Evidence for bidirectional, cross-lagged associations between alliance and psychological distress in an unguided mobile health intervention

Simon B. Goldberg[1,2*], Zishan Jiwani[1,2], Daniel M. Bolt[3], Kevin M. Riordan[1,2], Richard J. Davidson[2,4,5], Matthew J. Hirshberg[2]

[1]Department of Counseling Psychology, UW-Madison, Madison, WI, USA

[2]Center for Healthy Minds, UW-Madison, Madison, WI, USA

[3]Department of Educational Psychology, UW-Madison, Madison, WI, USA

[4]Department of Psychology, UW-Madison, Madison, WI, USA

[5]Department of Psychiatry, UW-Madison, Madison, WI, USA

*Corresponding author

Abstract

Bidirectional associations between changes in symptoms and alliance are established for in-person psychotherapy. Alliance may play an important role in promoting engagement and effectiveness within unguided mobile health (mHealth) interventions. Using models disaggregating alliance and psychological distress into within- and between-person components (random intercept cross-lagged panel model), we report bidirectional associations between alliance and distress over the course of a 4-week smartphone-based meditation intervention (n=302, 80.0% elevated depression/anxiety). Associations were stable across time with effect sizes similar to those observed for psychotherapy (βs=-.13 to -.14 and -.09 to -.10, for distress to alliance and alliance to distress, respectively). Alliance may be worth measuring to improve the acceptability and effectiveness of mHealth tools. Further empirical and theoretical work characterizing the role and meaning of alliance in unguided mHealth is warranted.

Keywords: working alliance; digital technology; smartphone-based interventions; mobile health

The association between therapeutic alliance and outcomes has been widely studied in psychotherapy. Large-scale meta-analyses have demonstrated a robust association between alliance and outcomes (r=.28; Flückiger et al., 2018). However, a causal interpretation of this association whereby alliance leads to improvement in outcomes has been questioned. It has been suggested that the alliance-outcome association may be a byproduct of patient characteristics that influence both alliance and outcome (e.g., personality psychopathology) or an epiphenomenon of early symptom change (DeRubeis et al., 2005; Flückiger et al., 2013). Given the methodological and ethical difficulty in experimentally manipulating alliance (e.g., assigning patients to high vs. low alliance conditions), researchers have used longitudinal approaches to examine cross-lagged relationships between alliance and outcomes. Although unable to prove causality, these approaches can evaluate the temporal ordering of changes in each construct (Falkenström et al., 2013; Flückiger et al., 2020).

The most definitive study of this kind used data from 17 studies and 5,350 participants (Flückiger et al., 2020). Flückiger et al. found evidence that early alliance predicted post-treatment outcomes, replicating results from prior meta-analyses (e.g., Flückiger et al., 2018). However, in longitudinal models, they detected bidirectional relationships between symptom improvement and alliance in the first seven sessions, consistent with the view that alliance both impacts and is impacted by symptom change.

Not all studies evaluating the link between alliance and outcome have detected these bidirectional relationships. Using a more conservative detrended model accounting for longitudinal changes in symptoms and alliance over time, Whelen and colleagues (2021) failed to find a relationship between alliance and symptom change in 191 patients receiving cognitive behavioral therapy for depression. They concluded that alliance plays at best a small role in producing beneficial outcomes within this treatment context.

Most research on the alliance-outcome association has been focused on interventions delivered in person (Flückiger et al., 2018). Recently there has been exponential growth in public use of mental health apps (Wasil et al., 2020) and an increasingly number of randomized controlled trials (RCTs) testing such approaches (Linardon et al., 2019; Goldberg et al., 2022). The successful application of mobile health (mHealth) interventions has enormous potential to lower the cost of care; increase access by reducing barriers related to geography and provider availability; protect patient privacy; and ultimately to improve mental health at the population level (Steinhubl et al., 2013; Wehmann et al., 2020). There is some evidence supporting the efficacy of mHealth interventions, with meta-analyses of RCTs showing positive effects on common mental health symptoms (e.g., depression, anxiety; Goldberg et al., 2022; Linardon et al., 2019; Weisel et al., 2019). At once, there is clear evidence that mHealth interventions, especially unguided smartphone apps, have high and rapid rates of disengagement (Baumel et al., 2019; Linardon & Fuller-Tyszkiewicz, 2020) and effect sizes are generally smaller than in-person interventions (Goldberg et al., 2022).

It has been proposed that understanding alliance in mHealth may help address key limitations associated with this delivery format (Goldberg, in press; Goldberg et al., in press; Henson et al., 2019; Wehmann et al., 2020). In theory, a strong alliance with the technology itself could help encourage users to engage with unguided interventions (Goldberg et al., in press) which may ultimately produce larger treatment effects (Nahum-Shani et al., 2022). Within the psychotherapy literature, alliance is traditionally conceptualized as patient-therapist agreement on the tasks and goals of psychotherapy along with an emotional bond characterized by trust and acceptance (Horvath & Greenberg, 1989). The degree to which a digital corollary of alliance exists within unguided mHealth interventions is an area of active theoretical and empirical investigation (Henson et al., 2019; Tong et al., 2022). The meaning of the bond dimension within unguided mHealth is especially unclear. Traditional bond items (e.g., "I feel that my therapist appreciates me"; Horvath & Greenberg, 1989) do not readily translate to a relationship with technology. While some have adapted these items to avoid anthropomorphizing technology (e.g., "The app supports me to overcome challenges"; Henson et al., 2019), further work is needed to clarify whether and how bond appears absent a human therapist.

As theoretical work continues to refine our understanding of alliance in an unguided mHealth context, findings from empirical studies examining the association between outcomes and existing measures purported to assess digital alliance within unguided mHealth interventions have been mixed, with some detecting the expected association (e.g., Miloff et al., 2020) and others not (e.g., Kiluk et al., 2014). In a study validating a measure of alliance for use with unguided mHealth interventions (Digital Working Alliance Inventory; Goldberg et al., in press), we used data from the active arm of a recently completed RCT testing a meditation-based smartphone app (Hirshberg et al., in press). We found that digital working alliance was positively associated with participants' use of the meditation app (days of use). Using traditional methods for examining the alliance-outcome association (i.e., examining alliance measured at a particular time point as a predictor of pre-post outcomes; Flückiger et al., 2018), alliance measured in Weeks 3 or 4 of the 4-week intervention period, but not alliance measured in Weeks 1 or 2, were associated with pre-post changes in distress.

An important limitation of previous work examining the alliance-outcome association in mHealth is that, to our knowledge, no analyses have disaggregated alliance and outcomes into within- and between-person components. Associations thus reflect the composite effect (combination of within- and between-person components) which are difficult to interpret (Curran et al., 2014). The observed alliance-outcome associations could be due to a between-person effect (those who tend to report higher alliance scores tend to improve more on outcomes), a within-person effect (alliance scores are rated more highly by a particular individual at times when their symptoms are improving more), or a combination of these (Zilcha-Mano & Fisher, 2022). A key next step for understanding the role of alliance within mHealth is examining linkages between alliance and outcome disaggregated into within- and between-person components. As has been done for alliance in psychotherapy (Flückiger et al., 2020), such analyses can clarify the temporal ordering of linkages between alliance and outcome over time.

The Current Study

The current study aimed to evaluate the associations between alliance and psychological distress disaggregated into within- and between-person components within the context of an unguided mHealth intervention: a meditation-based smartphone app. To do so, we used data drawn from the active arm (n=302) of a recently conducted RCT comparing the Healthy Minds Program (HMP) app with a waitlist control (Hirshberg et al., in press). We used random intercept cross-lagged panel models (RI-CLPM) to evaluate bidirectional associations between alliance and distress assessed weekly during the 4-week intervention.

Transparency and Openness

Preregistration

The RCT from which these data were drawn was preregistered (NCT04426318; https://osf.io/eqgt7). However, the analyses reported here were exploratory and not preregistered.

Data, Materials, Code, and Online Resources

Data and Mplus code are available online (https://osf.io/t8qxm/files/osfstorage).

Reporting

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study.

Ethical Approval

Study procedures were approved by the UW-Madison Institutional Review Board (2020-0533).

Method

Participants and Procedure

Data were drawn from the active arm of an RCT (NCT04426318) testing a meditation-based smartphone app (HMP) in 662 Wisconsin public school employees during the initial months of the COVID-19 pandemic (enrollment occurred between June and August 2020). Efficacy results have been reported elsewhere (Hirshberg et al., in press). The preregistered sample size was 400 which was estimated to provide 80% power to detect a between-group effect size of Cohen's $d \geq 0.38$ assuming 43.4% attrition (Linardon & Fuller-Tyszkiewicz, 2020) with $\alpha = .050$. This magnitude effect is similar to that observed in meta-analyses of meditation-based smartphone apps (e.g., Gál et al., 2021). As noted in the preregistration, more than 400 participants may be recruited if additional funding was secured.

Individuals with prior meditation experience and/or severe depressive symptoms were not eligible. Participants completed baseline measures and were randomly assigned to a waitlist control or to use the HMP app for four weeks. The preregistered primary outcome was psychological distress computed as a composite of depression, anxiety, and stress measures. To allow evaluation of both stability and cross-lagged paths, we examined distress and alliance measured weekly via REDCap during the 4-week intervention period (i.e., four assessment timepoints). Participants were emailed to complete weekly measures (including alliance ratings) regardless of their HMP app use.

Participants randomized to receive HMP who completed at least one of the weekly alliance assessments (n=302) were included. The sample was on average 42.77 years old (SD=10.95); 88.4% were female, 10.6% male, 0.3% gender non-binary, and 0.6% of unknown gender; 89.4% were non-Hispanic White, 2.0% Black, 0.3% Latinx, 1.3% Asian/Pacific Islander, 4.3% multiracial, and 2.6% of unknown race/ethnicity; 89.1% had completed college; 15.9% had an annual income ≤$50,000. The majority of participants (80.0%) reported symptoms of depression and/or anxiety in the elevated range (T-score ≥55 on Patient-Reported Outcomes Measure Information System [PROMIS] Depression and Anxiety scales; HealthMeasures, n. d.) at baseline.

Measures

Alliance

The 6-item Digital Working Alliance Inventory (DWAI; Goldberg et al., in press; Henson et al., 2019) was used to assess participants' alliance with the HMP app. This measure includes items adapted from the Working Alliance Inventory (WAI; Horvath & Greenberg, 1989) for the context of an unguided mHealth intervention. Two items are included from each of the three WAI domains: Task (e.g., "I believe the app tasks will help me to address my problem,"), Goal (e.g., "I trust the app to guide me towards my personal goals,"), and Bond (e.g., "The app supports me to overcome challenges"). Items are rated on a 7-point scale from 1 (*strongly agree*) to 7 (*strongly disagree*). Internal consistency was $\alpha s \geq .88$.

Psychological Distress

Symptoms of depression and anxiety were assessed using the computer adaptive versions (v1.0) of the PROMIS Depression and Anxiety measures (Pilkonis et al., 2011). Both measures have shown strong convergent validity with legacy measures (Choi et al., 2014; Schalet et al., 2014). Items reflect symptoms of depression (e.g., “I felt worthless”) and anxiety (e.g., “I felt fearful”) and are rated based on the past 7 days on a 5-point scale ranging from 1 (*never*) to 5 (*always*). The computer adaptive versions yield a T-score (i.e., population mean=50, *SD*=10). Internal consistency cannot be computed for the computer adaptive PROMIS measures, although fixed form versions of these measures have shown adequate internal consistency (αs≥.90; Pilkonis et al., 2011).

Psychological stress was assessed using the 10-item Perceived Stress Scale (PSS; Cohen & Williamson, 1988). The PSS is a widely used measure of psychological stress that assesses experiences in the past month (e.g., “How often have you felt that you were unable to control the important things in your life?”). Items are rated on a 5-point scale from 1 (*never*) to 5 (*very often*). The 10-item version has shown evidence for convergent and discriminant validity (Roberti et al., 2006). A total score was computed by summing across all items (α=.85).

Based on high correlations between measures of depression, anxiety, and stress seen in previous work (Goldberg et al., 2020), the preregistered primary outcome for the RCT was a composite psychological distress variable. To compute this composite, total scores on the three psychological symptom measures were z-transformed and then averaged (α=.87).

Analyses

We implemented RI-CLPM in Mplus (Version 8.8; Muthén & Muthén, 2017) to examine the relationships between alliance and psychological distress over the four weeks of the intervention period. RI-CLPM has been proposed as a method for addressing the inability of traditional cross-lagged panel models (CLPMs) to account for potential trait-like (i.e., time-invariant) elements within repeated measures (Hamaker, Kuiper, & Grasman, 2015). Traditional CLPMs attempts to account for stability through the inclusion of autoregressive relationships (i.e., controlling for Time [T] - 1 when predicting T), but in fact only account for *temporal* stability as they still assume that individuals vary over time around the same mean (Hamaker et al., 2015). In contrast, RI-CLPM separates within-person processes from trait-like, between-person differences by adding a random intercept. We used maximum likelihood with robust standard errors to handle missing data which is robust to data that are missing at random (Graham, 2009) and deviations from normality (Muthén & Muthén, 2017).

We compared fit across several different models in order to arrive at the most parsimonious but well-fitting model. Our initial model (Model 1) included stability and cross-lagged paths as well as synchronous associations, but no additional constraints. We then examined whether detrending (i.e., modeling temporal changes in distress and alliance; Model 2) improved fit and whether constraining stability and cross-lagged paths to be equivalent across time decreased model fit (Model 3). A final model examined whether removing temporal trends in alliance decreased model fit (Model 4). Absolute fit was evaluated using the root mean square error of approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker-Lewis Fit Index (TLI) which were interpreted using standard criteria (i.e., acceptable fit requires RMSEA<.05, CFI and TLI>.95; Brown, 2015). As models were nested, we compared fit using the log-likelihood χ^2 test. The Akaike and Bayesian Information Criteria were also examined. As a sensitivity analysis, we re-estimated this series of models restricted to those in the clinical range at baseline.

Results

Descriptive statistics for the repeated psychological distress and alliance assessments are reported in Supplemental Materials Table 1. Model results are summarized in Table 1. The initial model (Model 1) that included only stability and cross-lagged paths and synchronous associations fit the data poorly (RMSEA=.13, CFI=.90, TLI=.82). Detrending distress and alliance (Model 2) substantially improved absolute fit (RMSEA=0.00, CFT=1.00, TLI=1.00) as well as relative fit (χ^2[10]=98.12, *p*<.001). Constraining the stability and cross-lagged paths to be equivalent across time (Model 3) did not substantially worsen absolute (RMSEA=0.00, CFT=1.00, TLI=1.00) or relative fit (χ^2[8]= 7.24, *p*=.512). In this model, distress showed a significant decrease over time (random slope *B*=-0.12, *p*<.001) although alliance did not (random slope *B*=0.23, *p*=.179). Thus, we examined a model that removed the alliance random slope (Model 4). This modification did not substantially worsen absolute (RMSEA=0.00, CFT=1.00, TLI=1.00) or relative fit (χ^2[8]=4.57, *p*=.102). The BIC was also lowest for Model 4 (BIC=8006.19).

As Model 4 fit the data best, it was considered the final model and interpreted. As shown in Figure 1, we found evidence of cross-lagged associations from alliance to distress (*B*=-0.015, βs=-.09 to -.10, *p*=.020) as well as from distress to alliance (*B*=-0.87, βs=-.13 to -.14, *p*=.003). Synchronous paths were non-significant (*B*s=-0.06 to -0.27, βs=-.05 to -.25, *p*≥.135), with the exception of week 3 (*B*=-0.20, β=-.21, *p*=.036).

Models 1 to 4 were re-estimated restricting the sample to the portion (80.0%) with clinically elevated symptoms at baseline. Model 4 again emerged as the best fitting model. As for the full sample, we found evidence of cross-lagged associations from alliance to distress (*B*=-0.015, βs=-.10, *p*=.044) as well as from distress to alliance (*B*=-0.95, βs = -.13 to -.14, *p*=.009;

Supplemental Materials Figure 1). Synchronous paths were non-significant (*B*s=-0.10 to -0.26, βs=-.10 to -.24, $p \geq .070$).

Discussion

The delivery of psychological interventions via mHealth has the potential to substantially increase access to evidence-based strategies for promoting mental health (Steinhubl et al., 2013). However, engagement with these tools is often poor (Baumel et al., 2019; Linardon & Fuller-Tyszkiewicz, 2020) and effect sizes modest (Goldberg et al., 2022). A deeper understanding of the therapeutic processes that underlie mHealth interventions may help improve effectiveness.

To our knowledge, this is the first study to investigate potential bidirectional linkages between psychological distress and alliance within the mHealth context. Our findings largely replicate patterns observed within the in-person psychotherapy literature (Flückiger et al., 2020). In particular, we found evidence for bidirectional associations between distress and alliance, such that improvements in distress were linked with increases in alliance, and increases in alliance were linked with improvements in distress. Importantly, these associations emerged in models disaggregating alliance and distress ratings into within- and between-person components as well as when modeling temporal trends (i.e., changes in distress over the course of the intervention period). Effect sizes in the current study were similar to those observed by Flückiger et al. (βs=-.19 vs. -.13 to -.14 in the current study and -.07 vs. -.09 to -.10 in the current study, for cross-lagged paths from distress to alliance and alliance to distress, respectively). Interestingly, we found that constraining cross-lagged paths to be equivalent across time did not decrease model fit. This suggests that linkages between alliance and psychological distress emerge early (e.g., after one week) and may be relatively stable across time. These patterns emerging in tandem argues against the possibility that alliance is merely an epiphenomenon of early symptom change (i.e., alliance ratings increase due to early symptom improvement; DeRubeis et al., 2005; Flückiger et al., 2013) as equivalent evidence was found supporting the notion that symptom change is preceded by alliance improvement. Thus, as concluded by Flückiger et al., results suggest that within unguided mHealth, participants' alliance with a mental health tool and their early symptom change may complement each other, rather than one emerging as more important than the other.

These results have implications for future research and design of mHealth interventions. A potential role for alliance within mHealth supports ongoing theoretical and measure development work clarifying the meaning of the alliance construct in unguided mHealth interventions (Tong et al., 2022) as well as efforts to build tools conducive to users experiencing a strong therapeutic alliance. User-centered design principles (i.e., involving end-users in the design process; Schnall et al., 2016) may be a vital support in these efforts. The development of criteria-based rating scales such as Baumel and colleagues' (2017) Enlight scales may be helpful for guiding intervention development. The Enlight scales have been used to assess key aspects of mHealth interventions including usability (i.e., ease of learning to use the intervention), user engagement (i.e., intervention attracts users to use it), therapeutic persuasiveness (i.e., intervention encourages users to make positive changes in behavior), and therapeutic alliance (i.e., intervention creates an alliance with users). Therapeutic persuasiveness and therapeutic alliance rated using Enlight scales have emerged as particularly strong predictors of real-world usage within unguided mHealth (Baumel et al., 2018). It will be important for further theoretical work to clarify the precise boundaries of therapeutic alliance within unguided mHealth, including differentiation from theoretically proximal constructs such as therapeutic persuasiveness (Baumel et al., 2017).

The unguided mHealth context may be a fruitful place to investigate a causal role for alliance. Given methodological and ethical challenges with manipulating the alliance in psychotherapy (e.g., providing intentionally low alliance psychotherapy), the vast majority of alliance-outcome studies has examined naturally occurring variation in alliance (Flückiger et al., 2018). In a review of studies examining alliance as a mediator of outcomes (Baier et al., 2020), only five studies examined whether alliance mediated the effect of differences across treatment conditions and none were specifically designed to manipulate alliance. However, it may be methodologically and ethically possible to manipulate alliance within unguided mHealth, for example by removing elements designed to augment alliance (e.g., removal of a character who directs users through an mHealth intervention; Baumel et al., 2017).

This study has several important limitations. The sample was predominantly non-Hispanic White and female and participants were drawn from one midwestern state, limiting generalizability. Of those randomized to the intervention arm within the parent RCT (n=344), 12.2% (n=42) did not complete alliance assessments and were thus not included in the current study. These omitted observations may have biased results if data were missing not at random. We included only one brief measure of therapeutic alliance. As the meaning of alliance within an unguided mHealth context is still being clarified, it would have been valuable to include alternative measures of alliance and related constructs (e.g., therapeutic persuasiveness; Baumel et al., 2017). The intervention period was relatively short (4 weeks) and we may have missed important variations in alliance-outcome associations that emerge later. Distress ratings were made in reference to a specific timeframe while alliance ratings were not, which may have impacted the strength of the cross-lagged associations. Our selection of Model 4 as the best fitting model and the conclusion that stability and cross-lagged effects are stable across time may have been due to low statistical power to detect variations in these associations that could be disconfirmed in a large sample.

Nonetheless, results support a role for alliance within the context of unguided mHealth. It will be worthwhile continuing to clarify the meaning of alliance in this context and investigating factors that influence and are influenced by digital alliance. Future research can productively determine factors associated with variability in alliance across users which may be impacted by users' cultural identities (Koo et al., 2016), among other factors. Alliance may prove to be a worthwhile variable to guide intervention customization and ultimately for increasing intervention effectiveness by matching interventions with user characteristics and responsiveness (Collins & Varmus, 2015).

References

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*, 101921.

Baumel, A., Faber, K., Mathur, N., Kane, J. M., & Muench, F. (2017). Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. *Journal of Medical Internet Research, 19*(3), e82.

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.

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

Brown, T. A. (2015). *Confirmatory factor analysis for applied research (2nd ed.).* New York: Guilford Press.

Choi, S. W., Schalet, B., Cook, K. F., & Cella, D. (2014). Establishing a common metric for depressive symptoms: Linking the BDI-II, CES-D, and PHQ-9 to PROMIS depression. *Psychological Assessment, 26*(2), 513-527.

Cohen, S., & Williamson, G. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan & S. Oskamp (Eds.), *The social psychology of health: Claremont Symposium on Applied Social Psychology* (pp. 31–67). Newbury Park, CA: Sage.

Collins, F. S., & Varmus, H. (2015). A new initiative on precision medicine. *New England Journal of Medicine, 372*(9), 793-795.

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: a latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology, 82*(5), 879-894.

DeRubeis, R., Brotman, M., & Gibbons, C. (2005). A conceptual and methodological analysis of the nonspecifics argument. *Clinical Psychology: Science and Practice, 12*, 174-183.

Falkenström, F., Granström, F., & Holmqvist, R. (2013). Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology, 60*(3), 317-328.

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.

Flückiger, C., Grosse Holtforth, M., Znoj, H. J., Caspar, F., & Wampold, B. E. (2013). Is the relation between early post-session reports and treatment outcome an epiphenomenon of intake distress and early response? A multi-predictor analysis in outpatient psychotherapy. *Psychotherapy Research, 23*(1), 1-13.

Flückiger, C., Rubel, J., Del Re, A. C., Horvath, A. O., Wampold, B. E., Crits-Christoph, P., ... & Errázuriz, P. (2020). The reciprocal relationship between alliance and early treatment symptoms: A two-stage individual participant data meta-analysis. *Journal of Consulting and Clinical Psychology, 88*(9), 829-843.

Gál, É., Ștefan, S., & Cristea, I. A. (2021). The efficacy of mindfulness meditation apps in enhancing users' well-being and mental health related outcomes: a meta-analysis of randomized controlled trials. *Journal of Affective Disorders, 279*, 131-142.

Goldberg, S. B. (in press). A common factors perspective on mindfulness-based interventions. *Nature Reviews Psychology*.

Goldberg, S. B., Baldwin, S. A., Riordan, K. M., Torous, J., Dahl, C. J., Davidson, R. J., & Hirshberg, M. J. (in press). Alliance with an unguided smartphone app: Validation of the Digital Working Alliance Inventory. *Assessment*.

Goldberg, S. B., Imhoff-Smith, T., Bolt, D. M., Wilson-Mendenhall, C. D., Dahl, C. J., Davidson, R. J., & Rosenkranz, M. A. (2020). Testing a multi-component, self-guided, smartphone-based meditation app: three-armed randomized controlled trial. *JMIR Mental Health, 7*(11), e23825.

Goldberg, S. B., Lam, S., Simonsson, O., Torous, J., & Sun, S. (2022). Mobile phone-based interventions for mental health: A systematic meta-review of 14 meta-analyses of randomized controlled trials. *PLOS Digital Health, 1*(1), e0000002.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549-576.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102-116.

HealthMeasures. (n. d.). *PROMIS Score Cut Points.* Retrieved on July 5, 2020 from: https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis/promis-score-cut-points

Henson, P., Wisniewski, H., Hollis, C., Keshavan, M., & Torous, J. (2019). Digital mental health apps and the therapeutic alliance: Initial review. *BJPsych Open, 5*(1).

Hirshberg, M. J., Frye, C., Dahl, C. J., Riordan, K. M., Vack, N. J., Sachs, J., Goldman, R., Davidson, R. J., & Goldberg, S. B. (in press). A randomized controlled trial of a smartphone-based well-being training in public school system employees during the COVID-19 pandemic. *Journal of Educational Psychology*.

Horvath, A. & Greenberg, L. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology, 36*(2), 223-233.

Kiluk, B. D., Serafini, K., Frankforter, T., Nich, C., & Carroll, K. M. (2014). Only connect: The working alliance in computer-based cognitive behavioral therapy. *Behaviour Research and Therapy, 63*, 139-146.

Koo, K. H., Tiet, Q. Q., & Rosen, C. S. (2016). Relationships between racial/ethnic minority status, therapeutic alliance, and treatment expectancies among veterans with PTSD. *Psychological Services, 13*(3), 317–321.

Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2019). The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry, 18*(3), 325-336.

Linardon, J., & Fuller-Tyszkiewicz, M. (2020). Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *Journal of Consulting and Clinical Psychology, 88*(1), 1-13.

Miloff, A., Carlbring, P., Hamilton, W., Andersson, G., Reuterskiöld, L., & Lindner, P. (2020). Measuring Alliance Toward Embodied Virtual Therapists in the Era of Automated Treatments With the Virtual Therapist Alliance Scale (VTAS): Development and Psychometric Evaluation. *Journal of Medical Internet Research, 22*(3), e16660.

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide. Eight Edition.* Los Angeles, CA: Muthén & Muthén.

Nahum-Shani, I., Shaw, S. D., Carpenter, S. M., Murphy, S. A., & Yoon, C. (2022). Engagement in digital interventions. American Psychologist, 77(7), 836–852.

Pilkonis, P. A., Choi, S. W., Reise, S. P., Stover, A. M., Riley, W. T., Cella, D., & PROMIS Cooperative Group. (2011). Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS): Depression, anxiety, and anger. *Assessment, 18*(3), 263-283.

Roberti, J. W., Harrington, L. N., & Storch, E. A. (2006). Further psychometric support for the 10-item version of the perceived stress scale. *Journal of College Counseling, 9*(2), 135-147.

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*(2), 243-248.

Schalet, B. D., Cook, K. F., Choi, S. W., & Cella, D. (2014). Establishing a common metric for self-reported anxiety: Linking the MASQ, PANAS, and GAD-7 to PROMIS Anxiety. *Journal of Anxiety Disorders, 28*(1), 88-96.

Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Dieguez, A., Carry, M., ... & Travers, J.

(2016). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *Journal of Biomedical Informatics, 60*, 243-251.

Steinhubl, S. R., Muse, E. D., & Topol, E. J. (2013). Can mobile health technologies transform health care? *JAMA, 310*(22), 2395-2396.

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*, 819623.

Wasil, A. R., Gillespie, S., Patel, R., Petre, A., Venturo-Conerly, K. E., Shingleton, R. M., ... & DeRubeis, R. J. (2020). Reassessing evidence-based content in popular smartphone apps for depression and anxiety: Developing and applying user-adjusted analyses. *Journal of Consulting and Clinical Psychology, 88*(11), 983-993.

Wehmann, E., Köhnen, M., Härter, M., & Liebherz, S. (2020). Therapeutic Alliance in Technology-Based Interventions for the Treatment of Depression: Systematic Review. *Journal of Medical Internet Research, 22*(6), e17195.

Weisel, K. K., Fuhrmann, L. M., Berking, M., Baumeister, H., Cuijpers, P., & Ebert, D. D. (2019). Standalone smartphone apps for mental health: A systematic review and meta-analysis. *npj Digital Medicine, 2*(1), 1-10.

Whelen, M. L., Murphy, S. T., & Strunk, D. R. (2021). Reevaluating the alliance–outcome relationship in the early sessions of cognitive behavioral therapy of depression. *Clinical Psychological Science, 9*(3), 515-523.

Zilcha-Mano, S., & Fisher, H. (2022). Distinct roles of state-like and trait-like patient–therapist alliance in psychotherapy. *Nature Reviews Psychology, 1*(4), 194-210.

Table 1

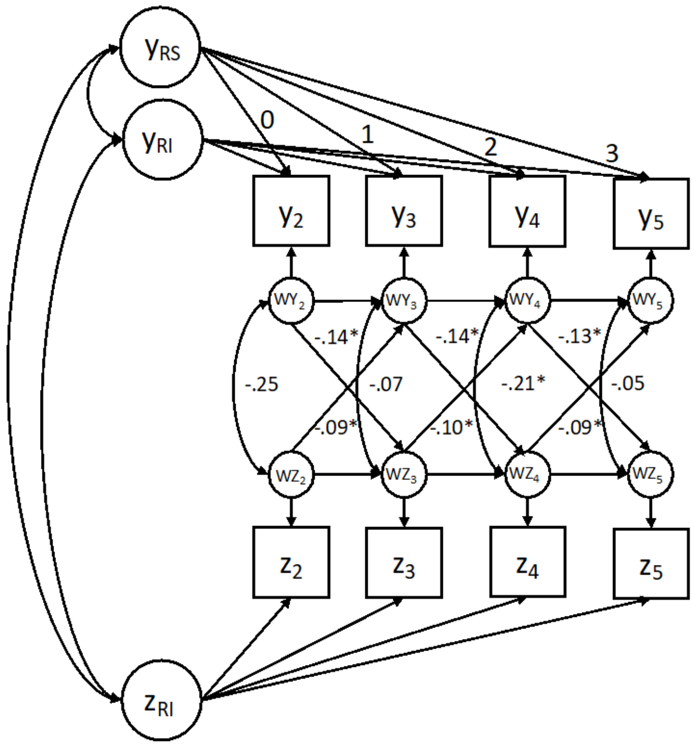
Absolute and Relative Model Fit

					Full sample (n = 302)						
Model	**AIC**	**BIC**	**χ^2**	**df**	**Comparison**	**χ^2 diff**	**χ^2 diff df**	***P***	**RMSEA**	**CFI**	**TLI**
Model 1	7989.74	8093.63	99.81	16	n/a	n/a	n/a	n/a	0.13	0.90	0.82
Model 2	7907.38	8048.37	1.69	6	Model 1	116.17	10	< .001	0.00	1.00	1.00
Model 3	7899.98	8011.29	8.93	14	Model 2	7.60	8	.473	0.00	1.00	1.00
Model 4	7902.30	8006.19	13.50	16	Model 3	3.79	2	.150	0.00	1.00	1.00
					Clinical sample (n = 240)						
Model	**AIC**	**BIC**	**χ^2**	**df**	**Comparison**	**χ^2 diff**	**χ^2 diff df**	***P***	**RMSEA**	**CFI**	**TLI**
Model 1	6321.35	6418.81	88.73	16	n/a	n/a	n/a	n/a	0.14	0.86	0.76
Model 2	6244.47	6376.74	1.74	6	Model 1	107.05	10	< .001	0.00	1.00	1.00
Model 3	6232.35	6336.77	5.00	14	Model 2	3.38	8	.908	0.00	1.00	1.00
Model 4	6233.13	6330.59	8.25	16	Model 3	2.57	2	.277	0.00	1.00	1.00

Note. Model 1 = lagged and cross-lagged paths only; Model 2 = detrended distress and alliance; Model 3 = detrended distress and alliance with paths constrained to be equivalent across time; Model 4 = detrended distress with paths constrained to be equivalent across time. AIC = Akaike Information Criterion; Bayesian Information Criterion; df = degrees of freedom; Comparison = comparison model for assessing relative fit; χ^2 diff = differences in χ^2 relative to comparison model scaled following Satorra and Bentler (2010); p = p-value for log-likelihood difference test; RMSEA = root mean square error of approximation; CFI = Comparative Fit Index; TLI = Tucker Lewis Fit Index.

Figure 1

Random Intercept Cross-lagged Panel Model (RI-CLPM) Examining Longitudinal Linkages Between Changes in Distress and Changes in Alliance



Note. y = distress; z = alliance; RI = random intercept; RS = random slope (with numbers 0 to 3 indicating factor loadings); W = within-person component. Values represent standardized effect sizes (STDYX standardization) associated with cross-lagged and synchronous paths derived from Model 4 (stability and cross-lagged paths constrained to be equivalent over time, distress detrended). Note that the RI-CLPM with an added slope factor is equivalent to the latent curve model with structured residuals (LCM-SR; Curran et al., 2014). $n = 302$. $^*p < .05$

Supplemental Materials Table 1

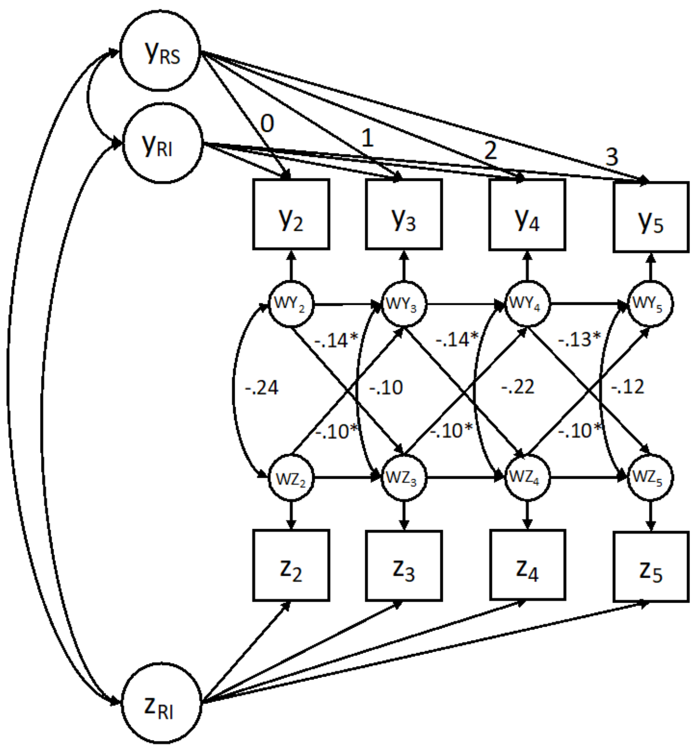
Distress and Alliance Intercorrelations and Descriptive Statistics

	$Distress_2$	$Distress_3$	$Distress_4$	$Distress_5$	$Alliance_2$	$Alliance_3$	$Alliance_4$	$Alliance_5$
$Distress_2$	-							
$Distress_3$	.85***	-						
$Distress_4$	.82***	.86***	-					
$Distress_5$	.78***	.81***	.88***	-				
$Alliance_2$	.03	.03	.06	.01	-			
$Alliance_3$	-.04	-.09	-.06	-.10	.65***	-		
$Alliance_4$	-.04	-.12	-.16*	-.18**	.64***	.65***	-	
$Alliance_5$	.03	-.02	-.02	-.06	.54***	.60***	.68***	-
n	277	257	257	280	275	256	255	280
% missing	8.28	14.90	14.90	7.28	8.94	15.23	15.56	7.28
Mean	-0.29	-0.40	-0.55	-0.66	33.69	33.58	34.29	34.22
SD	0.89	0.85	0.88	0.89	5.14	5.69	5.23	6.02
Skew	-0.21	-0.12	0.00	-0.12	-0.33	-1.25	-0.71	-1.27
Kurtosis	0.49	0.13	-0.04	0.26	0.06	3.38	0.62	2.71

Note. Subscripts represent week. Intercorrelations and descriptive statistics based on the full sample. Distress = composite of PROMIS Depression, PROMIS Anxiety, and Perceived Stress Scale; Alliance = Digital Working Alliance Inventory. Participants completed four (72.2%), three (15.9%), two (6.3%), or one (5.6%) assessments of the distress measures. Participants completed four (71.9%), three (16.2%), two (5.6%), one (5.6%), or zero (0.7%) assessments of alliance. Distress means were based on a composite variable of the three distress measures standardized (i.e., z-scored) using baseline means and SDs. Negative values reflect reductions in distress relative to the baseline. *$p < .05$, **$p < .01$, ***$p < .001$

Supplemental Materials Figure 1

Random Intercept Cross-lagged Panel Model (RI-CLPM) Examining Longitudinal Linkages Between Changes in Distress and Changes in Alliance in the Clinical Sample



Note. y = distress; z = alliance; RI = random intercept; RS = random slope (with numbers 0 to 3 indicating factor loadings); W = within-person component. Values represent standardized effect sizes (STDYX standardization) associated with cross-lagged and synchronous paths derived from Model 4 (stability and cross-lagged paths constrained to be equivalent over time, distress detrended). $n = 240$. *$p < .05$