The Role of Digital Therapeutic Alliance in Mental Health Outcomes: Secondary Analysis of a Randomized Controlled Trial

Aino Alina Simone Delcos

Student number: 568941

Supervisor: dr. Marta Anna Marciniak

Second supervisor: dr. Linda Dekker

Master Specialization: Clinical Psychology

Social and Behavioral Science

Erasmus University, Rotterdam

Date: 04.12.2024

Word count: 8547

Abstract

Digital therapeutic alliance (DTA), as suggested by the latest empirical evidence, influences engagement and outcomes in digital mental health apps. However, little is known about whether DTA plays a role in moderating mental health outcomes in fully automated apps. The present study examined whether DTA moderates changes in perceived stress, anxiety, and depressive symptoms during the use of ReApp, a fully automated mental health app aiming to enhance positive cognitive reappraisal (PCR). Ninety-two young adults used ReApp for three weeks and completed questionnaires before and after the intervention. DTA was measured at follow-up with the mobile Agnew Relationship Measure (mARM-G). Linear mixed-effects models were used to test whether general DTA and its five subscales (bond, partnership, confidence, openness, and client initiative) moderated mental health outcomes. The results showed a significant main effect of the openness subscale on perceived stress ($\beta = -2.12$, $p = .04$), although openness did not function as a moderator. Similarly, none remaining mARM-G subscales showed moderation effects, and general DTA did not moderate changes in mental health outcomes. These findings suggest that the mechanism through which users develop a sense of alliance with mental health apps may differ from alliance processes in traditional therapy. Given these conceptual differences, measures adapted from therapist-based frameworks, including the mARM-G, may not adequately capture DTA or its digital-specific processes. Future research should use newly developed alliance measures designed specifically for digital tools to more accurately assess the role of DTA in automated mental health apps.

Keywords: digital therapeutic alliance, ecological momentary intervention, mobile agnew relationship measure, anxiety, stress, depression, moderation

The Role of Digital Therapeutic Alliance in Digital Mental Health Outcomes: Secondary Analysis of a Randomized Controlled Trial

Mental health disorders affect over one billion people worldwide, causing major human and economic burdens (World Health Organization: WHO, 2025). Globally, this challenge is intensified by limited government investment in mental health, which accounts for a median of only 2% of total health budgets (WHO, 2025). For instance, in the Netherlands, long waiting times in mental health care are a widely discussed issue (Netherlands Court of Audit [Rekenkamer], 2020). These delays are primarily caused by a shortage of qualified healthcare professionals, growing patient demand for long-term treatments, and financial constraints that limit institutional capacity (Netherlands Court of Audit [Rekenkamer], 2020; Oecd & Policies, 2023). The Netherlands has been reported as one of the leading countries in the digitalization of mental health services (Oecd & Policies, 2023). In 2022, Dutch healthcare stakeholders signed a national agreement to promote digital and hybrid care, aiming to improve access to mental health services and reducing waiting times. (Keij et al., 2024) Therefore, research on digital mental health tools is essential to support their effective implementation.

Alongside these national goals in the Netherlands, global research on the digitalization of psychological care has also grown over the past two decades (Linardon et al., 2024; Macrynikola et al., 2025; Taylor et al., 2025). Digital tools are increasingly viewed as a promising way to enhance the accessibility and affordability of mental health support, particularly through Ecological Momentary Interventions (EMIs) (Amo & Lieder, 2024; Macrynikola et al., 2025; Marciniak et al., 2020; Marciniak et al., 2023; Tong et al., 2025; Wang et al., 2025). EMIs deliver real-time support within individuals' daily environments, allowing users to practice therapeutic strategies during everyday moments of need, thereby strengthening their psychological well-being (Heron & Smyth, 2009). Reflecting this growing interest, recent meta-

analyses demonstrate that digital mental health interventions produce significant small to moderate improvements in common conditions such as anxiety and depression (Linardon et al., 2024; Malouin-Lachance et al., 2025; Taylor et al., 2025). More specifically, studies that measured changes within the same participants from before to after using EMIs reported medium to large effect sizes, while studies comparing participants using EMIs to a control group (waitlist (n=82), placebo (n=317), or active treatment control (n=123)) found small to medium effect sizes (Versluis et al., 2016).

On the other hand, a recent meta-analysis of 176 randomized controlled trials found that about one in four participants (23.6%) stopped using their assigned mental health app before completing the program **(**Linardon et al., 2024**)**. Specifically, apps targeting depressive symptoms showed a mean dropout rate of 26.2%, increasing to 47.8% after correcting for publication bias (Torous et al., 2019). The dropout rates in digital interventions are often associated to lower engagement and lack of accountability measures (Ashur et al., 2024; Malouin-Lachance et al., 2025). This low engagement in fully automated digital interventions may stem from a lack of therapeutic alliance (TA) (Baumel & Kane, 2018).

Traditionally, the TA is defined as a measure of the quality of the relationship between a therapist and their client (Horvath & Luborsky, 1993). Bordin's pan-theoretical model (1979) is the most widely used theory of alliance, consisting of agreement on goals, agreement on tasks, and a strong emotional bond built on trust and mutual understanding (Bordin, 1979). It is recognized as the foundation of effective treatment and central predictor of successful traditional psychotherapy (Lederman & D'Alfonso, 2021; Malouin-Lachance et al., 2025).

Following the rise of digital treatments, the concept of the Digital Therapeutic Alliance (DTA) has been introduced in research over the past decade (Berry et al., 2018; Clarke et al.,

2016; Snipes et al., 2025; Taylor et al., 2025; Tong et al., 2025; Von Wulffen, Marciniak et al., 2023). Proposing that a client can form a working relationship with a digital mental health intervention (Berry et al., 2018; D'Alfonso et al., 2020), DTA aims to conceptually capture and measure the therapeutic quality of digital mental health interventions. Studies have shown that higher DTA is associated with greater improvements in anxiety and depressive symptoms (Macrynikola et al., 2025**;** Snipes et al., 2025).

While studies have established the concept of DTA, its impact on clinical outcomes varies widely across contexts, resulting in mixed findings. (Amo & Lieder, 2024; Ashur et al., 2024; Clarke et al., 2016) These inconsistencies mainly stem from methodological and conceptual challenges in applying traditional psychotherapy frameworks to digital interventions. To measure DTA, Berry et al. (2018) adapted the validated Agnew Relationship Measure (ARM) for digital contexts. The resulting questionnaire, the mobile Agnew Relationship Measure (mARM) assesses DTA on five subscales: bond, partnership, confidence, openness, and client initiative. While overall findings indicate that the mARM provides a valid measure of DTA with good face and content validity (Snipes et al., 2025; Von Wulffen, Marciniak et al., 2023), the individual subscales, particularly client initiative (Von Wulffen, Marciniak et al., 2023) and openness (Teague et al., 2025), may exhibit somewhat weaker psychometric properties.

To our knowledge, few studies have investigated the moderating role of the DTA in EMIs, indicating the need for further research. The Acceptance and Commitment Therapy in Daily Life (ACT-DL) trial by van Aubel et al. (2022) is the only published EMI study to date that has examined whether TA moderates treatment effects. The study found that higher alliance enhanced psychological flexibility among young people with early psychosis. However, the trial

assessed alliance using a traditional TA measure rather than a digital-specific instrument (van Aubel et al., 2022).

Building on emerging evidence, the present study examines whether DTA moderates changes in users' mental health outcomes during the use of an EMI. Despite growing interest in DTA, most existing studies have focused on apps that combine several therapeutic methods, such as cognitive restructuring techniques, mindfulness, or artificial intelligence-based chatbots (Amo & Lieder, 2024; Macrynikola et al., 2025; Taylor et al., 2025; Teague et al., 2025). Given that these apps integrate multiple therapeutic methods and, in some studies, include therapist or chatbot involvement (Macrynikola et al., 2025; Teague et al., 2025), it becomes challenging to isolate the unique contribution of the alliance to changes in mental health outcomes. Thus, it is still unknown whether DTA affects outcomes in fully automated interventions that use only one therapeutic method and involve no relational or human-like features.

To address this gap in literature, this study draws on data from Marciniak et al. (2023, 2024), who developed ReApp, a digital EMI app that exclusively targets positive cognitive reappraisal (PCR). PCR is an evidence-based emotion regulation strategy in which individuals reinterpret stressful situations to identify positive elements or potential benefits (Folkman & Lazarus, 1984). Low tendency to use PCR is considered a potential risk factor for mental health difficulties (Marciniak, Homan, et al., 2024). Since ReApp focuses only on PCR, it allows us to examine whether users' sense of alliance influences changes in mental health outcomes without the confounding effects of additional therapeutic features.

Previous findings from Marciniak et al. (2023, 2024), show that ReApp increases the use of PCR and leads to general improvements in stress, anxiety, and depressive symptoms. However, the role of DTA in moderating the mental health outcomes of ReApp, and EMIs more

generally, remains unknown. To address this gap, the current study extends on the findings of Marciniak et al. (2023, 2024) and is the first to examine whether general DTA and its specific subscales (bond, partnership, confidence, openness, and client initiative) moderate changes in depressive symptoms, anxiety, and perceived stress among young adults. Subsequently, the research questions are as follows:

1. Does general DTA, as measured by the mARM-G total score, moderate the effects of the ReApp on reductions in depressive symptoms (BDI-II), anxiety symptoms (STAI), and perceived stress (PSS)?
2. Are there differences between the five key mARM-G subscales of DTA (bond, partnership, confidence, openness, and client initiative) in how they moderate reductions in depressive symptoms (BDI-II), anxiety symptoms (STAI), and perceived stress (PSS) from baseline to follow-up?

Based on previously mentioned literature (Macrynikola et al., 2025; Snipes et al., 2025; van Aubel et al., 2022) the hypotheses are: (1) Higher general DTA, as measured by the mARM-G total score, will be associated with greater reductions in anxiety symptoms (STAI) from baseline to follow-up; (2) Higher general DTA as measured by the mARM-G total score, will be associated with greater reductions in depressive symptoms (BDI-II) from baseline to follow-up; (3) Higher general DTA, as measured by the mARM-G total score, will be associated with greater reductions in perceived stress (PSS) from baseline to follow-up, and (4) The five key mARM-G subscales of DTA (bond, partnership, confidence, openness, and client initiative) will moderate the reductions in mental health outcomes (BDI-II, STAI, and PSS). This is, however, a non-directional hypothesis due to lack of evidence in the literature, investigating which specific facets of digital TA are most impactful.

Methods

Study Design

The present study is a randomized-control trial (RCT) study. The data for this study were drawn from Marciniak et al. (2023, 2024), three-armed randomized controlled trial (ClinicalTrials.gov Identifier: NCT05784831). For the purpose of this study, the original intervention groups with differing prompts schedules (Burst vs. Continuous) were merged into a single intervention group, as Marciniak et al. (2024) found no significant differences in mental health outcomes or adherence between them. The control arm did not receive an intervention and is not included in the present study (see Marciniak et al., 2023, for further details). The study proposal was approved by the Ethics Committee of the Faculty of Arts and Social Sciences at the University of Zurich (approval #21.2.12)

Participants

Participants were recruited through online recruitment methods, including announcements on university websites, mailing lists for students, and social media. The current experiment has a sample size of 92 participants, of which four were male and 89 were female. Participants were German-speaking students from higher education institutions, and they were invited from Switzerland, Austria, Germany, and the Netherlands. The inclusion criteria of participants were: (1) Being a student of a higher education institution; (2) Being 18–29 years old; (3) Having sufficient knowledge of the German language; (4) Being a smartphone user; (5) Obtaining 12 points or less on a 20-point Positive Reappraisal subscale of the Cognitive Emotion Regulation Questionnaire, (CERQ; cutoff based on Marciniak et al. 2023). Participants were excluded if they reported a mental illness or were undergoing psychotherapy. All participants included in the study provided written informed consent.

Marciniak et al. (2024) estimated that, for an expected effect size of $d = 0.6$, with $\alpha = .05$ and a power of 80%, the required sample size was 94 participants. As the current sample size was closely aligned with this recommendation, it was unlikely to meaningfully affect the analyses.

Procedure

At baseline, individuals interested in participating completed a brief online screening to assess eligibility. This baseline questionnaire included basic demographic information (i.e., age, gender, education level) and included assessment of their tendency toward PCR using the CERQ. Those who met the inclusion criteria were contacted and received a personal code to be used throughout the study. Random assignment to experimental conditions was carried out using a block randomization procedure (block size of three), ensuring that participants were randomized in groups of three (Control, Burst, or Continuous).

Before beginning the intervention, all participants attended a remote meeting via ZOOM (required due to COVID-19 restrictions). During this meeting, participants were given instructions on how to download and use the digital EMI app, ReApp. They also received training specifically targeting PCR, including an explanation of its definition, theoretical background, and practical strategies for applying PCR techniques in everyday situations.

Participants then used ReApp for three weeks (21 days). The app consisted of two components, the EMI and an Ecological Momentary Assessment (EMA) rating participants' daily mood. Five EMA prompts were delivered per day within random three-hour intervals between 8AM and 10:30PM. Two of these daily EMA prompts were paired with EMIs between 11:30AM and 6:30PM. Participants assigned to the Burst condition did not receive any prompts

during the second week; instead, they were encouraged to use the intervention when needed. All participants could independently activate the EMI at any time felt it might be helpful.

In the EMA part, participants were asked each day about their current mood, stress level, and feelings. This measurement of daily mood was adapted from Vaessen et al. (2019). They received short questions like "I feel cheerful" or "I feel stressed" and answered on a scale from 1 (*not at all*) to 7 (*very much*). In total, 13 questions were asked each time. After the mood rating, participants were also asked about their social setting, online contact with others, and what they were doing at that moment. In the EMI part, participants were asked to think of a stressful event that had happened since the last app notification. They were then instructed to shift their perspective on the event by generating three possible positive reappraisals. If participants were feeling stressed about a future situation, they had the option to reappraise that anticipated event instead. The final step of the PCR training involved evaluating the reappraisal exercise. Using a 7-point Likert scale (1 = *not at all*, 7 = *very much*), participants rated how difficult the reappraisal was and whether it helped them cope more easily with the event.

After 21 days of app use, participants completed follow-up questionnaires which also included the assessment of DTA, received charts of their mood fluctuations, and were compensated with up to 105 Swiss francs or six university credit points.

Measures

Mental health outcomes

To investigate the effects of ReApp on mental health outcomes, we examined whether DTA moderated changes in depressive symptoms, anxiety, and perceived stress. The German version of the Beck Depression Inventory II (BDI-II; Kühner et al., 2007) was used to measure

depressive symptoms. This 21-item self-questionnaire shows high internal consistency of Cronbach alpha ≥ .84, good test–retest reliability, and strong convergent validity with other depression measures. Anxiety symptoms were measured with the German version of the State-Trait Anxiety Inventory (STAI; Laux et al., 1981). This self-questionnaire consists of 20 items assessing state anxiety (momentary anxiety) and 20 items assessing trait anxiety (dispositional anxiety). High internal and construct validity for both state and trait anxiety were reported. The German Perceived Stress Scale (PSS; Klein et al., 2016) was used to measure perceived stress. This 10-item self-questionnaire assesses the extent to which individuals perceive daily life situations as unpredictable, uncontrollable, and overwhelming. A high internal consistency of Cronbach alpha .77 was reported.

Digital Therapeutic alliance

The moderating role of DTA on mental health outcomes was assessed using the German version of the mobile Agnew Relationship Measure (mARM-G) (Von Wulffen, Marciniak et al., 2023). The mARM-G is a validated self-report instrument developed to measure DTA in the context of mobile interventions. It consists of 25 items rated on a 7-point Likert scale and captures five central dimensions of TA: (1) the bond between the app and user (α=.76); (2) the partnership or collaboration between user and app (α=.77); (3) user confidence in the app's competency (α=.83); (4) the openness or freedom the user feels to personally disclose (α=.61); and (5) the client or user initiative (feelings of control and empowerment) *(*α=.42 but α=.55 if item 20 removed). The internal consistency for the entire scale was excellent (α=.91). Von Wulffen, Marciniak et al. (2023) reported a negative loading for item 20 and advised that it be used with caution or removed from analyses.

Statistical Analyses

All analyses were performed in R (version 4.4.2) using RStudio (version 4.4.1).

Raw datasets from baseline, follow-up, and alliance (mARM-G) questionnaires were merged and cleaned by removing cases without valid baseline or follow-up data. Composite scores of the questionnaires were computed for depressive symptoms (BDI), anxiety symptoms (STAI), and perceived stress (PSS). For the mARM-G questionnaire, total and subscale scores (bond, partnership, confidence, openness, and initiative) were calculated as the mean of their corresponding items. Descriptive statistics (means and standard deviations) were calculated for each time point.

The distributions of outcome variables were assessed visually using histograms and Q–Q plots. When normality assumptions were violated, appropriate transformations were applied: logarithmic for BDI, square-root for STAI, and no transformation for PSS, which was approximately normal. Normality was further evaluated with the change in skewness and kurtosis after transformation.

Within-subject consistency and heterogeneity were assessed using intraclass correlation coefficients (ICC[1] and ICC[2]) derived from a two-way random effects model.

To address the first research question, whether general DTA, moderated the effects of ReApp on mental health outcomes, linear mixed-effects models (LMMs) were performed using the *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages. Random intercepts were included for participants to account for repeated measures. Separate models were fitted for each outcome (BDI, STAI, and PSS), with Time × mARM-G interaction specified as fixed effect. The second research question, which was exploratory, assessed whether the five subscales of DTA, namely, bond, partnership, confidence, openness, and client initiative, varied in moderating reductions in mental health outcomes. Therefore, additional moderation models were

conducted for each outcome (BDI–II, STAI, and PSS), replacing the general DTA score with each mARM–G subscale in the Time × Subscale interaction term. Partial η^2 effect sizes and their 90% confidence intervals were calculated.

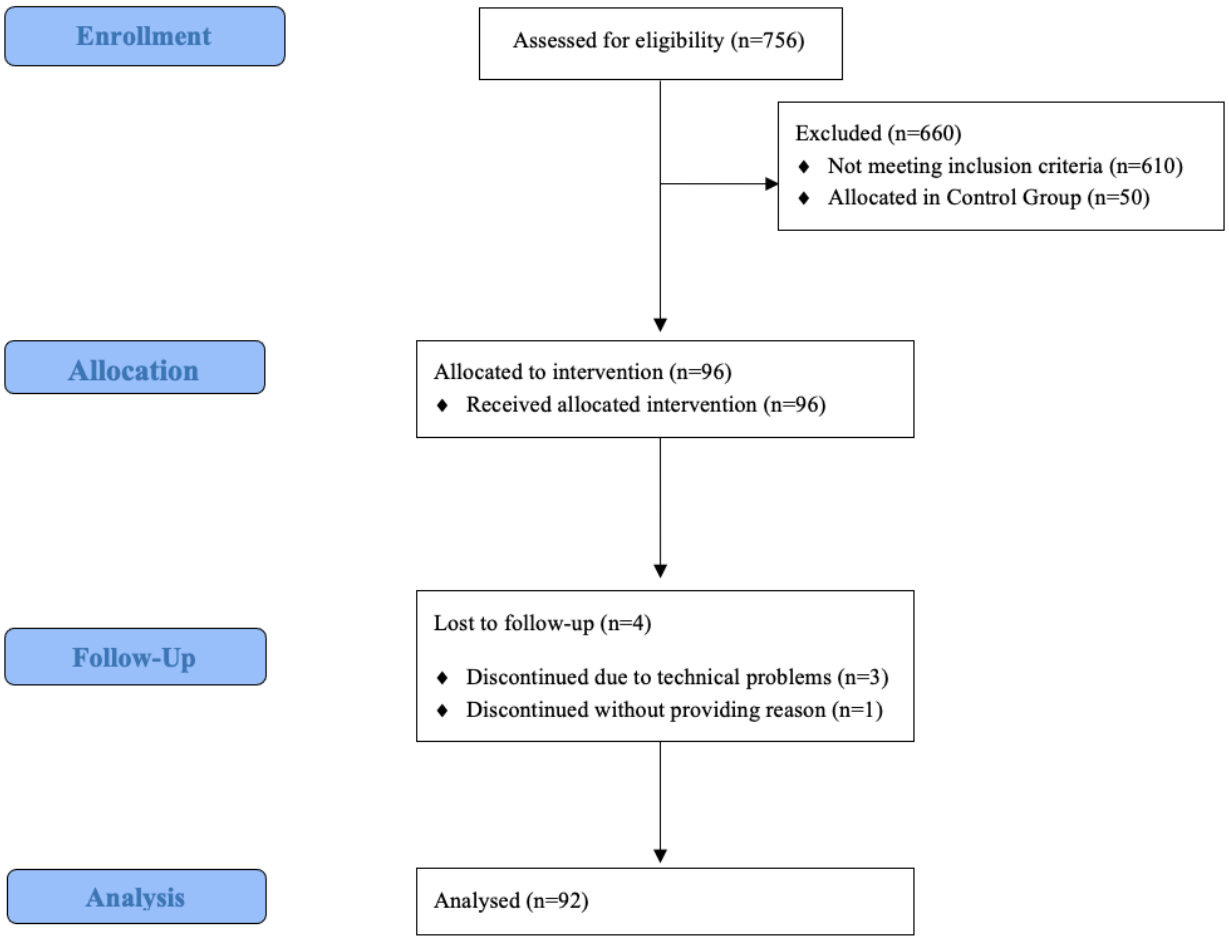
An additional analysis beyond the research questions was performed to assess the degree of overlap among the alliance subscales. The intercorrelations between the five mARM-G subscales were examined to determine whether they represented sufficiently distinct constructs to justify their use as separate moderators.

Results

Sample

Of the N = 756 individuals who completed the online screening, n = 146 met the inclusion criteria and were randomly assigned to one of three groups. Fifty participants in the self-monitoring control condition were excluded from the present analyses (see Marciniak et al., 2023). The two intervention groups (Burst and Continuous) from Marciniak et al. (2024), were merged into one. Four participants withdrew from the study due to technical issues (i.e., the app was not compatible with their phones), and one participant discontinued participation without giving a reason. In total, ninety-two participants completed the study procedure (see Figure 1).

Figure 1



Of the 92 participants, 88 identified as female and four as male, with no participants reporting another gender (see Table 1). The participants' mean age was 22.01 years (*SD* = 2.42), ranging from 18 to 29 years (see Table 1). All participants were German-speaking students enrolled in higher education institutions. Means (*M*) and standard deviations (*SD*) for the outcome measures (PSS, STAI, and BDI-II) at baseline and follow-up are presented in Table 1.

Table 1

Demographic Characteristics of the Sample

Measure	**n (%) or M (SD)**
Student	92 (100%)

Gender - Female	88 (95.7%)
Gender - Male	4 (4.3%)
Gender - Diverse	0 (0%)
Age (in years)	22.01(2.42)
Baseline score in PSS	20.29 (5.72)
Baseline score in STAI	42.86 (9.40)
Baseline score in BDI-II	13.95 (8.71)
Follow-up score in PSS	18.71 (5.81)
Follow-up score in STAI	40.59 (8.71)
Follow-up score in BDI-II	9.95 (8.18)

Note. PSS = Perceived Stress Scale; STAI = State–Trait Anxiety Inventory; BDI–II = Beck Depression Inventory–II. Values represent counts (*n*), percentages (%), and means (*M*) with standard deviations (*SD*) for the current sample after merging both intervention conditions (Burst and Continuous).

Perceived stress scale

Preliminary Analyses

The intraclass correlation coefficient was significant ICC(1) = 0.36, $p < .001$, ICC(2) = 0.52, indicating moderate consistency across measurement occasions, with a substantial amount of variability remaining within individuals. Visual inspection indicated that the distribution of PSS scores was approximately normal, and therefore no transformation was required (see Appendix, Table A1).

Therapeutic Alliance Moderation

The main effect of Time was not significant, $\beta = 3.80$, $p = .339$, 90% CI [−2.71, 10.31], indicating no overall change in perceived stress from baseline to follow-up (See Figure 2A, Table 2). The main effect of Therapeutic Alliance (mARM-G total) was also not significant, $\beta = -0.37$, $p = .657$, 90% CI [−1.71, 0.99], suggesting that the level of alliance did not relate to perceived stress levels (See Table 2). The Time × Alliance interaction was not significant, $\beta = -1.27$, $p = .171$, 90% CI [−2.78, 0.24], $\eta^2 = .02$, 90% CI [0.00, 1.00], indicating that therapeutic alliance did not moderate changes in perceived stress over time (See Table 2, Figure 3).

Subscales Moderation

There was a significant main effect of the Openness subscale of the therapeutic alliance on perceived stress, $\beta = -2.12$, $p = .04$, 95% CI [−3.80, −0.43], $\eta^2 = .042$, indicating that higher perceived openness in the therapeutic alliance was associated with lower levels of perceived stress (see Table 4). However, the interaction between Time and Openness was not significant, $\beta = 0.86$, $p = .458$, 95% CI [−1.04, 2.76], $\eta^2 = .006$, indicating that Openness did not moderate changes in perceived stress over time (see Table 4).

For the Bond subscale, the interaction between Time and Bond was not significant, $\beta = -0.24$, $p = .668$, 95% CI [−2.32, 0.31], $\eta^2 = .037$, suggesting that the emotional connection with the intervention did not moderate changes in perceived stress (See Table 7). Similarly, the Partnership subscale did not significantly moderate the effect of Time, $\beta = -0.91$, $p = .259$, 95% CI [−1.57, 0.64], $\eta^2 = .032$, indicating that collaboration or teamwork with the intervention did not influence stress reduction (See Table 6).

For the Confidence subscale, no significant moderation effect was found, $\beta = -0.76$, $p = .424$, 95% CI [−2.32, 0.80], $\eta^2 = .007$, showing that confidence in the program's effectiveness did not moderate changes in perceived stress across time (See Table 5). Finally, the Initiative

subscale did not significantly moderate changes in stress, $\beta = -0.80$, p = .245, 95% CI [−1.91, 0.32], $\eta^2 = .015$, indicating that proactive engagement in the intervention was not significantly related to changes in perceived stress (See Table 3).

Figure 2

Mental health outcomes of the ReApp intervention

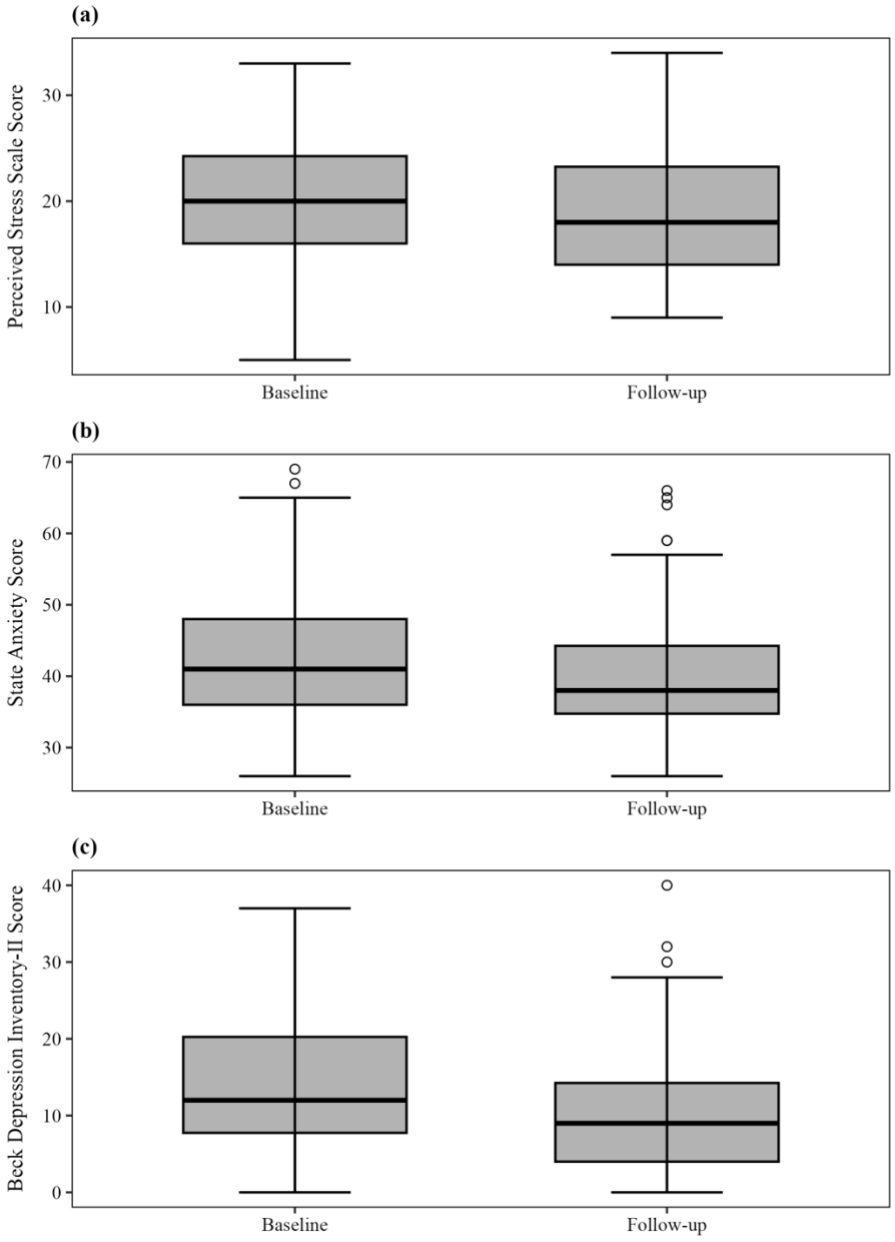


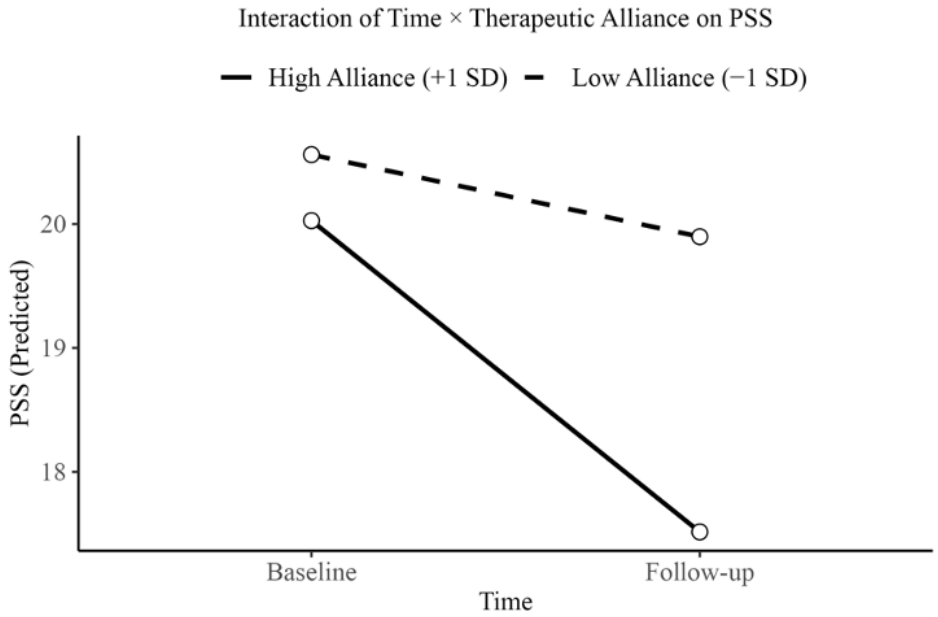
Table 2

Results of Interaction Analyses (Total Alliance Moderation)

Outcome	Value (β)	SE	DF	t-Value	p-Value	Eta Squared [90% CI]
Perceived stress (PSS)						
Intercept	21.85	3.54	157.84	6.17	< .001	
Time	3.80	3.96	90	0.96	0.339	0.010[0.00, 1.00]
Alliance	-0.37	0.82	157.84	-0.45	0.667	0.023[0.00, 01.00]
Time × Alliance	-1.27	0.92	90	-1.38	0.171	0.021[0.00, 1.00]
Anxiety symptoms (STAI)						
Intercept	7.08	0.42	144.29	17.02	< .001	
Time	0.19	0.42	90	0.46	0.646	0.002[0.00,1.00]
Alliance	-0.14	0.10	144.29	-1.40	0.164	0.048[0.01, 01.00]
Time × Alliance	-0.09	0.10	90	-0.89	0.378	0.009[0.00, 01.00]
Depressive symptoms (BDI-II)						
Intercept	2,58	0.51	134.17	5.02	< .001	
Time	-0.25	0.47	90	-0.54	0.593	0.003[0.00, 1.00]
Alliance	-0.02	0.12	134,17	-0.18	0.854	0.002[0.00, 1.00]
Time × Alliance	-0.04	0.11	90	-0.37	0.716	0.001[0.00, 1.00]

Figure 3

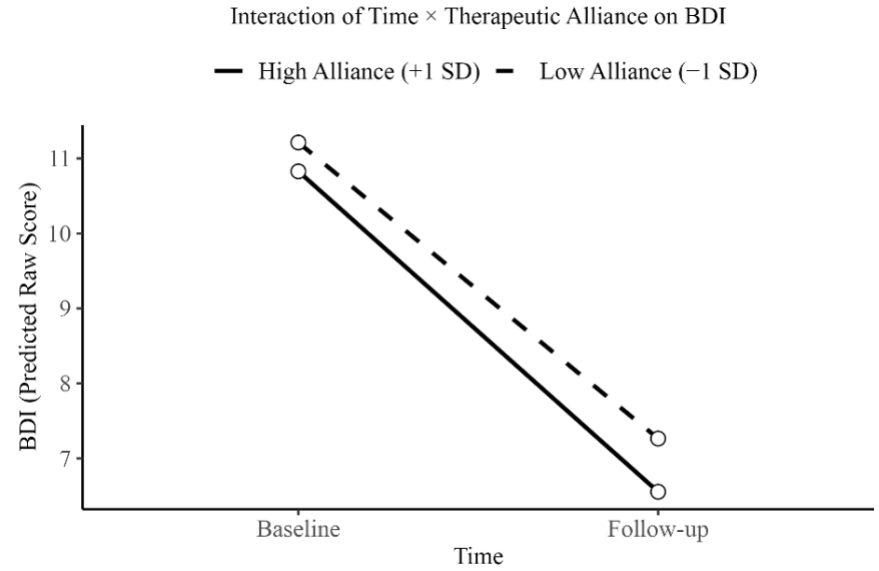
Interaction of Time × Therapeutic Alliance on PSS



Note. Predicted perceived stress (PSS) as a function of Therapeutic Alliance (mARM-G total) at Baseline and Follow-up. The dashed line represents low alliance (−1 *SD*), and the solid line represents high alliance (+1 *SD*).

Figure 4

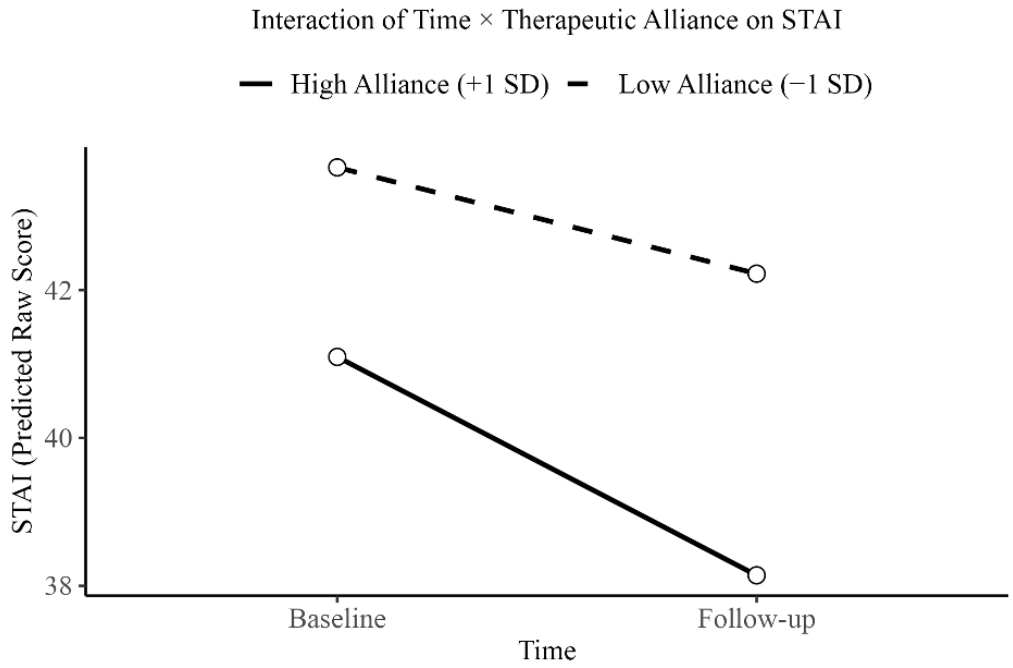
Interaction of Time × Therapeutic Alliance on BDI



Note. Predicted depressive symptoms (BDI) as a function of Therapeutic Alliance (mARM-G total) at Baseline and Follow-up. The dashed line represents low alliance (−1 *SD*), and the solid line represents high alliance (+1 *SD*).

Figure 5

Interaction of Time × Therapeutic Alliance on STAI



Note. Predicted anxiety symptoms (STAI) as a function of Therapeutic Alliance (mARM-G total) at Baseline and Follow-up. The dashed line represents low alliance (−1 *SD*), and the solid line represents high alliance (+1 *SD*).

BDI-II

Preliminary Analyses

The intraclass correlation coefficient was significant ICC(1) = 0.49, $p < .001$, ICC(2) = 0.52, indicating a moderate heterogeneity and reliability of the scores. Due to the non-normal distribution of residuals, the data was transformed. Visual inspection, along with assessments of skewness and kurtosis, indicated that the model using log-transformed data provided the best fit (see Appendix, Table A1).

Total Alliance Moderation

The main effect of Time was not significant, $\beta = -0.25$, $p = .593$, 90% CI [−1.02, 0.52], indicating no significant change in depressive symptoms across time (See Figure 2C, Table 2). The main effect of Therapeutic Alliance (mARM-G total) was not significant, $\beta = -0.02$, p

= .854, 90% CI [−0.22, 0.17], indicating that the level of alliance did not relate to depression symptom levels (See Table 2). The Time × Alliance interaction was also not significant, β = −0.04, p = .716, 90% CI [−0.22, 0.14], η^2 = .00, 90% CI [0.00, 1.00], indicating that alliance did not moderate the effect of time on depressive symptoms (See Table 2, Figure 4).

Subscales Moderation

For the Bond subscale, the interaction between Time and Bond was not significant, β = 0.08, p = .301, 95% CI [−0.07, 0.23], η^2 = .012 (See Table 7). Similarly, the Partnership subscale did not significantly moderate the effect of Time, β = 0.07, p = .849, 95% CI [−0.39, 0.53], η^2 = .000 (See Table 6). The Confidence subscale also did not moderate the Time effect, β = −0.05, p = .761, 95% CI [−0.43, 0.32], η^2 = .000 (See Table 5). The Openness subscale was not a significant moderator, β = −0.09, p = .842, 95% CI [−0.97, 0.79], η^2 = .017 (See Table 4). Finally, the Initiative subscale showed no significant moderation effect, β = 0.00, p = .996, 95% CI [−0.13, 0.13], η^2 = .000 (See Table 3).

STAI

Preliminary Analyses

The intraclass correlation coefficient was significant ICC(1) = 0.49, p < .001, ICC(2) = 0.66, indicating moderate heterogeneity and low reliability of the scores. As the residuals were not normally distributed, data transformation was applied. Visual inspection and evaluation of skewness and kurtosis suggested that the model using square root-transformed data provided the best fit (see Appendix, Table A1).

Total Alliance Moderation

The main effect of Time was not significant, β = 0.19, p = .646, 90% CI [−0.49, 0.88], suggesting no significant change in anxiety between baseline and follow-up (See Figure 2B,

Table 2). The main effect of Therapeutic Alliance (mARM-G total) was not significant, $\beta = -0.14$, $p = .164$, 90% CI [−0.29, 0.02], indicating that the level of alliance did not relate to anxiety symptom levels (See Table 2). The Time × Alliance interaction was not significant, $\beta = -0.09$, $p = .378$, 90% CI [−0.25, 0.07], $\eta^2 = .01$, 90% CI [0.00, 1.00], suggesting that alliance did not significantly moderate changes in anxiety levels over time (See Table 2, Figure 5).

Subscales Moderation

For the Bond subscale, the interaction between Time and Bond was not significant, $\beta = -0.10$, $p = .152$, 95% CI [−0.22, 0.02], $\eta^2 = .003$ (See Table 7). Similarly, the Partnership subscale did not significantly moderate the effect of Time, $\beta = -0.06$, $p = .668$, 95% CI [−0.12, 0.07], $\eta^2 = .002$ (See Table 6). The Confidence subscale also did not moderate the Time effect, $\beta = -0.13$, $p = .181$, 95% CI [−0.30, 0.03], $\eta^2 = .020$ (See Table 5). The Openness subscale was not a significant moderator, $\beta = -0.10$, $p = .414$, 95% CI [−0.30, 0.10], $\eta^2 = .007$ (See Table 4). Lastly, the Initiative subscale did not significantly moderate the relationship between Time and anxiety, $\beta = -0.04$, $p = .556$, 95% CI [−0.16, 0.08], $\eta^2 = .004$ (See Table 3).

Table 3

Moderation Analysis Results for initiative Subscale

Outcome	Value (β)	SE	DF	t-Value	p-Value	Eta Squared [90% CI]
Depressive symptoms (BDI-II)						
Intercept	2.44	0.40	134.17	6.04	< .001	
Time	-0.42	0.37	90.00	-1.14	0.26	0.014[0.00,1.00]

Initiative	0.01	0.09	134.17	0.12	0.905	0.000[0.00, 1.00]
Time × Initiative	-0.00	0.08	90.00	-0.01	0.996	0.000[0.00,1.00]
Anxiety symptoms (STAI)						
Intercept	6.80	0.33	143.28	20.57	< .001	
Time	0.02	0.33	90.00	0.05	0.957	0.000[0.00,1.00]
Initiative	-0.07	0.07	143.28	-0.90	0.368	0.021[0.00, 1.00]
Time × Initiative	-0.04	0.07	90.00	-0.59	0.556	0.004[0.00, 1.00]
Perceived stress (PSS)						
Intercept	20.91	2.80	157.60	7.48	< .001	
Time	1.98	3.12	90.00	0.64	0.527	0.004[0.00,1.00]
Initiative	-0.14	0.61	157.60	-0.23	0.821	0.012[0.00, 1.00]
Time × Initiative	-0.80	0.68	90.00	-1.17	0.245	0.015[0.00,1.00]

Table 4

Moderation Analysis Results for Openness Subscale

Outcome	**Value (β)**	**SE**	**DF**	**t-Value**	**p-Value**	**Eta Squared [90% CI]**

Depressive symptoms (BDI-II)						
Intercept	2.62	0.66	133.81	3.98	< .001	
Time	0.33	0.60	90.00	0.55	0.585	0.003[0.00,1.00]
Openness	-0.03	0.15	133.81	-0.20	0.842	0.008[0.00, 1.00]
Time × Openness	-0.17	0.14	90.00	-1.26	0.211	0.017[0.00, 1.00]
Anxiety symptoms (STAI)						
Intercept	6.79	0.54	142.76	12.50	< .001	
Time	0.27	0.54	90.00	0.49	0.624	0.003[0.00,1.00]
Openness	-0.06	0.12	142.76	-0.52	0.603	0.012[0.00, 1.00]
Time × Openness	-0.10	0.12	90.00	-0.82	0.414	0.007[0.00,1.00]
Perceived stress (PSS)						
Intercept	29.59	4.54	159.36	6.52	< .001	
Time	-5.38	5.14	90.00	-1.05	0.298	0.012[0.00,1.00]
Openness	-2.12	1.02	159.36	-2.07	0.04	0.042[0.005,1.00]
Time × Openness	0.86	1.16	90.00	0.74	0.458	0.006[0.00,1.00]

Table 5

Moderation Analysis Results for Confidence Subscale

Outcome	Value (β)	SE	DF	t-Value	p-Value	Eta Squared [90% CI]
Depressive symptoms (BDI-II)						
Intercept	2.75	0.50	134.27	5.47	< .001	
Time	-0.48	0.46	90.00	-1.05	0.295	0.012[0.00,1.00]
Confidence	-0.06	0.12	134.27	-0.52	0.601	0.003[0.00, 1.00]
Time × Confidence	0.02	0.11	90.00	0.14	0.888	0.000[0.00, 1.00]
Anxiety symptoms (STAI)						
Intercept	6.92	0.41	143.47	16.97	< .001	0.009[0.00,1.00]
Time	0.37	0.41	90.00	0.90	0.369	
Confidence	-0.10	0.10	143.47	-1.02	0.308	0.041[0.004,1.00]
Time × Confidence	-0.13	0.10	90.00	-1.35	0.181	0.020[0.00,1.00]
Perceived Stress (PSS)						
Intercept	21.36	3.49	157.81	6.12	< .001	

Time	1.50	3.90	90.00	0.38	0.702	0.002[0.00,1.00]
Confidence	-0.27	0.85	157.81	-0.31	0.756	0.009[0.00, 1.00]
Time × Confidence	-0.76	0.95	90.00	-0.80	0.424	0.007[0.00,1.00]

Table 6

Moderation Analysis Results for partnership Subscale

Outcome	**Value (β)**	**SE**	**DF**	**t-Value**	**p-Value**	**Eta Squared [90% CI]**
Depressive symptoms (BDI-II)						
Intercept	2.54	0.32	134.15	8.02	< .001	
Time	-0.48	0.29	90.00	-1.65	0.102	0.029[0.00,1.00]
Partnership	-0.01	0.07	134.15	-0.19	0.849	0.000[0.00, 1.00]
Time × Partnership	0.01	0.07	90.00	0.21	0.836	0.000[0.00,1.00]
Anxiety symptoms (STAI)						
Intercept	6.89	0.26	144.40	26.7	< .001	
Time	-0.07	0.26	90.00	-0.25	0.801	0.001[0.00, 1.00]
Partnership	-0.09	0.06	144.40	-1.55	0.124	0.044[0.006, 1.00]

Time × Partnership	-0.03	0.06	90.00	-0.43	0.668	0.002[0.00,1.00]
Perceived Stress (PSS)						
Intercept	20.22	2.19	156.88	9.24	< .001	
Time	2.42	2.43	90.00	1.00	0.321	0.011[0.00,1.00]
Partnership	0.02	0.50	156.88	0.04	0.971	0.013[0.00, 1.00]
Time × Partnership	-0.96	0.56	90.00	-1.72	0.09	0.032[0.001, 1.00]

Table 7

Moderation Analysis Results for Bond Subscale

Outcome	**Value (β)**	**SE**	**DF**	**t-Value**	**p-Value**	**Eta Squared [90% CI]**
Depressive symptoms (BDI-II)						
Intercept	2.46	0.39	133.79	6.41	< .001	
Time	-0.07	0.3	90.00	-0.19	0.853	0.000[0.00,1.00]
Bond	0.01	0.09	133.79	0.06	0.955	0.002[0.00, 1.00]
Time × Bond	-0.08	0.08	90.00	-1.04	0.301	0.012[0.00, 1.00]
Anxiety symptoms (STAI)						
Intercept	6.95	0.31	144.21	22.18	< .001	

Time	-0.02	0.31	90.00	-0.05	0.958	0.000[0.00, 1.00]
Bond	-0.10	0.07	144.21	-1.44	0.152	0.041[0.005, 1.00]
Time × Bond	-0.04	0.07	90.00	-0.51	0.612	0.003[0.00,1.00]
Perceived Stress (PSS)						
Intercept	20.66	2.65	157.02	7.80	< .001	
Time	3.73	2.95	90.00	1.27	0.208	0.018[0.00,1.00]
Bond	-0.09	0.60	157.02	-0.14	0.886	0.022[0.00, 1.00]
Time × Bond	-1.24	0.67	90.00	-1.85	0.067	0.037[0.003, 1.00]

Alliance Subscale Correlations

All mARM-G subscales were positively and significantly correlated ($p < .001$). Bond showed strong associations with Partnership ($r = .81$), Confidence ($r = .66$), and Initiative ($r = .63$), and a moderate association with Openness ($r = .34$). Partnership was strongly correlated with Confidence ($r = .69$) and Initiative ($r = .72$), and moderately correlated with Openness ($r = .33$). Confidence also showed a strong correlation with Initiative ($r = .70$) and a moderate correlation with Openness ($r = .31$). Openness showed the weakest correlations overall, though still significant, ranging from $r = .31$ to .46. Overall, the results indicate that the mARM-G subscales are highly interrelated (See Table 8).

Table 8

Correlation Matrix for mARM-G Subscales

Subscale	1.	2.	3.	4.	5.

1.Bond	1	0.81***	0.66***	0.34***	0.63***
2.Partnership	0.81***	1	0.69***	0.33***	0.72***
3.Confidence	0.66***	0.69***	1	0.31***	0.70***
4.Openness	0.34***	0.33***	0.31***	1	0.46***
5.Initiative	0.63***	0.72***	0.70***	0.46***	1

Note. Values represent Pearson correlations. *** $p < .001$.

Discussion

The current study investigated whether DTA moderated changes in perceived stress, anxiety symptoms, and depressive symptoms during the use of ReApp, a fully automated EMI targeting PCR.

The primary findings indicated no significant moderation effects of general DTA on changes in depressive symptoms, anxiety symptoms, or perceived stress. Similarly, none of the five alliance subscales moderated symptom improvement. Among the five subscales of DTA, only openness demonstrated a significant main effect, indicating that participants who experienced greater openness with the app reported lower overall perceived stress. However, openness did not moderate changes in stress over time.

Several conceptual and methodological considerations may help explain the absence of moderation effects. Growing evidence suggests that the DTA in fully automated apps may work differently compared to the TA that develops in traditional face-to-face therapy (Ashur et al., 2024; Brotherdale et al., 2024; Taylor et al., 2025; Tong et al., 2022, 2023, 2025). Indeed, previous studies indicate that some core elements of TA, especially the bond subscale, may not function in a meaningful way when no human therapist is present (Tong et al., 2022, 2023;

Brotherdale et al., 2024). Although users often describe digital apps as supportive, helpful and nonjudgmental, they usually do not ascribe human qualities such as warmth or care to them (Tong et al., 2023; Brotherdale et al., 2024). This raises the possibility that DTA and TA are conceptually distinct, meaning that mechanisms underlying traditional alliance may not fully translate to fully automated apps like the ReApp.

As a result of this conceptual difference, the use of TA tools that are adapted for digital contexts, such as the mARM-G, may be limited in fully automated apps. Adapting the digital version of the ARM questionnaire, Berry et al. (2018) cautioned that simply replacing "therapist" with "app" may be insufficient to capture the nuances of the DTA, and similar concerns have been raised regarding some mARM-G items not carrying the equivalent relational meaning to face-to-face therapy (Litt, 2024). For instance, the partnership item "The more I use the app, the more I get out of it" is argued to not translate to the same interpersonal meaning as "The more I work with this person, the more I get out of the relationship". Recent empirical work further supports this interpretation. Ashur et al. (2024) compared traditional TA measures adapted for digital use with tools developed specifically to assess DTA and found that the latter predicted clinical outcomes more strongly. Their findings suggest that conventional TA measures, even when adapted for digital use, may not fully reflect the aspects of DTA that are most relevant in app-based interventions. Instead, only digital-specific factors such as perceived emotional investment and app-related accountability were associated with outcomes, whereas traditional TA components, which are comparable to those captured by the mARM-G, were not. Thus, the mechanisms captured by current measures such as the mARM-G likely reflect merely processes that strengthen TA in traditional psychotherapy, which may not operate in the same way within

DTA. Consequently, moderation effects may remain undetected if the constructs being measured are not those mechanisms that actually drive DTA.

Emerging research has begun to identify what these digital-specific mechanisms may be. Instead of relational warmth or interpersonal bonding, DTA appears to be characterized by elements such as flexibility and emotional experiences (Clarke et al., 2016; Taylor et al., 2025; Tong et al., 2022, 2023). Flexibility refers to being able to use the app at any time, but also to feeling that the app provides a safe and nonjudgmental environment and that users can choose when and how they access support. Emotional experiences refer to users feeling less pressure and guilt because apps do not impose expectations or make them feel like a burden, and to users' sense that greater app interactivity enhances support and strengthens their trust in the app. However, since the mARM-G does not assess flexibility or emotional experiences, this further indicates that measures derived from Bordin's (1979) three-component model may not fully capture the digital-specific dimensions of DTA (Brotherdale et al., 2024). If flexibility and related emotional experiences are central to how users form a sense of alliance with an app, but the measurement tool does not include these aspects, potential moderation effects may remain undetected. This aligns with the increasing recognition that new DTA measures are needed that are specifically developed for digital use (Berry et al., 2018; Goldberg et al., 2021; Malouin-Lachance et al., 2025). Future research may benefit from testing moderation using measures developed specifically for digital interventions and from assessing digital-specific constructs such as flexibility and emotional experience, for example with the newly validated Melbourne-Manchester Digital Therapeutic Alliance Scale (MM-DTA) scale for mental health apps (Tong et al., 2025).

Several limitations need to be considered. The sample consisted of a homogeneous group of German-speaking university students, and most participants were women. These characteristics limit the generalizability of the findings to broader or more diverse populations (Sherman, 2024).

Moderation may have been influenced by the psychometric characteristics of the mARM-G. In the present study, the subscales were strongly intercorrelated, particularly between bond and partnership (r = .81) and between partnership and initiative (r = .72) (see Table 8). Such substantial overlap suggests that participants may not have fully distinguished between the theoretical alliance components, reflecting conceptual overlap in the items while still allowing for some uniqueness among the subscales (Rönkkö & Cho, 2020). This interpretation aligns with critiques that the mARM items are broad and nonspecific (Litt, 2024), as illustrated by confidence items such as “I have confidence in the app and the things it suggests” and “I feel optimistic about my progress,” as well as bond items such as “I feel understood by the app”. When predictors share substantial variance, it becomes difficult to detect unique moderation effects. In addition, the initiative subscale showed low reliability (α = .42) (Von Wulffen, Marciniak et al., 2023). The initiative subscale measures the user’s sense of control, accountability, and active engagement. These aspects are identified in digital alliance research as promising predictors of outcomes (Tong et al., 2023; Ashur et al., 2024). The low reliability of initiative makes the measure less sensitive and may prevent the detection of moderation.

In addition, the intervention period was relatively short. The three-week duration was chosen for feasibility reasons and to reduce participant burden, as users received five daily reminders to use the app. Future studies should use longer or longitudinal designs to ensure that users have enough time to form a deeper alliance with the app (Tong et al., 2025).

Despite these limitations, the study offers several notable strengths. First, it is one of the first studies to examine whether DTA moderates outcomes in a fully automated EMI. By focusing on a single therapeutic component, PCR, the study offers a controlled test of alliance effects without the influence of additional psychotherapeutic techniques.

Second, the study contributes empirical support to the growing view that DTA may need to be conceptualized and measured differently from traditional TA. The absence of moderation effects suggests that traditional frameworks of TA may not fully capture the mechanisms of DTA that matter in automated interventions.

Third, participants were able to initiate the app whenever they felt it was useful. This reflects real-world use and aligns with previous findings that digital interventions give users a high degree of choice and control, which may promote a sense of DTA in fully automated apps (Brotherdale et al., 2024).

Finally, the study was conducted in participants' everyday lives over several weeks, which increases ecological validity. The app was easy to integrate into daily routines, and adherence was high (Marciniak et al., 2024).

In summary, DTA, as measured with the mARM-G, did not moderate changes in perceived stress, anxiety, or depressive symptoms. These findings suggest that DTA may work differently in fully automated apps than in traditional psychotherapy, and that current measures may not capture the processes that matter in digital settings.

The study has several implications. The results contribute to the growing discussion that DTA may require a distinct conceptualization and measurement approach. The study also demonstrates that a single-component digital intervention can be delivered effectively in daily life, with high adherence and high feasibility.

Future studies should investigate DTA using newly developed measures specifically for digital mental health apps. Research should focus on digital-specific processes, such as flexibility and emotional experiences. Longer intervention periods and more diverse samples would help to better understand how users interact with fully automated apps. The research on digital mental health apps remains an emerging field within psychology, and the present findings further demonstrate the substantial potential these interventions hold to increase access to support and reduce waiting times. As research in this area progresses, the conceptual framework of DTA is expected to further develop, supported by the development of measures that more accurately capture its unique, digital-specific features.

References

D'Alfonso, S., Lederman, R., Bucci, S., & Berry, K. (2020). The Digital Therapeutic Alliance and Human-Computer Interaction. *JMIR Mental Health*, *7*(12), e21895. https://doi.org/10.2196/21895

Amo, V., & Lieder, F. (2024). Evaluating the Effectiveness of the InsightApp: a longitudinal randomized controlled trial on anxiety, valued action, and Psychological Resilience (Preprint). *JMIR Mental Health*, *12*, e57201. https://doi.org/10.2196/57201

Ashur, O., Saar, C. R., Brandes, O., & Baumel, A. (2024). Are there unique facets of therapeutic alliance for users of digital mental health interventions? An examination with the eHealth Therapeutic Alliance Inventory. *Internet Interventions*, *38*, 100783. https://doi.org/10.1016/j.invent.2024.100783

Baumel, A., & Kane, J. M. (2018). Examining Predictors of Real-World User Engagement with Self-Guided eHealth Interventions: Analysis of Mobile Apps and Websites Using a Novel Dataset. *Journal of Medical Internet Research*, *20*(12), e11491. https://doi.org/10.2196/11491

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects models usinglme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Berry, K., Salter, A., Morris, R., James, S., & Bucci, S. (2018). Assessing therapeutic alliance in the context of MHealth interventions for mental health problems: Development of the Mobile Agnew Relationship Measure (MARM) questionnaire. *Journal of Medical Internet Research*, *20*(4), e90. https://doi.org/10.2196/jmir.8252

Bordin, E. S. (1979). THE GENERALIZABILITY OF THE PSYCHOANALYTIC CONCEPT OF THE WORKING ALLIANCE. *Psychotherapy: Theory, Research & Practice*, *16*(3), 01437416-197916030–00002.

Brotherdale, R., Berry, K., & Bucci, S. (2024). A qualitative study exploring the digital therapeutic alliance with fully automated smartphone apps. *Digital Health*, *10*. https://doi.org/10.1177/20552076241277712

Clarke, J., Proudfoot, J., Whitton, A., Birch, M., Boyd, M., Parker, G., Manicavasagar, V., Hadzi-Pavlovic, D., & Fogarty, A. (2016). Therapeutic alliance with a fully automated mobile phone and Web-Based intervention: secondary analysis of a randomized controlled trial. *JMIR Mental Health*, *3*(1), e10. https://doi.org/10.2196/mental.4656

Folkman, S., PhD, & Lazarus, R. S., PhD. (1984). *Stress, appraisal, and coping*.

Goldberg, S. B., Baldwin, S. A., Riordan, K. M., Torous, J., Dahl, C. J., Davidson, R. J., & Hirshberg, M. J. (2021). Alliance with an Unguided smartphone app: Validation of the Digital Working Alliance inventory. *Assessment*, *29*(6), 1331–1345. https://doi.org/10.1177/10731911211015310

Heron, K. E., & Smyth, J. M. (2009). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, *15*(1), 1–39. https://doi.org/10.1348/135910709x466063

Horvath, A., Gaston, L., & Luborsky, L. (1993). The therapeutic alliance and its measures. In N.E. Miller, L. Luborsky, J.P. Barber & J.P. Docherty (Eds.), *Psychodynamic treatment research* (pp. 247–273). Basic Books.

Keij, B., Alblas, E., Snijders, B., Lhd, V. T., & Vugts, M. (2024). IZA-deelmonitor. Naar meer hybride zorg: nulmeting. In *Rivm (National Institute for Public Health and the Environment)*. https://doi.org/10.21945/rivm-2024-0094

Klein, E. M., Brähler, E., Dreier, M., Reinecke, L., Müller, K. W., Schmutzer, G., Wölfling, K., & Beutel, M. E. (2016). The German version of the Perceived Stress Scale – psychometric characteristics in a representative German community sample. *BMC Psychiatry*, *16*(1). https://doi.org/10.1186/s12888-016-0875-9

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). LMERTest Package: Tests in Linear Mixed Effects models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Kühner, C., Bürger, C., Keller, F., & Hautzinger, M. (2006). Reliabilität und Validität des revidierten Beck-Depressionsinventars (BDI-II). *Der Nervenarzt*, *78*(6), 651–656. https://doi.org/10.1007/s00115-006-2098-7

Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981). *Das State-Trait-Angstinventar (STAI): theoretische Grundlagen und Handanweisung*. https://fis.uni-bamberg.de/handle/uniba/26756

Litt, M. D. (2024). Do you need a therapist to have a therapeutic alliance? Comment on Benitez et al., "The connection still matters: Therapeutic alliance with digital treatment for alcohol use disorder." *Alcohol Clinical and Experimental Research*, *48*(3), 459–461. https://doi.org/10.1111/acer.15260

Lederman, R., & D'Alfonso, S. (2021). The Digital Therapeutic Alliance: Prospects and Considerations. *JMIR Mental Health*, *8*(7), e31385. https://doi.org/10.2196/31385

Linardon, J., Torous, J., Firth, J., Cuijpers, P., Messer, M., & Fuller-Tyszkiewicz, M. (2024). Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. *World Psychiatry*, *23*(1), 139–149. https://doi.org/10.1002/wps.21183

Litt, M. D. (2024). Do you need a therapist to have a therapeutic alliance? Comment on Benitez et al., "The connection still matters: Therapeutic alliance with digital treatment for alcohol use disorder." *Alcohol Clinical and Experimental Research*, *48*(3), 459–461. https://doi.org/10.1111/acer.15260

Macrynikola, N., Chang, S., & Torous, J. (2025). Is digital alliance associated with engagement & outcomes in guided digital interventions? An analysis of data from two studies. *Journal of Affective Disorders*. https://doi.org/10.1016/j.jad.2025.04.127

Malouin-Lachance, A., Capolupo, J., Laplante, C., & Hudon, A. (2025). Does the Digital Therapeutic Alliance exist?: An Integrative Review (Preprint). *JMIR Mental Health*, *12*, e69294. https://doi.org/10.2196/69294

Marciniak, M. A., Homan, S., Zerban, M., Schrade, G., Yuen, K. S. L., Kobylińska, D., Wieser, M. J., Walter, H., Hermans, E. J., Shanahan, L., Kalisch, R., & Kleim, B. (2024). Positive Cognitive Reappraisal Flexibility Is Associated with Lower Levels of Perceived Stress. *Behaviour Research and Therapy*, *183*, 104653. https://doi.org/10.1016/j.brat.2024.104653

Marciniak, M. A., Shanahan, L., Rohde, J., Schulz, A., Wackerhagen, C., Kobylińska, D., Tuescher, O., Binder, H., Walter, H., Kalisch, R., & Kleim, B. (2020). Standalone Smartphone Cognitive Behavioral Therapy–Based Ecological Momentary Interventions

to Increase Mental Health: Narrative review. *JMIR Mhealth and Uhealth*, *8*(11), e19836. https://doi.org/10.2196/19836

Marciniak, M. A., Shanahan, L., Veer, I., Walter, H., Binder, H., Hermans, E., Timmer, J., Kobylińska, D., Puhlmann, L., Tüscher, O., Kalisch, R., & Kleim, B. (2023). ReApp – an MHealth app increasing reappraisal: Results from two randomized controlled trials (Preprint). *PsyArXiv*. https://doi.org/10.31234/osf.io/u4f5e

Marciniak, M. A., Shanahan, L., Yuen, K. S. L., Veer, I. M., Walter, H., Tuescher, O., Kobylińska, D., Kalisch, R., Hermans, E., Binder, H., & Kleim, B. (2024). Burst versus continuous delivery design in digital mental health interventions: Evidence from a randomized clinical trial. *Digital Health*, *10*. https://doi.org/10.1177/20552076241249267

Netherlands Court of Audit (Rekenkamer) (2020), No Place for Big Problems - Tackling specialist mental healthcare waiting lists, https://english.rekenkamer.nl/publications/reports/2020/06/25/no-place-for-big-problems

Oecd, & Policies, E. O. O. H. S. A. (2023). Netherlands: Country Health Profile 2023. In *State of health in the EU*. https://doi.org/10.1787/3110840c-en

Rönkkö, M., & Cho, E. (2020). An updated guideline for Assessing Discriminant Validity. *Organizational Research Methods*, *25*(1), 6–14. https://doi.org/10.1177/1094428120968614

Sherman, J. W. (2024). There is nothing WEIRD about basic research: The critical role of convenience samples in psychological science. *American Psychologist*, *80*(7), 1073–1080. https://doi.org/10.1037/amp0001443

Snipes, C., Dorner-Ciossek, C., Hare, B. D., Besedina, O., Campellone, T., Petrova, M., Lakhan, S. E., & Pratap, A. (2025). Establishment and maintenance of a digital therapeutic alliance in people living with negative symptoms of schizophrenia: two exploratory Single-Arm studies. *JMIR Mental Health*, *12*, e64959. https://doi.org/10.2196/64959

Taylor, T., D'Alfonso, S., Dolan, M. J. T., Yiend, J., & Jacobsen, P. (2025). How do users of a mental health app conceptualise digital therapeutic alliance? A qualitative study using the framework approach. *BMC Public Health*, *25*(1). https://doi.org/10.1186/s12889-025-23603-5

Teague, S. J., Shatte, A. B. R., Fuller-Tyszkiewicz, M., & Hutchinson, D. M. (2025). User-centred design and evaluation of an mHealth app for fathers' perinatal mental health: a feasibility, acceptability, and usability study. *Behaviour and Information Technology*, 1–15. https://doi.org/10.1080/0144929x.2025.2502474

Tong, F., Lederman, R., D'Alfonso, S., Berry, K., & Bucci, S. (2022). Digital Therapeutic Alliance With Fully Automated Mental Health Smartphone Apps: A Narrative Review. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.819623

Tong, F., Lederman, R., D'Alfonso, S., Berry, K., & Bucci, S. (2023). Conceptualizing the digital therapeutic alliance in the context of fully automated mental health apps: A thematic analysis. *Clinical Psychology & Psychotherapy*, *30*(5), 998–1012. https://doi.org/10.1002/cpp.2851

Tong, F., Lederman, R., D'Alfonso, S., Berry, K., & Bucci, S. (2025). Development of a digital therapeutic alliance scale (MM-DTA) in the context of fully automated mental health apps. *Behaviour and Information Technology*, 1–15. https://doi.org/10.1080/0144929x.2025.2469672

Torous, J., Lipschitz, J., Ng, M., & Firth, J. (2019). Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. *Journal of Affective Disorders*, *263*, 413–419. https://doi.org/10.1016/j.jad.2019.11.167

Vaessen, T., Steinhart, H., Batink, T., Klippel, A., Van Nierop, M., Reininghaus, U., & Myin-Germeys, I. (2019). ACT in daily life in early psychosis: an ecological momentary intervention approach. *Psychosis*, *11*(2), 93–104. https://doi.org/10.1080/17522439.2019.1578401

Van Aubel, E., Vaessen, T., Van Winkel, R., Lafit, G., Beijer-Klippel, A., Viechtbauer, W., Batink, T., Van Der Gaag, M., Van Amelsvoort, T., Marcelis, M., Schirmbeck, F., De Haan, L., Reininghaus, U., & Myin-Germeys, I. (2022). Psychological flexibility and the moderating role of the therapeutic working alliance in Acceptance and Commitment Therapy in Daily Life (ACT-DL) in an early psychosis sample. *medRxiv (Cold Spring Harbor Laboratory)*. https://doi.org/10.1101/2022.01.19.22269524

Versluis, A., Verkuil, B., Spinhoven, P., Van Der Ploeg, M. M., & Brosschot, J. F. (2016). Changing Mental Health and Positive Psychological Well-Being Using Ecological Momentary Interventions: A Systematic Review and Meta-analysis. *Journal of Medical Internet Research*, *18*(6), e152. https://doi.org/10.2196/jmir.5642

Von Wulffen, C., Marciniak, M. A., Rohde, J., Kalisch, R., Binder, H., Tuescher, O., & Kleim, B. (2023). German version of the Mobile Agnew Relationship Measure: Translation and Validation study. *Journal of Medical Internet Research*, *25*, e43368. https://doi.org/10.2196/43368

Wang, P., Zhao, M., & Liu, W. (2025). Rethinking mental health. *Advances in Psychiatry and Behavioral Health*. https://doi.org/10.1016/j.ypsc.2025.01.003

World Health Organization: WHO. (2025, September 2). Over a billion people living with mental health conditions – services require urgent scale-up. *World Health Organization*. https://www.who.int/news/item/02-09-2025-over-a-billion-people-living-with-mental-health-conditions-services-require-urgent-scale-up

Appendix

Table A1

Skewness and Kurtosis Before and After Transformation

Variable	Skewness (Original)	Kurtosis (Original)	Skewness (Transformed)	Kurtosis (Transformed)	Transformation
BDI-II	0.800	0.075	–0.885	0.504	Log transformation
PSS	0.214	–0.567	0.214	–0.567	
STAI	0.808	0.159	0.579	–0.201	Square root transformation