students' self-esteem, based on the Fennell model. The results demonstrate the potential of this intervention, as students reported significantly higher levels of explicit self-esteem. This suggests that the model is suitable even for non-patients, confirming the promising findings previously reported in the literature (Kolubinski et al., 2018). Following our five-week program, students reported increased explicit self-esteem and reduced levels of depression and anxiety. However, there was no statistically significant result for implicit self-esteem (only marginally). This may be because the therapeutic techniques used in the intervention primarily impact explicit cognitions and, to a lesser extent, the implicit ones. Similarly, Sava et al. (2011) found no significant relationship between implicit self-esteem and irrational beliefs, suggesting that the classical CBT technique of identifying and disputing irrational beliefs may not be suitable for boosting implicit

self-esteem. In this case, cognitive training techniques or other types of approaches directly targeting implicit self-attitudes might be more efficient. Also, as the authors further suggested, implicit self-esteem might actually play a moderating role (in our case) for the efficacy of the intervention (i.e., initial low levels of implicit self-esteem being a factor of vulnerability; Sava et al., 2011). Hence, in future randomized controlled trials focused on the efficacy of the current intervention, we could also investigate this hypothesis.

These results can be further explained by the dual process theory (Chaiken & Trope, 1999), which has been extended to neural systems (Satpute & Lieberman, 2006). According to this theory, we process information in two distinct ways: consciously and deliberately (through the reflective system) and automatically (through the reflexive system). Thus, interventions targeting explicit aspects may not necessarily affect implicit, automatic attitudes. Additionally, a meta-analysis by Krizan and Suls (2008) revealed a small association between explicit and implicit self-esteem ($r = .11$), which aligns with this theory.

However, we want to also highlight the risk of type II error regarding implicit self-esteem results (since the effect was marginally significant). First of all, the initial sample size estimation was computed for a moderate effect size. However, the obtained effect on this outcome was small (PP $d = .29$; ITT $d = .30$). Combined with the high dropout rate (for the PP analysis), this suggests the sample might have been underpowered for detecting the effect of the intervention on implicit self-esteem. Second of all, 2 participants from the final sample of 19 (and an additional one from the 14 who dropped out) had identical NLT scores at pre and post-test. Because there were no objective indications of outliers, we decided to keep them in data analysis to avoid distorting the results. However, running an additional analysis for this outcome without the 2 mentioned participants results in improving the p -value from .067 to .060, which is closer to the cut-off of .050. Moreover, as Hoorens (2014) revealed, the validity and reliability of the competing algorithms for estimating the name-letter effect is still a topic of discussion. Thus, the type of score might also have an impact on the obtained result. Hence, the result of implicit self-esteem should be interpreted with caution.

Our intervention was designed as an online group program, particularly given the pandemic context. In line with other group interventions aimed at emotional distress (Koutra et al., 2010), we achieved significant results in reducing anxiety and depression symptoms.

The decrease in depressive symptoms may be explained by the vulnerability model (Beck, 1967) which suggests that negative self-evaluation is a causal risk factor for depression. Therefore, increasing self-esteem reduces depressive symptoms. However, it is challenging to pinpoint whether the reduction in symptoms is solely due to increased self-esteem (Sowislo & Orth, 2013), the group format, or other factors. When asked about the most helpful aspects of the program, our participants emphasized their relationship with the coordinator and gaining a deeper

understanding of themselves. These findings align with the classic perspective on therapeutic alliances, which are considered a factor for change in relational therapy or a necessary but not sufficient condition in classical cognitive approaches. Strong alliances have consistently been associated with positive outcomes across various psychotherapeutic approaches, as evidenced by recent meta-analyses (Flückiger et al., 2018; Baier et al., 2020). Self-understanding was another critical aspect for our participants, which is also a fundamental therapeutic factor often emphasized in group therapy (Behenck et al., 2016; Lovett & Lovett, 1991).

Another secondary outcome examined in this study was academic burnout. Although prior research has shown a negative association between overall self-esteem and burnout (Kupcewicz & Jóźwik, 2020), we did not find significant results in the present study. Additionally, the statistical analysis did not support our expected results regarding specific (academic) self-esteem. This discrepancy could be attributed to the timing of the post-test data collection, which occurred just before the exam session. According to Serebrennikova and Makarov (2019), emotional burnout indicators tend to be higher one month before the examination session, with the highest levels immediately before this stressful period. Furthermore, our sample consisted of first-year undergraduates, and the novelty of the evaluation itself could have represented an additional potential source of stress.

Regarding general feedback on our intervention, students expressed high satisfaction with the program, and the quality of information provided was generally well-received. On the 5-point Likert scale, 15 students indicated that they were “very satisfied with the program” (option 5), while 8 students selected the fourth level, indicating that they were “satisfied with the program”. Probably, the helpful aspects mentioned above (the relationship with the coordinator and gaining a deeper self-understanding), contributed to the overall satisfaction with the program. At the same time, some students noted that the program pace felt somewhat short, suggesting that while the content and support were appreciated, the duration of the intervention may need adjustment in future implementations. It is important to note that no qualitative data were collected beyond the Likert-scale ratings, which limits a more in-depth understanding of participants’ experiences.

Although the feasibility of the intervention was tested in the specific context of the pandemic, low self-esteem among students continues to represent a significant issue, often co-occurring with symptoms of anxiety and depression (Cao & Liu, 2024). Therefore, the results of this study remain transferable to the current post-pandemic context. Moreover, the strategies employed, such as group counselling and cognitive restructuring techniques, retain their relevance in supporting students’ well-being (Setiyowati et al., 2023). The skills fostered through the intervention, including the ability to manage anxious thoughts and to replace self-criticism with a more compassionate perspective, represent valuable resources for students’ mental health. Overall, we truly consider that the intervention holds potential for supporting students’ mental health in the post-pandemic context as well.

Limitations and future directions

The current study was an open-label, uncontrolled trial without follow-up measures. While this study suggests the promise of our proposed intervention, it must undergo randomized controlled trials to draw inferences about its efficacy. The absence of a control group impedes us from concluding that the changes in students' levels of explicit self-esteem, anxiety, and depression are caused by our intervention and not by other confounding variables, such as positive life events (e.g. new friends), faculty accommodation or passing hard partial exams. Hence, since through our study we found initial evidence for the feasibility of the intervention, future research should focus on testing its efficacy and effectiveness based on randomized controlled trial methodologies.

In addition to the study's inherent limitations, another concern is the dropout rate, which stood at 42%. It's worth noting that dropout rates are a common phenomenon in clinical interventions, particularly those delivered over the Internet. Melville et al. (2010) reported a mean dropout rate of 31%, influenced by various factors such as participants' initial expectations, psychological conditions, and treatment specifics, among others. Our data suggest a single difference between those who stayed and those who dropped: gender. This could result as a statistical artifact since there was an initial discrepancy in our sample distribution (70% females vs. 30% male), hence the higher probability of having more female than male dropouts. We also consider the pandemic context as a potential contributor to this rate. During the intervention period, all university courses were conducted online, which might have led some students to experience fatigue with excessive screen time. Indeed, this intervention should be tested in a context outside of social isolation. The online format primarily serves as a convenient means of connecting students from different faculties and universities. Moreover, since 70% of the students found the intervention to be too short, we recommend increasing the number of sessions to at least seven.

Also a limitation of this study is that the data was collected during the COVID-19 pandemic. We could not anticipate how the health status of participants and their loved ones might have affected our data, especially in terms of anxiety and depression rates. Furthermore, the economic impact of the pandemic, such as potential job loss or financial insecurity within their families, may have influenced participants' responses and behaviors. Therefore, we suggest assessing the practicality of this online intervention in the current, non-pandemic context.

Conclusions

The presented results suggest that our proposed intervention for improving self-esteem is feasible and well-received by university students. The program can be

further refined and implemented to enhance its impact on self-esteem and mental health, including the reduction of symptoms of depression and anxiety. With additional testing and development, this program has the potential to become an integral component of university counseling services.

Authors' Notes

Luana M. Alexa  <https://orcid.org/0009-0000-2731-0955>

Andrei Rusu  <https://orcid.org/0000-0003-0036-4149>

Funding: The work of Luana M. Alexa was supported by an internal grant from the Faculty of Sociology and Psychology (West University of Timișoara), no. 15636/15.03.2022. The work of Andrei Rusu was supported by funding from the NO Grants 2014-2021, under project contract no. 17/2020.

Acknowledgment. The authors declare there are no potential conflicts of interest associated with this publication.

Data availability. The data and all materials (e.g., presentations for participants, working sheets in Romanian) used in the program are available on request from the corresponding author.

References

- Adamson, J., Ozenc, C., Baillie, C., & Tchaturia, K. (2019). Self-Esteem Group: Useful Intervention for Inpatients with Anorexia Nervosa?. *Brain sciences*, 9(1), 12-22. <https://doi.org/10.3390/brainsci9010012>
- Albers, L., Rotteveel, M., & Dijksterhuis, A. P. (2009). Towards optimizing the name letter test as a measure of implicit self-esteem. *Self and Identity*, 8(1), 63-77. <https://doi.org/10.1080/15298860802091062>
- Arshad, M., Zaidi, S. M. I. H., & Mahmood, K. (2015). Self-Esteem & Academic Performance among University Students. *Journal of Education and Practice*, 6(1), 156-162.
- Arslan, G. (2019). Mediating role of the self-esteem and resilience in the association between social exclusion and life satisfaction among adolescents. *Personality and Individual Differences*, 151, 109514. <https://doi.org/10.1016/j.paid.2019.109514>
- Bachman, J. G. (1970). *The impact of family background and intelligence on tenth-grade boys (Vol. 2)*. Institute for Social Research, University of Michigan.
- Baier, A. L., Kline, A. C., & Feeny, N. C. (2020). Therapeutic alliance as a mediator of change: A systematic review and evaluation of research. *Clinical Psychology Review*, 82. <https://doi.org/10.1016/j.cpr.2020.101921>

- Beattie, S., & Beattie, D. (2018). An investigation into the efficacy of a cognitive behavioural therapy group for low self-esteem in a primary care setting. *The Cognitive Behaviour Therapist*, 11, e12. <https://doi.org/10.1017/S1754470X18000168>
- Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York, NY: Harper and Row.
- Beck, J. S. (2011). *Cognitive Behavior Therapy: Basics and Beyond (2nd ed.)*. New York: Guilford Press.
- Berg, M., Lindegaard, T., Flygare, A., Sjöbrink, J., Hagvall, L., Palmebäck, S., Klemetz, H., Ludvigsson, M. & Andersson, G. (2022). Internet-based CBT for adolescents with low self-esteem: a pilot randomized controlled trial. *Cognitive Behaviour Therapy*, 51(5), 388-407. <https://doi.org/10.1080/16506073.2022.2060856>
- Behenck, A. S., Gomes, J. B., & Heldt, E. (2016). Patient rating of therapeutic factors and response to cognitive-behavioral group therapy in patients with obsessive-compulsive disorder. *Issues in mental health nursing*, 37(6), 392-399. <https://doi.org/10.3109/01612840.2016.1158335>
- Billingham, S. A., Whitehead, A. L., & Julious, S. A. (2013). An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. *BMC medical research methodology*, 13, 1-6. <https://doi.org/10.1186/1471-2288-13-104>
- Bodnaru, A., Rusu, A., Blonk, R. W. B., Virgă, D., Iliescu, D., & Van den Broeck, A. (2023). iJobs - An online implementation of the JOBS II program for fostering reemployment: A feasibility and acceptability study. *Internet interventions*, 34, 100674. <https://doi.org/10.1016/j.invent.2023.100674>
- Brooke, J. (1996). Sus: a “quick and dirty” usability. In Jordan, P. W., Thomas, B., McClelland, I. L., & Weerdmeester, B. (Eds.). *Usability evaluation in industry*. CRC Press.
- Brown, J., Elliott, S., Boardman, J., Ferns, J., & Morrison, J. (2004). Meeting the unmet need for depression services with psycho-educational self-confidence workshops: Preliminary report. *The British Journal of Psychiatry*, 185(6), 511-515. <https://doi.org/10.1192/bjp.185.6.511>
- Cao, X., & Liu, X. (2024). Self-esteem as a predictor of anxiety and academic self-efficacy among Chinese university students: a cross-lagged analysis. *Current Psychology*, 43(22), 19628-19638. <https://doi.org/10.1007/s12144-024-05781-4>
- Chandra, E. K., Wibowo, M. E., & Sunawan, S. (2019). Cognitive Behaviour Group Counseling with Self-Instruction and Cognitive Restructuring Techniques to Improve Students' Self-Confidence. *Islamic Guidance and Counseling Journal*, 2(1), 11-17. <https://doi.org/10.25217/igcj.v2i1.305>
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. Guilford Press.

- Coffey, J. K., & Warren, M. T. (2020). Comparing adolescent positive affect and self-esteem as precursors to adult self-esteem and life satisfaction. *Motivation and emotion*, 44(5), 707–718. <https://doi.org/10.1007/s11031-020-09825-7>
- Cotiga, A. C., Zanzfirescu, Ș.A., Iliescu, D., Ciumăgeanu, M., Gotca, I., & Popa, C. O. (2023). Psychometric Characteristics of the Romanian Adaptation of the GAD-7. *Journal of Psychopathology and Behavioral Assessment*, 45, 691–701. <https://doi.org/10.1007/s10862-023-10062-8>
- Crisp, D., Griffiths, K., Mackinnon, A., Bennett, K., & Christensen, H. (2014). An online intervention for reducing depressive symptoms: secondary benefits for self-esteem, empowerment and quality of life. *Psychiatry Research*, 216(1), 60-66. <https://doi.org/10.1016/j.psychres.2014.01.041>
- David, D. (2020). *Clinical Assessment System*. SC RTS Romanian Psychological Testing Services. <https://www.rtscluj.ro/>
- Di Giunta, L., Alessandri, G., Gerbino, M., Kanacri, P. L., Zuffiano, A., & Caprara, G. V. (2013). The determinants of scholastic achievement: The contribution of personality traits, self-esteem, and academic self-efficacy. *Learning and Individual Differences*, 27, 102-108. <https://doi.org/10.1016/j.lindif.2013.07.006>
- Eurostat (2021). Education and Training Monitor 2021 - Romania key indicators. <https://op.europa.eu/webpub/eac/education-and-training-monitor-2021/en/romania.html>
- Fennell, M. J. (1997). Low self-esteem: A cognitive perspective. *Behavioural and Cognitive Psychotherapy*, 25(1), 1-26. <https://doi.org/10.1017/S1352465800015368>
- Fennell, M. (2016). *Overcoming low self-esteem: A self-help guide using cognitive behavioural techniques*. Hachette UK.
- Flückiger, C., Del Re, A. C., Wampold, B. E. & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, 55(4), 316-340. <https://doi.org/10.1037/pst0000172>
- Giofrè, D., Borella, E., & Mammarella, I. C. (2017). The relationship between intelligence, working memory, academic self-esteem, and academic achievement. *Journal of Cognitive psychology*, 29(6), 731-747. <https://doi.org/10.1080/20445911.2017.1310110>
- Hawi, N. S., & Samaha, M. (2017). The relations among social media addiction, self-esteem, and life satisfaction in university students. *Social Science Computer Review*, 35(5), 576-586. <https://doi.org/10.1177/0894439316660340>
- Harris, M. A., & Orth, U. (2020). The link between self-esteem and social relationships: A meta-analysis of longitudinal studies. *Journal of personality and social psychology*, 119(6), 1459-1477. <https://doi.org/10.1037/pspp0000265>
- Hiçdurmaz, D., İnci, F. & Karahan, S. (2017). Predictors of mental health symptoms, automatic thoughts, and self-esteem among university students. *Psychological reports*, 120(4), 650-669. <https://doi.org/10.1177/0033294117707945>

- Hoorens, V. (2014). What's really in a Name-Letter Effect? Name-letter preferences as indirect measures of self-esteem. *European Review of Social Psychology*, 25(1), 228-262. <https://doi.org/10.1080/10463283.2014.980085>
- Horrell, L., Goldsmith, K., Tylee, A., Schmidt, U., Murphy, C., Bonin, E., Beecham, J., Kelly, J., Raikundalia, S., & Brown, J. (2014). One-day cognitive-behavioural therapy self-confidence workshops for people with depression: Randomised controlled trial. *The British Journal of Psychiatry*, 204(3), 222-233. <https://doi.org/10.1192/bjp.bp.112.121855>
- Hu, J., Huang, Y., Liu, J., Zheng, Z., Xu, X., Zhou, Y., & Wang, J. (2022). COVID-19 related stress and mental health outcomes 1 year after the peak of the pandemic outbreak in china: the mediating effect of resilience and social support. *Frontiers in Psychiatry*, 13. <https://doi.org/10.3389/fpsyg.2022.828379>
- Isbăsoiu, A. B., Tulbure, B. T., Rusu, A., & Sava, F. A. (2021). Can We Boost Treatment Adherence to an Online Transdiagnostic Intervention by Adding Self-Enhancement Strategies? Results From a Randomized Controlled Non-inferiority Trial. *Frontiers in psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.752249>
- Jan, M., Soomro, S., & Ahmad, N. (2017). Impact of social media on self-esteem. *European Scientific Journal*, 13(23), 329-341.
- Javed, Z., & Tariq, O. (2016). Career-decisions, Self-efficacy and Self-esteem among Students of Private and Government Academic Institutions. *Pakistan Journal of Social and Clinical Psychology*, 14(2), 42-46.
- Jeffries, V., & Salzer, M. S. (2020). Mental health symptoms and academic achievement factors. *Journal of American College Health*, 1-4. <https://doi.org/10.1080/07448481.2020.1865377>
- Keltikangas-Järvinen, L. (1992). Self-esteem as a predictor of future school achievement. *European Journal of Psychology of Education*, 7(2), 123-130. <https://doi.org/10.1007/BF03172889>
- Kolubinski, D. C., Frings, D., Nikčević, A. V., Lawrence, J. A., & Spada, M. M. (2018). A systematic review and meta-analysis of CBT interventions based on the Fennell model of low self-esteem. *Psychiatry research*, 267, 296-305. <https://doi.org/10.1016/j.psychres.2018.06.025>
- Koutra, A., Katsiadrami, A., & Diakogiannis, G. (2010). The effect of group psychological counselling in Greek university students' anxiety, depression, and self-esteem. *European Journal of Psychotherapy and Counselling*, 12(2), 101-111. <https://doi.org/10.1080/13642537.2010.482733>
- Krizan, Z., & Suls, J. (2008). Are implicit and explicit measures of self-esteem related? A meta-analysis for the Name-Letter Test. *Personality and Individual Differences*, 44(2), 521-553. <https://doi.org/10.1016/j.paid.2007.09.017>
- 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>

- Kupcewicz, E., & Jóźwik, M. (2020). Association of burnout syndrome and global self-esteem among Polish nurses. *Archives of medical science*, 16(1), 135-145. <https://doi.org/10.5114/aoms.2019.88626>
- Leahy, R. L. (2003). *Cognitive therapy techniques: A practitioner's guide*. Guilford Press.
- Lewis, J. R. (2018). The system usability scale: past, present, and future. *International Journal of Human-Computer Interaction*, 34(7), 577-590. <https://doi.org/10.1080/10447318.2018.1455307>
- Liu, M., Wu, L., & Ming, Q. (2015). How does physical activity intervention improve self-esteem and self-concept in children and adolescents? Evidence from a meta-analysis. *PloS one*, 10(8), e0134804. <https://doi.org/10.1371/journal.pone.0134804>
- Lovett, L., & Lovett, J. (1991). Group therapeutic factors on an alcohol in-patient unit. *The British Journal of Psychiatry*, 159(3), 365-370. doi: 10.1192/bjp.159.3.365
- Lupascu, N., Timar, B., Albai, A., Roman, D., Potre, O., & Timar, R. (2019). Validation and cross-cultural adaptation of the depression Patient's Health Questionnaire – 9 in the Romanian population of patients with Type 2 Diabetes Mellitus. *Diabetes, Metabolic Syndrome and Obesity*, 12, 841-849. <https://doi.org/10.2147/DMSO.S203099>
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of personality*, 74(2), 403-456. <https://doi.org/10.1111/j.1467-6494.2005.00380.x>
- Melville, K. M., Casey, L. M., & Kavanagh, D. J. (2010). Dropout from Internet-based treatment for psychological disorders. *British Journal of Clinical Psychology*, 49(4), 455-471. <https://doi.org/10.1348/014466509X472138>
- McLellan, S., Muddimer, A., & Peres, S. C. (2012). The effect of experience on System Usability Scale ratings. *Journal of usability studies*, 7(2), 56-67.
- Morton, L., Roach, L., Reid, H., & Stewart, S. (2012). An Evaluation of a CBT Group for Women with Low Self-Esteem. *Behavioural and Cognitive Psychotherapy*, 40(2), 221-225. <https://doi.org/10.1017/S1352465811000294>
- Nuttin Jr, J. M. (1987). Affective consequences of mere ownership: The name letter effect in twelve European languages. *European journal of social psychology*, 17(4), 381-402. <https://doi.org/10.1002/ejsp.2420170402>
- Orth, U., & Robins, R. W. (2022). Is high self-esteem beneficial? Revisiting a classic question. *American Psychologist*, 77(1), 5–17. <https://doi.org/10.1037/amp0000922>
- Pack, S., & Condren, E. (2014). An evaluation of group cognitive behaviour therapy for low self-esteem in primary care. *The Cognitive Behaviour Therapist*, 7, E7. <https://doi.org/10.1017/S1754470X14000051>
- Paloş, R., Maricuţoiu, L. P., & Costea, I. (2019). Relations between academic performance, student engagement and student burnout: A cross-lagged analysis

- of a two-wave study. *Studies in Educational Evaluation*, 60, 199-204. <https://doi.org/10.1016/j.stueduc.2019.01.005>
- Peyton, K., Huber, G. A., & Coppock, A. (2022). The Generalizability of Online Experiments Conducted During the COVID-19 Pandemic. *Journal of Experimental Political Science*, 9(3), 379–394. <https://doi.org/10.1017/XPS.2021.17>
- Polyakova, V., Streltsova, E., Iudin, I., & Kuzina, L. (2024). Irreversible effects? How the digitalization of daily practices has changed after the COVID-19 pandemic. *Technology in Society*, 76, 102447. <https://doi.org/10.1016/j.techsoc.2023.102447>
- Pullmann, H., & Allik, J. (2008). Relations of academic and general self-esteem to school achievement. *Personality and Individual Differences*, 45(6), 559-564. <https://doi.org/10.1016/j.paid.2008.06.017>
- Randal, C., Pratt, D., & Bucci, S. (2015). Mindfulness and self-esteem: a systematic review. *Mindfulness*, 6(6), 1366-1378. <https://doi.org/10.1007/s12671-015-0407-6>
- Rigby, L., & Waite, S. (2006). Group Therapy for Self-Esteem, Using Creative Approaches and Metaphor as Clinical Tools. *Behavioural and Cognitive Psychotherapy*, 35(3), 361-364. <https://doi.org/10.1017/S1352465806003389>
- Robson, P. (1989). Development of a new self-report questionnaire to measure self esteem. *Psychological Medicine*, 19, 513-518. <https://doi.org/10.1017/s003329170001254x>
- Rosenberg, M. (1965). Rosenberg self-esteem scale. *Journal of Religion and Health*. <https://doi.org/10.1037/t01038-000>
- Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American sociological review*, 60(1), 141-156. <https://doi.org/10.2307/2096350>
- Rossi, A., Panzeri, A., Pietrabissa, G., Manzoni, G. M., Castelnovo, G., & Mannarini, S. (2020). The anxiety-buffer hypothesis in the time of COVID-19: when self-esteem protects from the impact of loneliness and fear on anxiety and depression. *Frontiers in Psychology*, 11, 2177-2227. <https://doi.org/10.3389/fpsyg.2020.02177>
- Satpute, A. B., & Lieberman, M. D. (2006). Integrating automatic and controlled processes into neurocognitive models of social cognition. *Brain research*, 1079(1), 86-97. <https://doi.org/10.1016/j.brainres.2006.01.005>
- Sava, F. A., Maricutoiu, L. P., Rusu, S., Macsinga, I., & Virga, D. (2011). Implicit and explicit self-esteem and irrational beliefs. *Journal of Cognitive and Behavioral Psychotherapies*, 11(1), 97-111. <https://doi.org/10.1016/j.paid.2016.05.022>
- Schaufeli, W. B., Martinez, I. M., Pinto, A. M., Salanova, M., & Bakker, A. B. (2002). Burnout and engagement in university students: A cross-national study. *Journal of cross-cultural psychology*, 33(5), 464-481. <https://doi.org/10.1177/0022022102033005>

- Serebrennikova, O. A., & Makarov, S. Y. (2019). Features of emotional burnout of students in the conditions of pre-examination and examination academic stress. *Biomedical and Biosocial Anthropology*, (34), 26-32. <https://doi.org/10.31393/bba34-2019-04>
- Setiyowati, H. D., Dantes, N., & Dwiawati, K. A. (2023). The Effectiveness of Group Counseling Cognitive Restructuring Techniques to Increase the Self-confidence of Students. *Bisma The Journal of Counseling*, 7(1), 25-33. <https://doi.org/10.23887/bisma.v7i1.58813>
- Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological bulletin*, 139(1), 213-240. <https://doi.org/10.1037/a0028931>
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, 166(10), 1092-1097. doi:10.1001/archinte.166.10.1092
- Suh, H., Sohn, H., Kim, T., & Lee, D. G. (2019). A review and meta-analysis of perfectionism interventions: Comparing face-to-face with online modalities. *Journal of Counseling Psychology*, 66(4), 473-486. <https://doi.org/10.1037/cou0000355>
- Swartzman, S., Kerr, J., & McElhinney, R. (2021). An evaluation of a group-based cognitive behavioural therapy intervention for low self-esteem. *Behavioural and Cognitive Psychotherapy*, 49(4), 485-494. <https://doi.org/10.1017/S1352465821000072>
- Taylor, T. L., & Montgomery, P. (2007). Can cognitive-behavioral therapy increase self-esteem among depressed adolescents? A systematic review. *Children and Youth Services Review*, 29(7), 823-839. <https://doi.org/10.1016/j.childyouth.2007.01.010>
- Triana, R., Keliat, B. A., & Sulistiowati, N. M. D. (2019). The relationship between self-esteem, family relationships, and social support as the protective factors and adolescent mental health. *Humanities & social sciences reviews*, 7(1), 41-47. <https://doi.org/10.18510/hssr.2019.715>
- Waite, P., McManus, F., & Shafran, R. (2012). Cognitive behaviour therapy for low self-esteem: A preliminary randomized controlled trial in a primary care setting. *Journal of behavior therapy and experimental psychiatry*, 43(4), 1049-1057. <https://doi.org/10.1016/j.jbtep.2012.04.006>
- Yalom, I. D., & Leszcz, M. (Collaborator). (2005). *The theory and practice of group psychotherapy (5th ed.)*. Basic Books/Hachette Book Group.

Appendix – Overview of the intervention

The intervention was designed for students based on Melanie Fennell's model and guide¹. The main themes were discussed during five weekly meeting sessions with the students who passed the screening phase. There were four working groups, and each student was assigned to one depending on their academic schedule and preferences. All materials (PowerPoint presentations, working sheets) were translated into Romanian language. In addition, they were sent via e-mail weekly to those who were not present at one of the meetings.

Session 1

All first sessions began with the coordinator introducing herself and the program. Next, all participants were invited to introduce themselves and tell more about their hopes and fears about the intervention. In addition, the group's rules were discussed, such as confidentiality, turning on the camera, having a notebook or something to write on, and whatever request was made by one participant or another.

After the introduction phase, the coordinator clarified the concept of "self-esteem" and pointed out some differences between this concept and similar others, such as "self-image," "self-concept," "self-confidence," and "self-efficacy."

Next, the coordinator presented the red flags that show low self-esteem, such as self-criticism, self-blame, and self-doubts, and what it means to have healthy self-esteem (a fundamental acceptance of who you are, with strengths and weaknesses).

Students were invited to participate in a game to balance the theoretical part with a more interactive and emotional one². Each participant was asked to take a 10 or 50 lei bill (about 2 or 10 euros), fold it several times, and then unfold it. The coordinator pointed out that, as the bill, no matter how many imperfections someone has, the personal value remains the same. This metaphorical exercise was meant to show that unique value doesn't change, no matter how many spots a person has. For debriefing, every student shared their thoughts and feelings about the exercise and the now-and-here moment.

Next, the coordinator linked the ending of the first session with the second one by giving homework in which students should write down their negative beliefs about themselves.

¹ Fennell, M. (2016). *Overcoming low self-esteem: A self-help guide using cognitive behavioral techniques*. Hachette UK.

² Belmont, J. A. (2016). *103 activități de grup. Idei de tratament și strategii practice*. TREI: 159-160

Session 2

Session 2 began with a discussion about the negative beliefs previously identified by students, followed by the coordinator's explications of the CBT model.

From a list, students were asked to choose and note those early or late experiences that could have shaped how they see themselves and their beliefs about the self (such as high expectations, bullying, toxic relationships, etc.). Every student had the chance to share their experiences, thus facilitating normalization and cohesion of the group.

After a recap of the CBT model, the coordinator presented biases in thinking and how human beings tend to perceive only the information confirming their beliefs.

Ultimately, the coordinator presented what "rules of life" mean from Fannell's model and how they work. The main aim of the second meeting was to make students understand how low self-esteem develops and what rules they must follow to feel good about themselves (to be perfect, to be pleasant for everyone, to have absolute control over themselves, etc.) based on Fannell's model. Like the previous session, students received homework to note their own living rules, which must be respected to feel good about themselves.

The following scheme might resume all the theoretical parts explained in the second session:

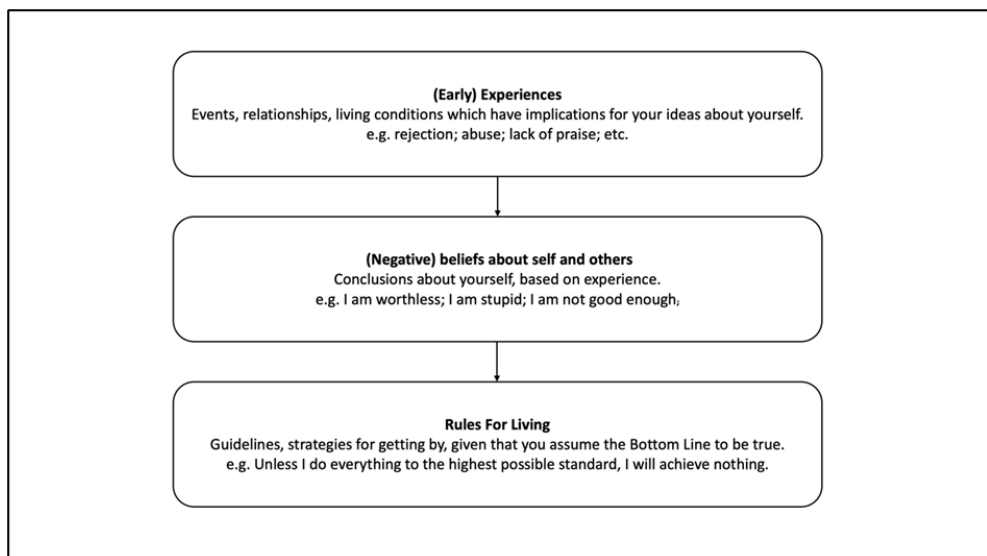


Figure 1. How a low self-esteem develops based on Melanie Fennell's model

The original name of the second box, “The Bottom Line” was replaced with the meaning of it („beliefs about self and others”) due to a lack of proper translation into Romanian.

Session 3

The third meeting focused on understanding how low self-esteem is being maintained, resumed by the following scheme, which was explained to students step by step.

First, the coordinator presented trigger situations and invited participants, based on their identified rules for living, to think about those moments in which rules might be broken. Next, the students were taught about the CBT model of anxiety. They were asked to reflect, note, and discuss their unhelpful behavior, such as avoidance, success minimization, unnecessary precautions, etc.

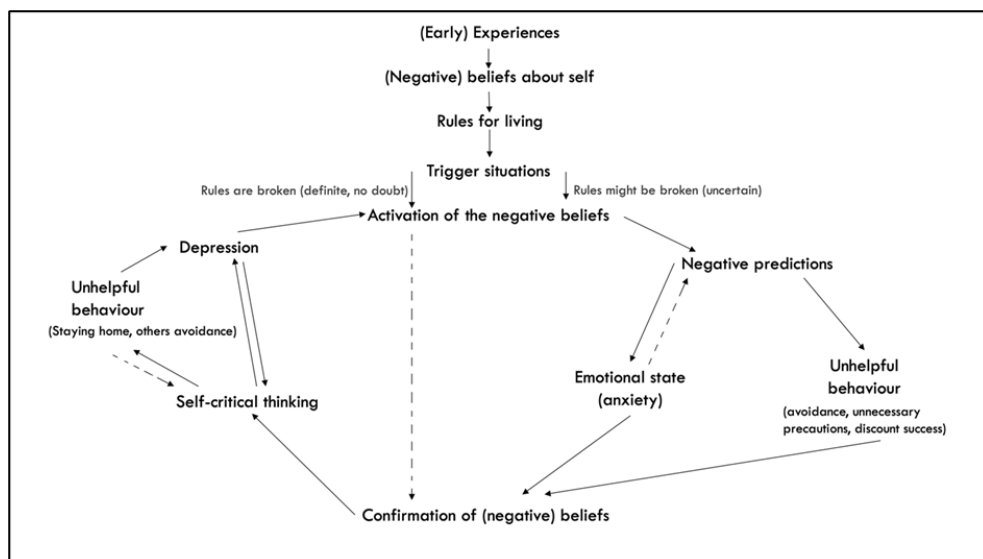


Figure 2. Melanie Fennell' cognitive model of low self-esteem

After that, the coordinator explained the other situation, in which the rules are certainly broken, a case in which negative beliefs are confirmed, and self-critical thoughts are activated. Consequently, negative feelings like sadness and hopelessness appear.

For homework, students were asked to map their vicious circle using a worksheet similar to Figure 2, plus blank boxes to write in.

Session 4

Session four started with checking homework and explaining to each student their vicious circle, followed by additional clarifications in case of unclear information.

Next, the coordinator pointed out that all participants have to break their vicious circle in two points (negative predictions and self-critical thinking), one for each situation (if rules for living might be broken or the rules are certainly broken). To achieve that, students were told about three steps to follow: *to be aware*, *to rethink*, and *to experience*.

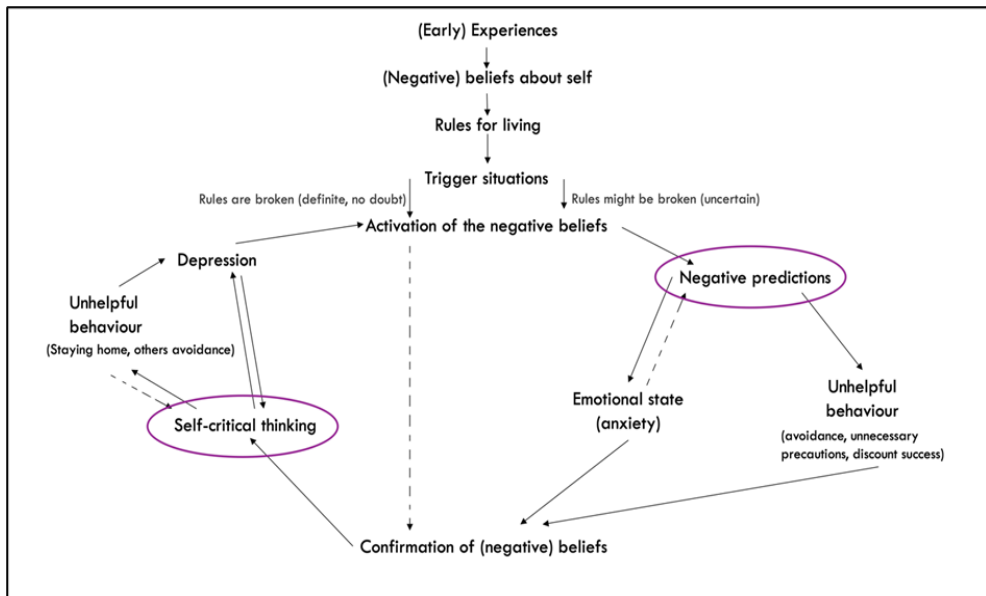


Figure 3. Points of action for breaking the vicious circle

To facilitate cognitive restructuring, students received worksheets based on those from Fannell's guide, one for anxious situations and one for self-critical situations. On the first page of each worksheet, there is a table that guided them to record the context, emotions and bodily sensations, related thoughts, their intensity, and the behaviors that followed (awareness stage). Afterwards, for each situation (negative predictions or self-criticism), participants were introduced to a set of key questions meant to help them find alternatives to their initial thoughts. These questions were presented on the second page of the worksheets.

On page three, the focus shifted to the rethink stage. For anxious predictions, students recorded the situation and emotions, rated the credibility of their initial thoughts, and then generated alternative perspectives using the key questions. They

also planned small experiments to test these new perspectives and reflected on what they did differently, what the outcomes were, and what they had learned.

For self-criticism, the page guided participants to reconsider the strength of their self-critical thoughts, explore alternative perspectives, and assess how strongly they believed in them. They then reflected on the results: how they felt after seeing alternatives, how strongly they still believed in the critical thoughts, and what concrete actions or experiments they could try.

These worksheets were also offered as homework, giving participants the opportunity to practice the rethink process in real-life situations throughout the following week.

Session 5

The last meeting started with checking each student's worksheets and sharing the previous week's opinions, thoughts, and experiences. As in session 4, there were pointed ways and questions for themselves to make rules for living more flexibly and restructuring irrational thoughts about using CBT-specific techniques.

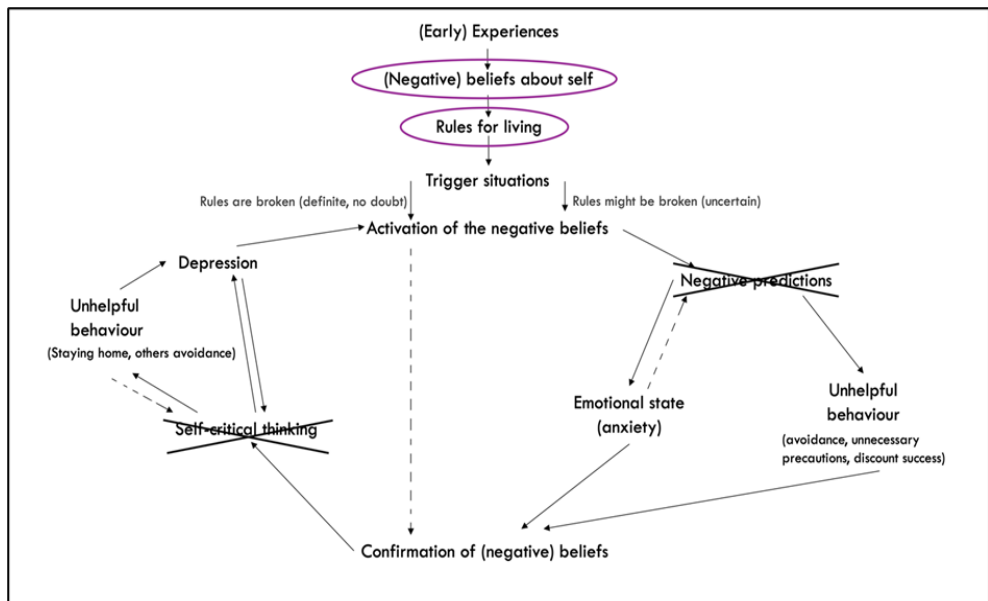


Figure 4. The main themes discussed in the last session of the intervention

After the theoretical part, students were invited to reflect on their own experiences in the group during the last five weeks and share with others the most relevant things they learned in the time they spent together. This reflective exercise not only reinforced learning but also provided an opportunity to consolidate

progress, recognize individual achievements, and highlight strategies they could continue to use beyond the sessions. The activity took up more than half of the session, giving each participant sufficient time and space to share their personal reflections, express their thoughts, and receive feedback from peers. This allowed for meaningful discussion, ensured that all voices were heard, and fostered a sense of closure and mutual support within the group.

INSTRUCTIONS TO AUTHORS

Manuscript preparation.

Manuscripts should not exceed 50 pages, must be prepared according to the *Publication Manual of the American Psychological Association*, 7th edition, and must be typed double-spaced throughout including the cover sheet, abstract and keywords, text, references, and tables/figures. Articles should be written in English. All articles should include an English abstract of no more than 250 words. Manuscripts are submitted for blind review; therefore, the author's name, degree, and affiliation should appear on the cover sheet only, which should also include the article title and the complete mailing address and telephone number of the corresponding author.

Submitting the manuscript.

Manuscripts can be currently submitted using the Journal's electronic submission system, using: <https://www.ejmanager.com/my/jebp/>. The corresponding author will receive a confirmation e-mail once a manuscript number has been assigned. For manuscript-related inquiries, please also contact the Editor (jebp@psychology.ro).

The corresponding author should provide a cover letter confirming that the manuscript is not being concurrently reviewed by another journal together with an anti-plagiarism declaration stating that the paper they are submitting does not contain any plagiarized content. The opinions and statements do not necessarily represent the policies of IIAS of PAMH or the views of the editor. By submitting a manuscript, the authors agree that the copyright for their article is transferred to the IIAS of PAMH, if and when the article is accepted for publication.

Subscriptions.

Starting with volume XXIII (2023), JEBP has become a Platinum Open Access journal which allows immediate access to the content of the journal without the payment of a subscription fee or licence. Authors pay no article publication charge and all the costs of publishing the journal are met by one or more sponsoring organizations.

We acknowledge the support of Babeş-Bolyai University Program for Research, Competitiveness and Excellence, and CNFIS-FDI-2022-0179 Project.

ARTICLES SECTION

An exploratory comparison of IEMT- versus EMDR-directed eye movements on changes in emotionality and distress during recall of negative memories 1

DALENA VAN HEUGTEN – VAN DER KLOET, ANOUK BOONSTRA,
NAIMA TROUK, AND ANNE TEN BRINKE

Feasibility and Usability of a VR-Based Training for Therapeutic Relationship Skills in Future Psychotherapists: A Pilot Study 19

DANA-ANDREEA CRISTOLȚAN

Maternal Depression and Children's Socio-Emotional Development in the First Two Years of Life: A Meta-Analysis 39

MĂDĂLINA RUXANDRA COSTIN, DIANA TĂUT,
ADRIANA SMARANDA BĂBAN, AND SEBASTIAN PINTEA

Perceived Parental Blame and Emotion Regulation as Predictors of Adult Psychological Distress: Development and Validation of the Perceived Parental Blame Scale 69

NESRULLAH OKAN

Effect of aerobic exercise on relapse tendency in individuals with methamphetamine use disorder in a virtual social environment: Mediated by cue-induced craving and interoceptive sensibility 93

XIAOXIA LIU AND ZEFAN FAN

Do Values Change After Major Life Events? A Meta-analytic Review 117

DORIS POP, OANA A. DAVID

Effectiveness of Group-Based Cognitive Behavioral Therapy in Reducing Depression and Anxiety Among Adolescents in Residential Institutions: A Randomized Controlled Trial 137

DING WAI ENG AND FIRDAUS MUKHTAR

Improving students' self-esteem with an online group intervention based on the Fennell Model: an open-label, uncontrolled feasibility trial 163

LUANA M. ALEXA, ANDREI RUSU

