

Which adolescents are well-suited to app-based mindfulness training?

A randomized clinical trial and data-driven approach for personalized recommendations

Christian A. Webb¹, Caroline M. Swords²,

Hannah R. Lawrence¹, and Lori M. Hilt²

¹Harvard Medical School & McLean Hospital, Boston, MA

²Lawrence University, Appleton, WI

Corresponding author:

Christian A. Webb, Ph.D.

Harvard Medical School & McLean Hospital

115 Mill Street, Belmont, MA 02478

Email: cwebb@mclean.harvard.edu

Funding: Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under Award Number R15MH116303 to LMH. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The first author (CAW) was partially supported by R01 MH116969, the Tommy Fuss Fund and a NARSAD Young Investigator Grant from the Brain & Behavior Research Foundation. HRL was supported by a grant from the American Foundation for Suicide Prevention (PRL-0-095-19). The opinions and assertions contained in this article should not be construed as reflecting the views of the sponsors.

Conflict of Interest: The authors declare that they have no conflict of interest.

Draft version 1, 6/30/21. This paper is currently under peer review. Please do not copy or cite without author's permission

Abstract

Objective: Rumination heightens risk for depression and anxiety, which increase substantially during adolescence. Smartphone apps offer a convenient and cost-effective means for adolescents to access mindfulness training, which may reduce rumination. Despite their increasing popularity, it is unclear which adolescents benefit from mindfulness apps.

Methods: Adolescents ($n = 152$) with elevated trait rumination were randomly assigned to 3 weeks of app-based mindfulness training or a mood monitoring control. Multilevel models tested group differences in state rumination change, assessed via ecological momentary assessment. Baseline adolescent characteristics were submitted to elastic net regularization models to develop a “Personalized Advantage Index” indicating an individual’s expected outcome from the mindfulness app relative to the mood monitoring control. Finally, we translated a predictive model (developed in an external sample) for personalized recommendations of expected benefit from the mindfulness app.

Results: Adolescents in the mindfulness app condition reported significantly greater reductions in rumination than adolescents in the control condition. Individuals predicted to have better outcomes from the mindfulness app relative to mood monitoring had significantly greater reductions in rumination if randomly assigned to the mindfulness condition. Between-condition differences in outcome were not significant for adolescents predicted to have better outcomes in the mood monitoring condition.

Conclusions: Findings support the efficacy of a mindfulness app to reduce state rumination in adolescents, particularly among adolescents high in trait rumination. A predictive model is put forth, which could be used to objectively communicate expected mindfulness app outcomes to adolescents prior to engagement in app-based mindfulness training.

Keywords: mindfulness; smartphone app; rumination; adolescents; prediction

Public Health Significance Statement

An increasing number of adolescents are turning to mindfulness apps as a means of coping with stress. This study demonstrates the potential of a data-driven approach – based on baseline clinical and demographic characteristics – to inform which adolescents are most likely to benefit from a mindfulness app.

Which adolescents are well-suited to app-based mindfulness training?

A randomized clinical trial and data-driven approach for personalized recommendations

Rates of anxiety and, in particular, depression surge during the adolescent years (Avenevoli et al., 2015; Merikangas et al., 2010). By the end of adolescence, an estimated 15% of youth have experienced at least one episode of major depression, and 32% have developed one or more anxiety disorders (Avenevoli et al., 2015; Merikangas et al., 2010). In addition, a recent study indicated that rates of emotional disorders and psychological distress among adolescents have increased during the past decade (Twenge et al., 2019). Despite these alarming statistics, only approximately half of adolescents with a mental disorder receive treatment (Avenevoli et al., 2015; Geiger & Davis, 2019; Whitney & Peterson, 2019), and 28% to 75% of those youth who do receive services do not complete treatment (De Haan et al., 2013).

Digital technologies, including smartphone apps, provide one avenue to help bridge the treatment gap and to substantially increase access to evidenced-based interventions for adolescents with anxiety or depressive symptoms. Apps may help overcome barriers associated with traditional, in-person psychotherapeutic or psychiatric treatment (e.g., geographical barriers, lack of available local clinicians, long waitlists, prohibitive costs, and stigma or discomfort disclosing mental health problems or engaging in in-person treatment; Grist et al., 2017). Perhaps not surprisingly, many adolescents have turned to mental health apps as a means of coping with stress and symptoms of anxiety and/or depression. Most U.S. adolescents now own a smartphone, which offers an accessible and low cost means for teens to access mental health content and therapeutic skills training (e.g., via mindfulness or cognitive behavioral therapy apps). In fact, whereas 67% of 13- to 18-year-olds owned a smartphone in 2015, that figure rose to 84% by 2019 (Rideout & Robb, 2019). According to current estimates, there are

approximately 10,000 apps focused on mental health (Lagan et al., 2021; Torous et al., 2018). A recent analysis of mental health apps available from the Apple iOS and Google Play stores revealed that the most common features were mood monitoring, journaling, and mindfulness training (Lagan et al., 2021). The most frequently used apps for depression (Headspace; 5 million monthly active users) and anxiety (Calm: 9 million monthly active users) are mindfulness apps (Wasil et al., 2020). Eleven percent of adolescents – and 18% of teens with elevated depressive symptoms – have used a mindfulness app (Rideout & Fox, 2018), of which there are over 260 available from the Apple iOS store and Google Play (Mani et al., 2015).

Given its emphasis on the cultivation of present moment attention and meta-awareness of thought content and processes, mindfulness training may be particularly beneficial for adolescents who ruminate (i.e., engage in repetitive negative thinking focused on the past; Hilt & Swords, 2021; Webb et al., 2021). A recent study found a mindfulness app to be beneficial for adolescents with elevated rumination (Hilt & Swords, 2021), with significant reductions in both trait rumination and anxiety symptoms. In addition to the latter mindfulness app trial, several studies have reported that conventional (i.e., in-person) mindfulness training reduces rumination among children and adolescents (Ames et al., 2014; Mendelson et al., 2010; van de Weijer-Bergsma et al., 2014), including brief (e.g., 8 minute) mindfulness exercises (Hilt & Pollak, 2012; Villa & Hilt, 2014). However, despite the growing popularity of mindfulness apps among adolescents – in particular among those adolescents with elevated stress, anxiety or depressive symptoms (Rideout & Fox, 2018) – little is known regarding who benefits from these apps. A critical question is: *Which* adolescents are well-suited to app-based mindfulness training, versus those who may experience limited or no benefit (or even deterioration or other adverse effects; Goodman et al., 2017; Van Dam et al., 2018)?

In a recent trial of a mindfulness app for ruminating adolescents, a combination of baseline characteristics including higher levels of trait rumination and lower emotional suppression predicted greater reduction in rumination from a 3-week course of app-based mindfulness training (Webb et al., 2021). However, this study was a single-arm trial and did not include a comparison or control condition. Thus, it is unknown whether the identified predictors were specific indicators of one's likelihood of benefiting from a mindfulness app (i.e., "prescriptive predictors"), or represented general or intervention non-specific predictors of outcome (i.e., "prognostic predictors"; Cohen & DeRubeis, 2018; Webb et al., 2018).

The Present Study

Within the context of a randomized clinical trial (RCT; $n = 152$), the present study tested whether the combination of baseline clinical and demographic characteristics predicted which adolescents with elevated trait rumination were likely to benefit from a mindfulness app relative to a mood monitoring control condition. As an initial step to testing external validity, a predictive model was developed in a sample from a former trial (Webb et al., 2021) and applied to adolescents in the present trial to test the model's predictive performance in a new sample of adolescents where comparisons between a mindfulness app and control condition was possible. We tested whether the subset of adolescents predicted to have greater reduction in rumination from the mindfulness app (relative to the control condition) had better outcomes if they had in fact been randomly assigned to the mindfulness condition. In contrast, for the subset of participants predicted to experience greater reductions in rumination in the control group (relative to the mindfulness condition), we expected no significant between-condition differences in observed outcomes. This study and our analytic plans were pre-registered (<https://tinyurl.com/x58phrsu>).

Method

Participants

The sample consisted of 152 community adolescents (12-15; M age = 13.71 years, SD = 0.90) with elevated levels of trait rumination, which were assessed during a phone screen. Two questions from the rumination subscale of the Children's Response Styles Questionnaire (CRSQ; Abela et al., 2002) were used to screen for elevated trait rumination ("When you feel sad or stressed, do you think about a recent situation, wishing it had gone better?" and "When you feel sad or stressed, do you think "why can't I handle things better?"). Similar to recent studies (Hilt & Swords, 2021; Webb et al., 2021), adolescents were eligible if their average score indicated that they "sometimes," "often," or "almost always" ruminate in response to sadness or stress. Based on these criteria, forty adolescents were not eligible to participate (see **Supplemental Figure 2** for CONSORT flow diagram). **Table 1** presents clinical and demographic characteristics of the sample.

Procedure

This study was approved by the [omitted for blinded submission] IRB. Participants were recruited from a moderately sized midwestern community from 2019-2020 for a study on how a mobile app may help adolescents cope with negative emotions. Participants were recruited through word of mouth, in-person and online advertisements, study posters, and letters sent to the parents/guardians of sixth through ninth grade students enrolled in the local public-school district. Interested families completed a phone screen to determine eligibility. If eligible, parents/guardians and adolescents were scheduled to visit the lab (see Supplement). Parents/guardians and adolescents provided informed consent/assent and completed baseline questionnaires. Adolescents were subsequently randomly assigned to use a version of the app

(i.e., mindfulness with mood monitoring or mood monitoring-only control), which was then downloaded onto their smartphone. Participants who did not have their own smartphone borrowed one from the lab ($n = 19$). Participants were asked to use the app three times a day for three weeks. When participants used the app for the first time, they were prompted to enter their sleep and wake time, which provided a window of time (i.e., morning, late afternoon, and before bedtime) during which the app could notify participants without interrupting sleep. If participants were busy at the time they received a notification, they were asked to use the app as soon as they were able. To encourage app use, participants earned a \$5 bonus each week they used the app 21 times or more. A counter within the app helped adolescents keep track of how many times they had used it. Additionally, study personnel emailed parents weekly regarding the number of times adolescents used the app that week, and whether they had earned the weekly bonus.

Mood Monitoring Control Condition

Participants randomized to the mood monitoring control group only completed ecological momentary assessment (EMA) survey questions. Each time participants used the app, they were prompted to report on state rumination (i.e., focus on emotions and problems) and state mood (i.e., sadness, anxiety, happiness, and calmness) on a 0 (*not at all*) to 100 (*extremely*) scale.

Mindfulness Condition

Participants randomized to the mindfulness condition answered the same mood monitoring EMA survey questions as those randomized to the control condition. However, they were also randomly assigned to receive a brief mindfulness exercise two out of every three EMA surveys they completed (i.e., 67% chance of receiving a mindfulness exercise). Higher levels of state negative affect (i.e., sadness or anxiety ratings ≥ 90) increased the chances that an adolescent received a mindfulness exercise to 85%. When prompted to complete a mindfulness

exercise, participants indicated how much time they had available (i.e., about 1, 5, or 10 minutes) and were randomly assigned to one of several mindfulness activities within the timeframe selected. During the one-minute exercises, participants were prompted by written instructions to focus on their breath, sounds in the environment, or physical sensations displayed alongside a one-minute timer. All other mindfulness exercises were guided audio recordings ranging from 3-12 minutes, which similarly focused on breath, sounds, or body scans. These guided audio recordings were selected from a library of free and publicly available mindfulness exercises and are representative of mindfulness-based interventions commonly used for this age group. Immediately following the completion of the brief mindfulness exercise, participants again answered the same EMA questions.

Transparency and Openness

This study's analysis plan was preregistered; see <https://tinyurl.com/x58phrsu>. We report how we determined our sample size, data exclusions, manipulations, and all measures included in this study, and followed JARS (Kazak, 2018). Data and analysis code are available upon request. Data were analyzed using R, version 4.1.0 (R Core Team, 2021). Original trial registration can be found at clinicaltrials.gov (*NCT # omitted for blinded submission*).

Measures

Baseline Symptom Measures

Response Style. The Children's Response Style Questionnaire (CRSQ; Abela et al., 2002) is a 25-item measure adapted from the Response Styles Questionnaire for adults (Nolen-Hoeksema & Morrow, 1991) used to assess response styles to sad feelings in children and adolescents. The CRSQ contains three subscales that capture three different response styles to distress: rumination, distraction, and problem-solving. The 13-item rumination subscale assesses

how often an individual responds to distress with rumination and includes items such as, “When I am sad or stressed, I think, ‘Why can’t I handle things better.’” The 7-item distraction subscale assesses how often an individual avoids feeling symptoms of distress through engaging in distracting tasks or behaviors (e.g., watching TV or playing video games). The 5-item problem-solving subscale assesses how often an individual responds to distress by active engagement in solution-focused thoughts or behaviors (e.g., “Think of a way to make my problem better.”). For all items, respondents are asked to rate how often they react to sad feelings in the way modeled by the question on a 4-point Likert scale (0 = *almost never* to 3 = *almost always*). For the purposes of the study, we modified directions from the CRSQ to ask participants to report on what they do when they feel sad *or stressed*, in line with current conceptualizations (Nolen-Hoeksema et al., 2008; Shaw et al., 2019). In our sample, the CRSQ demonstrated excellent reliability for rumination ($\alpha = .92$), but lower internal consistency for distraction ($\alpha = .53$), and problem-solving ($\alpha = .73$).

Trait Mindfulness. The Five Facet Mindfulness Questionnaire (FFMQ; Baer et al., 2006) is a 39-item measure assessing levels of trait mindfulness across five dimensions (i.e., observing, describing, acting with awareness, nonjudgment of inner experiences, and nonreactivity to inner experiences). The observing subscale (8 items) measures the extent to which an individual notices internal and external stimuli/experiences as they are happening (e.g., “I notice the smells and aromas of things.”). The describing subscale (8 items) assesses an individual’s ability to verbally describe feelings or experiences (e.g., “I am good at finding words to describe my feelings.”). The acting with awareness subscale (8 items) measures an individual’s tendency to give their undivided attention to a task or experience (e.g., “I find myself doing things without paying attention” (reverse scored)). The nonjudgement subscale (8

items) measures the extent to which an individual takes a nonevaluative stance toward one's thoughts and experiences (e.g., "I think some of my emotions are bad or inappropriate and I should not feel them" (reverse scored)). The nonreactivity subscale (7 items) measures the extent to which an individual experiences their thoughts and feelings without getting caught up in them (e.g., "Usually when I have distressing thoughts or images, I step back and am aware of the thought or image without getting taken over by it."). Items are rated on a 5-point Likert scale (1 = *never or very rarely true* to 5 = *very often or always true*). The FFMQ has demonstrated validity and adequate-to-good reliability in samples of adult and emerging adult meditators and nonmeditators (Baer et al., 2008). Past research suggests the FFMQ demonstrates good-to-excellent reliability in adolescent samples (see Hambour et al., 2018; Johnstone et al., 2020; Ramler et al., 2016). In our sample, reliability for the subscales were good to excellent: observing ($\alpha = .81$), describing ($\alpha = .80$), awareness ($\alpha = .84$), nonjudgment ($\alpha = .90$) and nonreactivity ($\alpha = .77$).

Emotion Regulation. The Emotion Regulation Questionnaire (ERQ; Gross & John, 2003) is a 10-item self-report measure of two emotion-regulation strategies: suppression and reappraisal. The suppression subscale comprises four items (e.g., "When I am feeling negative emotions, I make sure not to express them."). The reappraisal subscale comprises 6 items, e.g., "When I want to feel more positive emotions, I change what I'm thinking about." For all questions, respondents are asked to rate how often they engage in each emotion regulation strategy on a 7-point Likert scale ranging from 0 (*strongly disagree*) to 6 (*strongly agree*). Past research has demonstrated that the ERQ has high internal consistency, test-retest reliability, and good convergent and discriminant validity among emerging adults and adults (Gross & John, 2003; John & Gross, 2004), and shows good-to-excellent reliability in samples of children and

adolescents (e.g., Carthy et al., 2010; Hollenstein et al., 2012). In our sample, the reliability for the suppression ($\alpha = .70$) and reappraisal ($\alpha = .89$) subscales were adequate.

Depressive Symptoms. The Children's Depression Inventory (CDI; Kovacs, 1992) is a commonly-used 27-item measure that assesses the frequency and severity of depressive symptoms in the last two weeks in children and adolescents (e.g., Craighead et al., 1995; Klein, et al., 2005). Questions on the CDI ask respondents to select which statements best describe how they have been thinking and feeling over the past two weeks (e.g., "I am sad once in a while," "I am sad many times," and "I am sad all the time"). Items are scored on a scale from 0 to 2 and higher scores indicate greater frequency and clinical severity of depressive symptoms. In our study, the CDI demonstrated excellent reliability ($\alpha = .90$).

Anxiety Symptoms. The Multidimensional Anxiety Scale for Children (MASC; March et al., 1997) is 39-item index of anxiety in youth. Items on the MASC are rated on a scale from 0 (*never true of me*) to 3 (*often true about me*). The MASC includes four subscales: physical symptoms, harm avoidance, social anxiety, and separation anxiety. This study focused on total scores, which had excellent reliability ($\alpha = .90$).

Worry. The Penn State Worry Questionnaire - Child (PSWQ-C; Chorpita et al., 1997) is a 14-item measure of worry in children and adolescents adapted from the Penn State Worry Questionnaire for adults (Meyer et al., 1990). Items assess frequency, severity, and controllability of worry on a 4-point Likert scale (0 = *never true* to 3 = *always true*) (e.g., "I am always worrying about something"). Higher scores on the PSWQ-C indicate greater worry. Past research suggests that the PSWQ-C demonstrates excellent test-retest reliability and high convergent and discriminant validity among child and adolescent samples (Chorpita et al., 1997). In this study, the PSWQ-C demonstrated excellent reliability ($\alpha = .93$).

Outcome Measures: Ecological Momentary Assessment of State Rumination

To assess state rumination, or the extent to which an individual is engaging in ruminative thinking in the present moment, participants rated the extent to which they were ruminating “just before” seeing the following EMA questions: “How much were you focusing on your problems?” (Rum_{Prob}) and “How much were you focusing on your emotions?” (Rum_{Emo}) on a 0 (*not at all*) to 100 (*extremely*) scale. These items were adapted from Moberly and Watkins (2008) and have been used to assess state rumination in other studies with adolescents (Hilt et al., 2017; Hilt & Pollak, 2012; Webb et al., 2021). Consistent with prior work (Webb et al., 2021), we computed daily means of Rum_{Prob} and Rum_{Emo} representing the mean value of these two variables for each day, for each subject, across the 21-day EMA period. Rum_{Prob} and Rum_{Emo} were moderately correlated in our sample (mean within-person $r = 0.37$, $SD = 0.32$).

Analytic Strategy

Overall RCT Outcomes

Given the nested (hierarchical) data structure (i.e., repeated assessments nested within individuals), we ran multilevel models (separately for the Rum_{Prob} and Rum_{Emo} outcomes) to test group (mindfulness app vs. control) differences in the slope of state rumination change across intervention days (*day* centered to represent estimated post-intervention state rumination scores), specifying a random slope and intercept. Treatment group x day interactions were modeled to test between-group differences in rumination change over time, adjusting for age and gender. Analyses were conducted with the `lme4` (Bates et al., 2014) and `lmerTest` (Kuznetsova et al., 2017) packages in R, using an intent-to-treat approach (i.e., all subjects included).

Generating Predicted Outcomes

Predictor variables in this RCT included age, gender, response styles (CRSQ subscales:

trait rumination, distraction & problem-solving), worry (PSWQ total score), mindfulness skills (FFMQ subscales: observe, describe, awareness, nonjudgment, nonreactivity), emotion regulation (ERQ subscales: cognitive reappraisal and expressive suppression), depressive symptom severity (CDI total score), and anxiety symptom severity (MASC total score). All predictor variables were assessed at baseline (i.e., pre-intervention). For our outcome measures, we computed the slope of state rumination change over the course of the 3-week trial from subject-specific regressions of rumination scores on intervention day (day 1- 21). This was done separately for Rum_{Prob} and Rum_{Emo}, resulting in two outcome measures assessing changes in problem-focused rumination and emotion-focused rumination, respectively. There were no missing values among the predictors and outcomes. Prior to analyses, all continuous predictor variables were z-standardized ($M = 0$; $SD = 1$) and dichotomous variables were recoded (-0.5 ; 0.5). Five outlier ($\pm 3 SD$ from the mean) values for the outcome variables were winsorized by replacing those values with the closest non-outlier value.

Predicted outcomes (i.e., slope of change in state rumination) were estimated for each adolescent ($n = 152$) in the trial, for each condition (mindfulness vs. control) separately. Separate analyses were conducted for the two rumination outcome variables (i.e., predicting the slope of change in Rum_{Prob} and Rum_{Emo}), given prior findings indicating their differential patterns of prediction (Webb et al., 2021) and their modest correlation (mean within-person $r = 0.37$). Consistent with a previous mindfulness app trial (Webb et al., 2021), the set of predictor variables were submitted to elastic net regularization (ENR)(glmnet package, Friedman et al., 2010) implemented via 10-fold cross-validation (CV) to generate predictions of outcome (repeated 100 times to generate stable estimates). Tuning of ENR's alpha and lambda parameters was conducted via the resampling grid search from the CARET package (Kuhn, 2008), testing

each combination of alpha (from 0 to 1 by 0.05) and lambda (from 0 to 2 by 0.01).

First, predicted outcomes for the mindfulness app condition were generated from ENR models derived from a previous trial of the same mindfulness app in adolescents with elevated trait rumination ($n = 80$; Webb et al., 2021). Specifically, for each participant in the present sample ($n = 152$), expected slope of change for both Rum_{Prob} and Rum_{Emo} were generated based on the set of predictor variables (and their associated coefficients) retained in ENR models from a separate sample in a prior trial (i.e., external validation of a prior mindfulness app prognostic model; Webb et al., 2021). Second, given that the latter single-arm trial did not include the control condition from the current RCT, expected control condition outcomes were derived from models developed in the current sample. Specifically, for those participants who were randomly assigned to the mindfulness app condition, their expected control condition outcome was estimated from an ENR model developed in the control condition sample (i.e., participants randomly assigned to the control condition). Finally, for those participants randomly assigned to the control condition, their predicted control condition outcome was derived from 10-fold CV ENR (to protect against overfitting; Kuhn & Johnson 2013). Specifically, the control group was split into 10 equal sized folds, and predictions for each of the held out 1/10 of the sample were generated from an ENR model developed in the other 9/10 of the data (repeated 100 times to generate stable estimates).

Personalized Mindfulness App Recommendations

As a final product of the above models, every participant had a predicted outcome for both the mindfulness and control conditions. Consistent with prior similar studies (e.g., see Cohen et al., 2020; Schwartz et al., 2021; Webb et al., 2018), these two predicted outcomes were subtracted (i.e., predicted slope of change for mindfulness app *minus* control condition) to create

a Personalized Advantage Index (PAI) to classify participants into those predicted to experience greater reduction in state rumination to the mindfulness app vs. the mood monitoring control condition. A negative PAI score indicates that a given adolescent is predicted to experience greater reductions in state rumination in the mindfulness app relative to the control condition (and vice versa for positive PAI scores).

Evaluating Recommendations

Two-tailed Welch's *t*-tests (Delacre et al., 2017) compared observed outcomes for those participants randomly assigned to their algorithm-indicated treatment condition (i.e., mindfulness app for PAI scores < 0 and control condition for PAI scores > 0) vs. the contraindicated condition. Cohen's *d* values were calculated to estimate effect sizes for group differences. In addition, to test whether PAI scores (as a continuous variable) moderated treatment group differences in outcomes, we tested a treatment group x PAI score interaction with change in Rum_{Prob} and Rum_{Emo} as the outcome (and while controlling for baseline trait rumination scores).

Comparison Model

For comparison, the above analyses were re-rerun substituting the multivariable ENR models (i.e., including all predictor variables) with a simple linear regression including only baseline trait rumination as the sole predictor. A second simpler comparison model excluded the step involving the prediction of control condition outcome. Specifically, as described above, PAI scores were computed by subtracting the predicted mindfulness app outcome *minus* control outcome for every individual. In this second comparison model, we tested whether predictions of mindfulness app outcome moderated treatment group differences in outcome (i.e., tested via a treatment group x predicted mindfulness app outcome interaction). In other words, can a model simply based on predictions of mindfulness app outcome from a prior trial (including baseline

trait rumination as the sole predictor) predict the likelihood an adolescent will benefit from a mindfulness app vs. a mood monitoring app? Finally, we used the parameter estimates from a final model to demonstrate the translation of a predictive model to personalized recommendations for app-based mindfulness training. Analyses were preregistered (<https://tinyurl.com/x58phrsu>), except for the group x PAI score interaction and the above two comparison models, which were added subsequently (see Supplement).

Determination of Sample Size

Sample size was determined via power analysis for the primary RCT aim focused on testing between-group differences in reduction of rumination [*citation omitted for blinded submission*]. Pilot work revealed a large between-group difference on changes in trait rumination ($d = .88$). The current sample size ($n = 152$) would provide excellent power ($>.99$; assuming $\alpha = 0.5$) to detect the latter effect. With an $\alpha = .05$ and power = 0.80, a sample size of 152 could detect a small-to-medium ($d = .46$) effect size for between-group differences in rumination change. However, this manuscript reports secondary analyses focused on predictions of outcome based on pre-treatment adolescent characteristics. With an $\alpha = .05$ and power = 0.80, the included sample size per group could detect a medium ($r = .31$) correlation between individual predictors and outcome.

Results

Attrition and Adherence

The mean number of EMA surveys completed over the course of the trial was 51.1 ($SD = 16.3$; or an average of 2.4 surveys per day) for the mindfulness group and 52.7 ($SD = 14.8$; or an average of 2.5 surveys per day) for the control group. Adolescents in the mindfulness condition completed a mean of 33.7 ($SD = 11.2$, or an average of 1.6 per day) mindfulness exercises over

the 3-week trial, and 83% completed at least 21 mindfulness exercises (i.e., averaged ≥ 1 per day). Participants received 1-minute mindfulness exercises most frequently (90% vs. 9% for 5-min and 1% for 10-min mindfulness exercise options). Over two-thirds of the sample (67.6%) received at least one 5-min mindfulness exercise, and 21.1% received at least one 10-min exercise.

Overall Outcomes

Multilevel models revealed significant treatment group x day interactions for Rum_{Prob} ($b = 0.57, t(134.5) = 3.20, p = .002$) and Rum_{Emo} ($b = 0.40, t(140.5) = 2.07, p = .04$), indicating greater improvement in problem-focused and emotion-focused state rumination in the mindfulness app relative to the control condition (**Figure 1**). However, as displayed by the distribution of subject-specific slopes in **Figure 2**, there was variability across subjects in rumination change, with some adolescents experiencing reductions in rumination, whereas others experiencing limited or no improvement (or even worsening rumination). Given the substantial between-person variability in outcomes, we sought to predict *which* adolescents are expected to benefit from engaging in app-delivered mindfulness training (see below).

Predicted Outcomes

Table 2 reports baseline predictor variables retained in the ENR models predicting change in Rum_{Prob} and Rum_{Emo} in the mindfulness app condition. Predictors of better mindfulness app outcome included lower emotional suppression and distraction, as well as higher trait rumination, anxiety symptoms, and problem-solving orientation. **Table 3** presents the variables retained in the ENR models predicting mood monitoring control condition outcome. Predictors of better outcome in the mood monitoring condition included lower anxiety symptoms, depressive symptoms, trait rumination, and suppression, and higher levels of certain mindfulness

facets (nonjudgment and awareness of experience), among other variables (see **Table 3**).

Personalized Mindfulness App Recommendations

Emotion-Focused Rumination

For analyses predicting slope of change in Rum_{Emo} , the mean PAI score was -0.03 ($SD = 0.78$; range = $-2.04 - 1.76$), indicating that, on average, the mindfulness app condition was only slightly favored by the model. Fifty-one percent of the sample ($n = 78/152$) were predicted to have better emotion-focused rumination outcomes in the mindfulness app vs. control condition (i.e., PAI score < 0). In contrast, 49% of the sample ($n = 74/152$) were predicted to have a more favorable outcome in the control condition (i.e., PAI score > 0).

Problem-Focused Rumination

For analyses predicting slope of change in Rum_{Prob} , the mean PAI score was -0.27 ($SD = 0.75$; range = $-1.84 - 2.26$). On average, the mindfulness app condition was favored by the model. Sixty-one percent of adolescents ($n = 93/152$) were predicted to have relatively better problem-focused rumination outcomes in the mindfulness app compared with the control condition (i.e., PAI score < 0). Thirty-nine percent of the sample ($n = 59/152$) were predicted to have a more favorable outcome in the control condition (i.e., PAI score > 0).

Evaluating Recommendations

Emotion-Focused Rumination

Among the subset of adolescents predicted to have a better outcome in the mindfulness app (PAI score < 0), those randomly assigned to the mindfulness app ($M = -0.64$; $SD = 1.34$) had significantly greater reductions in Rum_{Emo} relative to those randomized to the control condition ($M = 0.12$; $SD = 1.01$; $d = 0.65$, $t(73.9) = -2.87$, $p < .01$; **Figure 3a**). In contrast, for those participants predicted to have a relatively more favorable outcome in the control condition, there

were no significant differences in outcome among those randomly assigned to the mindfulness condition ($M = -0.54$; $SD = 1.66$) versus the control condition ($M = -0.35$; $SD = 0.99$; $d = 0.14$, $t(45.2) = -0.58$, $p = 0.56$; **Figure 3a**). There was a nonsignificant trend for the treatment group x PAI score interaction in predicting Rum_{Emo} improvement ($t(147) = 1.96$; $p = 0.05$). Namely, decreasing PAI scores (i.e., reflecting relatively greater predicted benefit for the mindfulness app relative to the control condition) were associated with larger group differences in outcome, favoring the mindfulness app.

Problem-Focused Rumination

Similar to the above findings, among the subset of adolescents predicted to have a better outcome to the mindfulness app, those randomly assigned to the mindfulness app ($M = -0.81$; $SD = 1.14$) had significantly greater reductions in Rum_{Prob} relative to those randomized to the mood monitoring condition ($M = -0.21$; $SD = 0.92$; $d = 0.58$, $t(84.8) = -2.78$, $p < 0.01$; **Figure 3b**). In contrast, for those adolescents predicted to have a relatively more favorable outcome in the control condition, there were no significant differences in outcome between those who received the mindfulness app ($M = -0.43$; $SD = 1.56$) and those who were in the control condition ($M = -0.03$; $SD = 0.81$; $d = 0.31$, $t(37.6) = -1.19$, $p = 0.24$; **Figure 3b**). The group x PAI interaction was not significant, $t(147) = 0.40$; $p = 0.69$.

Comparison Models

A similar pattern of findings emerged when the above multivariable prediction models were re-run only including a single predictor variable (i.e., baseline CRSQ trait rumination). Specifically, for analyses predicting Rum_{Emo}, the subset of adolescents predicted to have a better outcome in the mindfulness app exhibited significantly greater reductions in state rumination if assigned to the mindfulness app ($M = -0.82$; $SD = 1.34$) relative to the control condition ($M =$

0.02; $SD = 1.08$; $d = 0.68$, $t(75.2) = -3.13$, $p < 0.01$; **Supplemental Figure 1a**). In contrast, among the subset of participants predicted to have a relatively more favorable outcome in the control condition, there were no significant differences in outcome ($d = 0.01$, $t(48) = -0.05$, $p = 0.96$). Similarly, for analyses predicting Rum_{Prob} , between-group differences in outcome only emerged for the subset of adolescents predicted to have a better outcome in the mindfulness app ($d = 0.77$, $t(80.2) = -3.80$, $p < .001$; for analyses of control indicated adolescents: $d = 0.11$, $t(44.3) = 0.43$, $p = 0.67$; **Supplemental Figure 1b**). Treatment group x PAI score interactions emerged for both rumination outcomes, indicating that PAI scores moderated treatment group differences in outcome (Rum_{Emo} : $t(148) = 2.35$; $p = 0.02$; Rum_{Prob} : $t(148) = 3.41$; $p < 0.001$).

As a final comparison model, the above interaction test was re-run substituting the PAI score (i.e., computed as the predicted slope of change for the mindfulness app *minus* control condition) with the predicted slope of change for the mindfulness app. In other words, do predicted mindfulness app outcome scores from a model developed in a sample from a prior trial (Webb et al., 2021) moderate treatment group differences in outcome? Significant treatment group x predicted mindfulness outcome interactions emerged for both rumination outcomes, indicating that predicted mindfulness outcomes moderated differences in observed outcomes between intervention conditions (for Rum_{Emo} : $t(148) = 2.57$; $p = 0.01$; for Rum_{Prob} : $t(148) = 3.57$; $p < 0.001$). Better predicted mindfulness outcomes were associated with larger group differences in observed outcome, favoring the mindfulness group. The correlation between predicted mindfulness app Rum_{Emo} outcomes and observed outcomes was $r = .22$ ($p = 0.06$) for the mindfulness condition and $r = -.19$ ($p = 0.08$) for the control condition. For Rum_{Prob} , these values were $r = .38$ ($p < 0.01$) and $r = -.17$ ($p = 0.12$), respectively.

Demonstration of Personalized Recommendations

To demonstrate how such a predictive model can be translated to personalized mindfulness app recommendations, we used the above final comparison model as an example. Specifically, the parameter estimates from the above regression models were used to estimate predicted change in state rumination for the mindfulness app vs. mood monitoring control for a new adolescent based on their pre-intervention trait rumination score. First, as displayed in **Figure 4A** (R_{Emo}) and **4B** (R_{Prob}), we plotted the relationship between predicted mindfulness app outcome and observed outcome for the mindfulness app (green line) and the mood monitoring condition (red line). The dashed vertical grey line represents the point at which these two regression lines intersect. An individual with a predicted mindfulness app outcome score to the left of this line is predicted to have a better outcome to the mindfulness app relative to the mood monitoring app (and vice versa for adolescents with predicted scores to right of this line). The area to the left of this line is colored yellow to reflect a “cautious recommendation” for app-based mindfulness training. Second, we computed a 95% confidence interval for this intersection point via bootstrap resampling (Boot package; Canty & Ripley, 2021). Specifically, 1,000 samples (with replacement) were drawn, and we recomputed the two regression lines and their intersection point for each of these samples. The dashed vertical red line represents the left margin of the 95% confidence interval for this intersection point. In other words, if an individual’s predicted mindfulness outcome falls to the left of this line our confidence in the predicted benefit of app-based mindfulness training relative to mood monitoring increases. Third, we also implemented the Johnson-Neyman technique (Hayes & Matthes, 2009) to probe the group \times predicted mindfulness outcome interaction and to estimate the value of the moderator (predicted mindfulness outcome score) at which group differences in outcome become

statistically significant (for $Rum_{Emo} < -0.22$; for $Rum_{Prob} < -0.27$; see solid vertical grey line). If an adolescent's predicted mindfulness app outcome falls to the left of *both* the Johnson-Neyman threshold (solid grey line) and the 95% confidence interval (dotted red line) the plot area is colored green to reflect a more confident recommendation to use the mindfulness app.

To illustrate with a concrete example, and using the models focused on the prediction of change in Rum_{Prob} (**Figure 4b**), an adolescent with a trait rumination score half a standard deviation above the mean (i.e., 21) would have a predicted mindfulness outcome score of -0.47 (within the “green zone” of **Figure 4b**), and a predicted reduction in Rum_{Prob} of 16.5 points over 3 weeks of using the mindfulness app vs. 0.74 points for the mood monitoring app. Assuming a baseline Rum_{Prob} score half a standard deviation above the mean (i.e., 55.6), this would translate to a 29.7% ($16.5/55.6$) expected reduction in problem-focused rumination if using the mindfulness app vs. 1.3% ($0.74/55.6$) for the mood monitoring app. In sum, based on a brief assessment of trait rumination, our algorithm can provide individual adolescents with useful information regarding their expected benefit *prior* to deciding to enroll in a multi-week course of app-based mindfulness training.

Discussion

An initial single-arm trial demonstrated that adolescents with elevated trait rumination who used a mindfulness app for 3-weeks exhibited significant reductions in their levels of rumination (Hilt & Swords, 2021). The present RCT extends these findings by showing that on average, adolescents randomly assigned to a mindfulness app report greater reductions in rumination relative to a mood monitoring control group. As displayed in **Figure 2**, there was substantial between-individual variability in outcomes to the mindfulness app and mood monitoring control condition. Given these individual differences in outcome, coupled with the

increasing number of adolescents using mindfulness apps as a means of coping with emotional distress (Rideout & Fox, 2018; Lagan et al., 2021; Wasil et al., 2020), we tested whether baseline characteristics could inform which adolescents are most likely to benefit from these apps.

Predictive models identified a subset of adolescents, representing 51%-61% of the sample, who exhibited significantly greater reductions in rumination relative to the mood monitoring control ($d = .65$ for Rum_{Emo}; $d = .58$ for Rum_{Prob}). In contrast, between-condition differences in outcome among adolescents predicted to have better outcomes in the mood monitoring control condition were not significant ($d = .14$ for Rum_{Emo}; $d = .31$ for Rum_{Prob}). A comparison model yielded similar findings, suggesting that a multivariable machine learning (ENR) model added little predictive information relative to a linear regression simply incorporating baseline trait rumination scores. Results revealed that adolescents with a heightened tendency to ruminate experienced greater benefits from the mindfulness app relative to the mood monitoring condition. Critically, as an initial test of external validity, predicted mindfulness app outcomes for participants in the present trial were generated from a regression model developed in a prior sample (Webb et al., 2021). As illustrated in **Figure 4**, predicted mindfulness app outcomes were positively associated with observed outcomes for adolescents randomly assigned to the mindfulness app ($r = .22$ and $r = .38$ for Rum_{Emo} and Rum_{Prob}, respectively), but *inversely* associated with outcomes for participants who received the control condition ($r = -.19$ and $r = -.17$ for Rum_{Emo} and Rum_{Prob}, respectively). In summary, results indicated that a brief self-report measure of trait rumination (CRSQ rumination subscale) can inform which adolescents are most likely to benefit from app-based mindfulness training. Future research is needed to test whether predictive models can be developed to identify which adolescents are most likely to benefit from a mindfulness app vs. other popular mental health

apps (e.g., focused on psychoeducation, journaling, or cognitive behavioral skills; Lagan et al., 2021). The extent to which trait rumination predicts differential response to a mindfulness app vs. these other common mental health apps is an open question. In addition, future studies could directly compare mindfulness interventions that differ in length and content, given the many varieties of meditation practices that exist (e.g., focused attention, open monitoring, or loving-kindness meditations). It may be that some adolescents are more receptive and responsive to certain mindfulness practices, which could be predicted based on their individual characteristics.

Finally, we demonstrated how our predictive model can be used to generate personalized mindfulness app recommendations for an individual adolescent. First, the model provides a binary prediction of whether or not a given adolescent is expected to have a greater reduction in rumination from the mindfulness app or mood monitoring app (i.e., based on whether the predicted outcome score falls to the left or right of the intersection point in **Figure 4**). Second, the model generates an estimate of the expected magnitude of improvement (or deterioration) in state rumination for each app condition. Third, the model distinguishes between strengths of recommendation, demarcated by the green (confident recommendation) vs. yellow (cautious recommendation) sections of the figure (based on a bootstrapped confidence interval and Johnson-Neyman interval). Collectively, this information could be used to objectively communicate expected mindfulness app outcomes to a given adolescent prior to them engaging in app-based mindfulness training.

There are several notable strengths to the present study. First, we compared reductions in rumination from a mindfulness app vs. a mood monitoring control condition. This allowed us to test whether adolescent baseline characteristics were merely prognostic predictors of reductions in state rumination or whether they specifically predicted benefit from the mindfulness app.

Second, both the mindfulness and mood monitoring conditions were delivered via smartphone apps, and thus represent highly scalable interventions. Third, our outcome (change in state rumination) was assessed via repeated, daily EMA, which increased ecological validity and minimized recall bias relative to conventional, self-report questionnaires of trait rumination, which rely on accurate retrospective recall of one's tendency to ruminate. Fourth, we compared our multivariable machine learning (ENR) models with a simple linear regression including a single predictor. Fifth, as an initial test of external validity, our predictive model of mindfulness app outcome was developed in a sample of adolescents from a prior trial (Webb et al., 2021). Finally, we distinguished between rumination focused on problems vs. emotions, given prior findings indicating their differential patterns of prediction (Webb et al., 2021) and their modest correlation (mean within-person $r = 0.37$ in the present sample).

Several limitations should be noted. First, sample size was relatively small for such prognostic models (Luedtke et al., 2019; Riley et al., 2020). Second, although the inclusion of a mood monitoring control group was a strength relative to prior research, it will be important for future studies to compare a mindfulness app to alternative popular mental health apps (e.g., CBT-based apps) and to compare different variants of mindfulness training (e.g., varying lengths and content of meditation exercises). Third, we relied on a limited subset of potentially relevant predictors. For example, measures of attentional control (e.g., assessed via behavioral tasks; Wong et al., 2018; Lim et al., 2018), the propensity to mind-wander (Rahl et al., 2017; Webb et al., 2021; Xu et al., 2017), or personality characteristics (Karl et al., 2021; Krick & Felfe, 2020; Nyklíček & Irmischer, 2017; Tang & Braver, 2020) could inform the likelihood of benefiting from a mindfulness app above and beyond the contribution of the clinical and demographic variables considered in the present study. Fourth, higher scores on the emotion-focused

rumination EMA item do not necessarily reflect problematic rumination (e.g., individuals undergoing mindfulness training may focus on negative emotions in an adaptive, non-judgmental manner). This may help account for some of the surprising findings related to the analyses of emotion-focused rumination (e.g., only 51% predicted to have better outcomes to the mindfulness app vs. control condition; vs. 61% for problem-focused rumination). Future research is needed with a more nuanced assessment of emotion-focused rumination (e.g., differentiating between adaptive awareness of negative emotions vs. problematic dwelling on these emotions) These limitations notwithstanding, findings from the present study support the efficacy of a mindfulness app to reduce rumination in adolescents and highlight the potential of data-driven approaches to inform which adolescents may benefit most from app-based mindfulness training. Ultimately, a prospective test of this algorithm is needed to compare adolescents assigned to their algorithm-indicated intervention versus random assignment (or group assignment based on participant preference, uninformed by the algorithm).

References:

- Abela, J. R. Z., Brozina, K., & Haigh, E. P. (2002). An Examination of the Response Styles Theory of Depression in Third- and Seventh-Grade Children: A Short-Term Longitudinal Study. *Journal of Abnormal Child Psychology*, *30*(5), 515–527.
<https://doi.org/10.1023/A:1019873015594>
- Ames, C. S., Richardson, J., Payne, S., Smith, P., & Leigh, E. (2014). Mindfulness-based cognitive therapy for depression in adolescents. *Child and Adolescent Mental Health*, *19*(1), 74–78.
- Avenevoli, S., Swendsen, J., He, J.-P., Burstein, M., & Merikangas, K. R. (2015). Major Depression in the National Comorbidity Survey–Adolescent Supplement: Prevalence, Correlates, and Treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(1), 37–44.e2. <https://doi.org/10.1016/j.jaac.2014.10.010>
- Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using Self-Report Assessment Methods to Explore Facets of Mindfulness. *Assessment*, *13*(1), 27–45.
<https://doi.org/10.1177/1073191105283504>
- Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., Walsh, E., Duggan, D., & Williams, J. M. G. (2008). Construct Validity of the Five Facet Mindfulness Questionnaire in Meditating and Nonmeditating Samples. *Assessment*, *15*(3), 329–342.
<https://doi.org/10.1177/1073191107313003>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *ArXiv Preprint ArXiv:1406.5823*.
- Canty, A., & Ripley, B. (2021). *boot: Bootstrap R (S-Plus) Functions*.
- Carthy, T., Horesh, N., Apter, A., Edge, M. D., & Gross, J. J. (2010). Emotional reactivity and

- cognitive regulation in anxious children. *Behaviour Research and Therapy*, 48(5), 384–393.
- Chorpita, B. F., Tracey, S. A., Brown, T. A., Collica, T. J., & Barlow, D. H. (1997). Assessment of worry in children and adolescents: An adaptation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy*, 35(6), 569–581.
[https://doi.org/10.1016/S0005-7967\(96\)00116-7](https://doi.org/10.1016/S0005-7967(96)00116-7)
- Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment Selection in Depression. *Annual Review of Clinical Psychology*, 14(1), 209–236. <https://doi.org/10.1146/annurev-clinpsy-050817-084746>
- De Haan, A. M., Boon, A. E., de Jong, J. T., Hoeve, M., & Vermeiren, R. R. (2013). A meta-analytic review on treatment dropout in child and adolescent outpatient mental health care. *Clinical Psychology Review*, 33(5), 698–711.
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch’s t-test instead of Student’s t-test. *International Review of Social Psychology*, 30(1).
- F. Wong, K., A. A. Massar, S., Chee, M. W. L., & Lim, J. (2018). Towards an Objective Measure of Mindfulness: Replicating and Extending the Features of the Breath-Counting Task. *Mindfulness*, 9(5), 1402–1410. <https://doi.org/10.1007/s12671-017-0880-1>
- Geiger, A. W., & Davis, L. (2019). A growing number of American teenagers – particularly girls – are facing depression. *Pew Research Center*. <https://www.pewresearch.org/fact-tank/2019/07/12/a-growing-number-of-american-teenagers-particularly-girls-are-facing-depression/>
- Goodman, M. S., Madni, L. A., & Semple, R. J. (2017). Measuring mindfulness in youth: Review of current assessments, challenges, and future directions. *Mindfulness*, 8(6),

1409–1420.

- Grist, R., Porter, J., & Stallard, P. (2017). Mental health mobile apps for preadolescents and adolescents: A systematic review. *Journal of Medical Internet Research, 19*(5), e176.
- Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*(2), 348–362. <https://doi.org/10.1037/0022-3514.85.2.348>
- Hambour, V. K., Zimmer-Gembeck, M. J., Clear, S., Rowe, S., & Avdagic, E. (2018). Emotion regulation and mindfulness in adolescents: Conceptual and empirical connection and associations with social anxiety symptoms. *Personality and Individual Differences, 134*, 7–12.
- Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods, 41*(3), 924–936.
- Hilt, L. M., & Pollak, S. D. (2012). Getting out of rumination: Comparison of three brief interventions in a sample of youth. *Journal of Abnormal Child Psychology, 40*(7), 1157–1165. <https://doi.org/10.1007/s10802-012-9638-3>
- Hilt, L. M., Sladek, M. R., Doane, L. D., & Stroud, C. B. (2017). Daily and trait rumination: Diurnal cortisol patterns in adolescent girls. *Cognition and Emotion, 31*(8), 1757–1767.
- Hilt, L. M., & Swords, C. M. (2021). Acceptability and Preliminary Effects of a Mindfulness Mobile Application for Ruminative Adolescents. *Behavior Therapy*. <https://doi.org/10.1016/j.beth.2021.03.004>
- Hollenstein, T., McNeely, A., Eastabrook, J., Mackey, A., & Flynn, J. (2012). Sympathetic and parasympathetic responses to social stress across adolescence. *Developmental*

- Psychobiology*, 54(2), 207–214.
- John, O. P., & Gross, J. J. (2004). Healthy and unhealthy emotion regulation: Personality processes, individual differences, and life span development. *Journal of Personality*, 72(6), 1301–1334.
- Johnstone, J. M., Ribbers, A., Jenkins, D., Atchley, R., Gustafsson, H., Nigg, J. T., Wahbeh, H., & Oken, B. (2020). Classroom-Based Mindfulness Training Reduces Anxiety in Adolescents: Acceptability and Effectiveness of a Cluster-Randomized Pilot Study. *Journal of Restorative Medicine*, 10(1).
- Karl, J. A., Fischer, R., & Jose, P. E. (2021). The development of mindfulness in young adults: The relationship of personality, reinforcement sensitivity, and mindfulness. *Mindfulness*, 12(5), 1103–1114.
- Kovacs, M. (1992). *Children's Depression Inventory Manual*. Multi-Health Systems.
- Krick, A., & Felfe, J. (2020). Who benefits from mindfulness? The moderating role of personality and social norms for the effectiveness on psychological and physiological outcomes among police officers. *Journal of Occupational Health Psychology*, 25(2), 99–112. <https://doi.org/10.1037/ocp0000159>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13).
- Lagan, S., D'Mello, R., Vaidyam, A., Bilden, R., & Torous, J. (2021). Assessing Mental Health Apps Marketplaces with Objective Metrics from 29,190 Data Points from 278 Apps. *Acta Psychiatrica Scandinavica*.
- Lim, J., Teng, J., Patanaik, A., Tandi, J., & Massar, S. A. A. (2018). Dynamic functional connectivity markers of objective trait mindfulness. *NeuroImage*, 176, 193–202.

<https://doi.org/10.1016/j.neuroimage.2018.04.056>

Luedtke, A., Sadikova, E., & Kessler, R. C. (2019). Sample Size Requirements for Multivariate Models to Predict Between-Patient Differences in Best Treatments of Major Depressive Disorder. *Clinical Psychological Science*, 7(3), 445–461.

<https://doi.org/10.1177/2167702618815466>

Mani, M., Kavanagh, D. J., Hides, L., & Stoyanov, S. R. (2015). Review and Evaluation of Mindfulness-Based iPhone Apps. *JMIR MHealth and UHealth*, 3(3).

March, J. S., Parker, J. D., Sullivan, K., Stallings, P., & Conners, C. K. (1997). The Multidimensional Anxiety Scale for Children (MASC): Factor structure, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(4), 554–565. <https://doi.org/10.1097/00004583-199704000-00019>

Mendelson, T., Greenberg, M. T., Dariotis, J. K., Gould, L. F., Rhoades, B. L., & Leaf, P. J. (2010). Feasibility and preliminary outcomes of a school-based mindfulness intervention for urban youth. *Journal of Abnormal Child Psychology*, 38(7), 985–994.

Merikangas, K. R., He, J., Burstein, M., Swanson, S. A., Avenevoli, S., Cui, L., Benjet, C., Georgiades, K., & Swendsen, J. (2010). Lifetime Prevalence of Mental Disorders in US Adolescents: Results from the National Comorbidity Study-Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolescent Psychiatry*, 49(10), 980–989. <https://doi.org/10.1016/j.jaac.2010.05.017>

Moberly, N. J., & Watkins, E. R. (2008). Ruminative self-focus and negative affect: An experience sampling study. *Journal of Abnormal Psychology*, 117(2), 314.

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta Earthquake. *Journal of*

- Personality and Social Psychology*, 61(1), 115–121.
- Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking Rumination. *Perspectives on Psychological Science*, 3(5), 400–424.
- Nyklíček, I., & Irrmischer, M. (2017). For whom does mindfulness-based stress reduction work? Moderating effects of personality. *Mindfulness*, 8(4), 1106–1116.
- Rahl, H. A., Lindsay, E. K., Pacilio, L. E., Brown, K. W., & Creswell, J. D. (2017). Brief mindfulness meditation training reduces mind wandering: The critical role of acceptance. *Emotion*, 17(2), 224–230. <https://doi.org/10.1037/emo0000250>
- Ramler, T. R., Tennison, L. R., Lynch, J., & Murphy, P. (2016). Mindfulness and the College Transition: The Efficacy of an Adapted Mindfulness-Based Stress Reduction Intervention in Fostering Adjustment among First-Year Students. *Mindfulness*, 7(1), 179–188. <https://doi.org/10.1007/s12671-015-0398-3>
- Rideout, V., & Fox, S. (2018). *Digital Health Practices among Teens and Young Adults—Hopelab*. <https://www.hopelab.org/report/a-national-survey-by-hopelab-and-well-being-trust-2018/digital-health-practices/digital-health-practices-among-teens-and-young-adults/>
- Rideout, V., & Robb, M. B. (2019). *The Common Sense Census: Media Use by Tweens and Teens, 2019 | Common Sense Media*. <https://www.common sense media.org/research/the-common-sense-census-media-use-by-tweens-and-teens-2019>
- Riley, R. D., Ensor, J., Snell, K. I., Harrell, F. E., Martin, G. P., Reitsma, J. B., Moons, K. G., Collins, G., & van Smeden, M. (2020). Calculating the sample size required for developing a clinical prediction model. *Bmj*, 368.
- Shaw, Z. A., Hilt, L. M., & Starr, L. R. (2019). The developmental origins of ruminative

- response style: An integrative review. *Clinical Psychology Review*, 74, 101780.
<https://doi.org/10.1016/j.cpr.2019.101780>
- Tang, R., & Braver, T. S. (2020). Towards an Individual Differences Perspective in Mindfulness Training Research: Theoretical and Empirical Considerations. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.00818>
- Torous, J., Firth, J., Huckvale, K., Larsen, M. E., Cosco, T. D., Carney, R., Chan, S., Pratap, A., Yellowlees, P., Wykes, T., Keshavan, M., & Christensen, H. (2018). The Emerging Imperative for a Consensus Approach Toward the Rating and Clinical Recommendation of Mental Health Apps. *The Journal of Nervous and Mental Disease*, 206(8), 662–666.
- Twenge, J. M., Cooper, A. B., Joiner, T. E., Duffy, M. E., & Binau, S. G. (2019). Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005–2017. *Journal of Abnormal Psychology*, 128(3), 185–199. <https://doi.org/10.1037/abn0000410>
- Van Dam, N. T., van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., et al. (2018). Mind the Hype: A Critical Evaluation and Prescriptive Agenda for Research on Mindfulness and Meditation. *Perspectives on Psychological Science*, 13(1), 36–61.
- van de Weijer-Bergsma, E., Langenberg, G., Brandsma, R., Oort, F. J., & Bögels, S. M. (2014). The effectiveness of a school-based mindfulness training as a program to prevent stress in elementary school children. *Mindfulness*, 5(3), 238–248.
- Villa, C. D., & Hilt, L. M. (2014). Brief instruction in mindfulness and relaxation reduce rumination differently for men and women. *International Journal of Cognitive Therapy*, 7(4), 320–333. https://doi.org/10.1521/ijct_2014_07_02
- Wasil, A. R., Gillespie, S., Shingleton, R., Wilks, C. R., & Weisz, J. R. (2020). Examining the

- Reach of Smartphone Apps for Depression and Anxiety. *American Journal of Psychiatry*, *177*(5), 464–465. <https://doi.org/10.1176/appi.ajp.2019.19090905>
- Webb, C. A., Israel, E. S., Belleau, E., Appleman, L., Forbes, E. E., & Pizzagalli, D. A. (2021). Mind-Wandering in Adolescents Predicts Worse Affect and Is Linked to Aberrant Default Mode Network–Salience Network Connectivity. *Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(3), 377–387. <https://doi.org/10.1016/j.jaac.2020.03.010>
- Webb, C. A., Swords, C. M., Murray, L. M., & Hilt, L. M. (2021). *App-based Mindfulness Training for Adolescent Rumination: Predictors of Immediate and Cumulative Benefit*. PsyArXiv. <https://doi.org/10.31234/osf.io/k4ygw>
- Webb, C. A., Trivedi, M. H., Cohen, Z. D., Dillon, D. G., Fournier, J. C., Goer, F., Fava, M., McGrath, P. J., Weissman, M., Parsey, R., Adams, P., Trombello, J. M., Cooper, C., Deldin, P., Oquendo, M. A., McInnis, M. G., Huys, Q., Bruder, G., Kurian, B. T., ... Pizzagalli, D. A. (2018). Personalized prediction of antidepressant v. placebo response: Evidence from the EMBARC study. *Psychological Medicine*, 1–10. <https://doi.org/10.1017/S0033291718001708>
- Whitney, D. G., & Peterson, M. D. (2019). US national and state-level prevalence of mental health disorders and disparities of mental health care use in children. *JAMA Pediatrics*, *173*(4), 389–391.
- Xu, M., Purdon, C., Seli, P., & Smilek, D. (2017). Mindfulness and mind wandering: The protective effects of brief meditation in anxious individuals. *Consciousness and Cognition*, *51*, 157–165. <https://doi.org/10.1016/j.concog.2017.03.009>

Table 1.*Demographic and Clinical Characteristics for the Sample (N = 152).*

		Control (n = 80)				Mindfulness (n = 72)			
		n	%	Mean	SD	n	%	Mean	SD
Age				13.66	0.86			13.78	0.93
Sex									
	<i>Female</i>	45	56.25%			44	61.11%		
	<i>Male</i>	35	43.75%			28	38.89%		
Race									
	<i>White</i>	70	87.50%			55	76.39%		
	<i>Black; African American</i>	2	2.50 %			3	4.17%		
	<i>Asian; Asian American</i>	2	2.50%			1	1.39%		
	<i>Native Hawaiian or Pacific Islander</i>	0	0.00%			1	1.39%		
	<i>Native American</i>	0	0.00%			0	0.00%		
	<i>Other</i>	1	1.25%			1	1.39%		
	<i>Multi-racial</i>	5	6.25%			11	15.28%		
Ethnicity									
	<i>Non-Hispanic</i>	73	91.25%			63	87.50%		
	<i>Hispanic</i>	7	8.75%			9	12.50%		
Clinical Characteristics									
	<i>ERQ Suppression</i>			15.49	4.44			16.14	5.26
	<i>ERQ Reappraisal</i>			25.34	7.28			25.68	8.51
	<i>CRSQ Rumination</i>			15.68	8.16			17.81	9.73
	<i>CRSQ Distraction</i>			9.31	3.59			9.43	3.13
	<i>CRSQ Problem Solving</i>			5.74	3.28			6.24	3.10
	<i>FFMQ Observe</i>			21.38	6.20			22.28	6.23
	<i>FFMQ Describe</i>			24.69	5.75			24.26	5.75
	<i>FFMQ Awareness</i>			26.17	5.84			25.03	6.23

<i>FFMQ Nonjudgement</i>			27.49	7.58		25.56	7.46
<i>FFMQ Nonreactivity</i>			19.44	3.99		19.58	4.58
<i>CDI Total</i>			12.04	7.72		13.40	8.77
<i>MASC Total</i>			50.84	17.35		53.47	16.51
<i>PSWQ Total</i>			21.53	9.41		23.86	9.26
<i>History of NSSI</i>	11	13.75%			11	15.28%	

Table 2.*Baseline variables retained in elastic net models predicting mindfulness app outcome*

	Emotion-Focused Rumination	Problem-Focused Rumination
<i>Predictors</i>	Coefficient	Coefficient
<i>Age</i>		
<i>Gender</i>		
<i>CDI</i>		
<i>MASC Total</i>	0.12	
<i>ERQ Suppression</i>	-0.22	-0.46
<i>ERQ Reappraisal</i>	0.30	
<i>CRSQ Rumination</i>	0.21	
<i>CRSQ Distraction</i>	-0.48	-0.27
<i>CRSQ Problem Solving</i>	0.25	0.26
<i>FFMQ Observe</i>		
<i>FFMQ Describe</i>	-0.34	
<i>FFMQ Nonjudgement</i>		-0.36
<i>FFMQ Nonreactivity</i>		0.26
<i>FFMQ Awareness</i>		

Note. Derived from sample in prior trial (Webb et al., 2021). ERQ = Emotion Regulation Questionnaire; CRSQ = Children's Response Styles Questionnaire; FFMQ = Five Facet Mindfulness Questionnaire; MASC = Multidimensional Anxiety Scale for Children. Positive parameter estimates indicate that higher scores on the predictor variable are associated with better outcome (i.e., reduction in state rumination).

Table 3*Baseline variables retained in elastic net models predicting control condition outcome*

	Emotion-Focused Rumination	Problem-Focused Rumination
<i>Predictors</i>	<i>Coefficient</i>	<i>Coefficient</i>
<i>Age</i>		-0.03
<i>Gender</i>		0.06
<i>CDI</i>	-0.008	-0.04
<i>MASC</i>	-0.10	-0.01
<i>ERQ Suppression</i>		-0.04
<i>ERQ Reappraisal</i>		0.007
<i>CRSQ Rumination</i>		-0.02
<i>CRSQ Distraction</i>		0.05
<i>CRSQ Problem Solving</i>		0.04
<i>FFMQ Observe</i>		0.003
<i>FFMQ Describe</i>		-0.005
<i>FFMQ Nonjudgement</i>	0.08	0.003
<i>FFMQ Nonreactivity</i>		-0.006
<i>FFMQ Awareness</i>	0.10	0.04
<i>PSWQ-C</i>		-0.01

Note. ERQ = Emotion Regulation Questionnaire; CRSQ = Children's Response Styles Questionnaire; FFMQ = Five Facet Mindfulness Questionnaire; CDI = Children's Depression Inventory; MASC = Multidimensional Anxiety Scale for Children; PSWQ = Penn State Worry Questionnaire child version. Emotion-focused rumination model: $\alpha = 0.25$, $\lambda = 0.48$; Problem-focused rumination model: $\alpha = 0$, $\lambda = 2$. Positive parameter estimates indicate that higher scores on the predictor variable are associated with better outcome (i.e., reduction in state rumination).

Figure Captions

Figure 1. Condition x Day interaction for problem-focused rumination (top panel) and emotion-focused rumination (bottom panel).

Figure 2. Plotting distribution of change (slopes) in problem-focused rumination (top panel) and emotion-focused rumination (bottom panel) for each condition.

Figure 3. Comparison of change in rumination for participants randomized to the mindfulness app vs. control (leftmost bars labeled “RCT Results”). Comparison of change in state rumination for participants randomly assigned to their PAI-indicated condition vs. those assigned to their PAI-contraindicated condition (middle and rightmost bars). For this bar plot, outcomes (slopes) were multiplied by -1 so that positive scores reflect rumination reduction. Error bars represent standard error. * $p < .05$, ** $p < .01$

Figure 4. Plot of predicted mindfulness app outcome for each condition to inform personalized recommendations. Dashed vertical grey line represents the point at which the two regression lines intersect (with dashed vertical red line representing the left margin of a bootstrapped 95% confidence interval). Solid vertical grey line is derived from the Johnson-Neyman technique and represents value of the moderator (predicted mindfulness outcome score) at which group differences in outcome become statistically significant. See main text for description of use for personalized mindfulness app recommendation with an example.

Figure 1

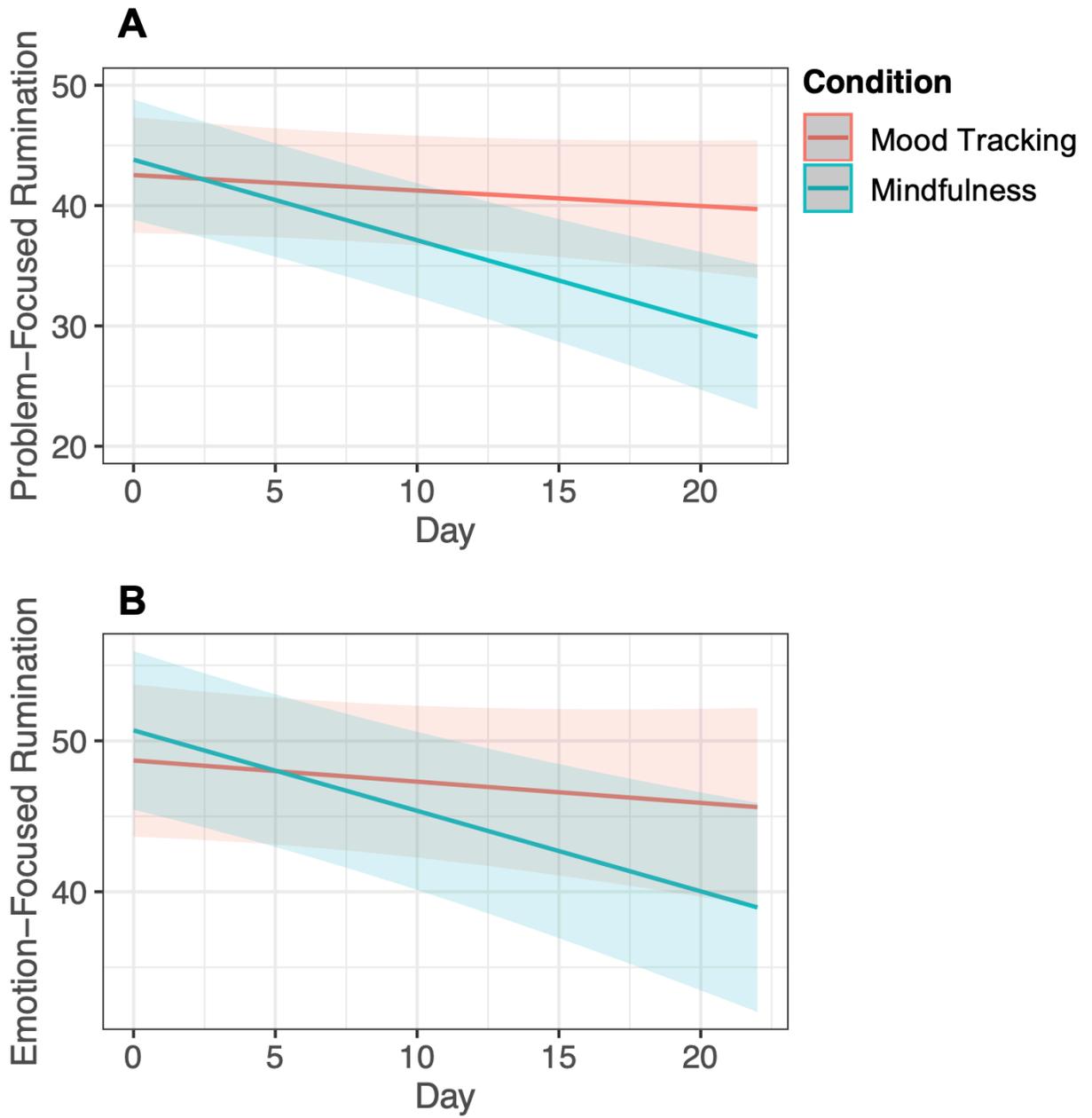


Figure 2

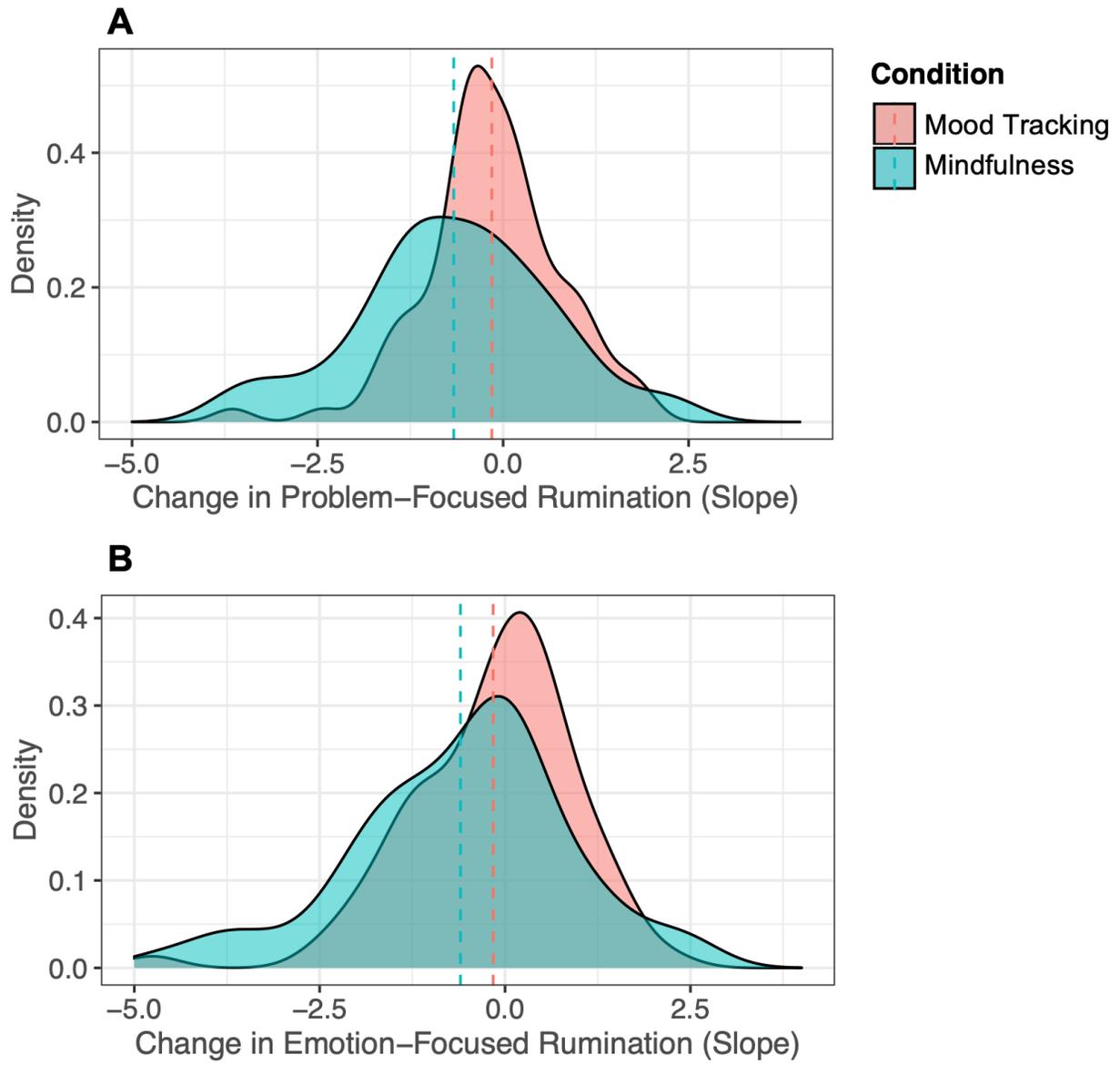
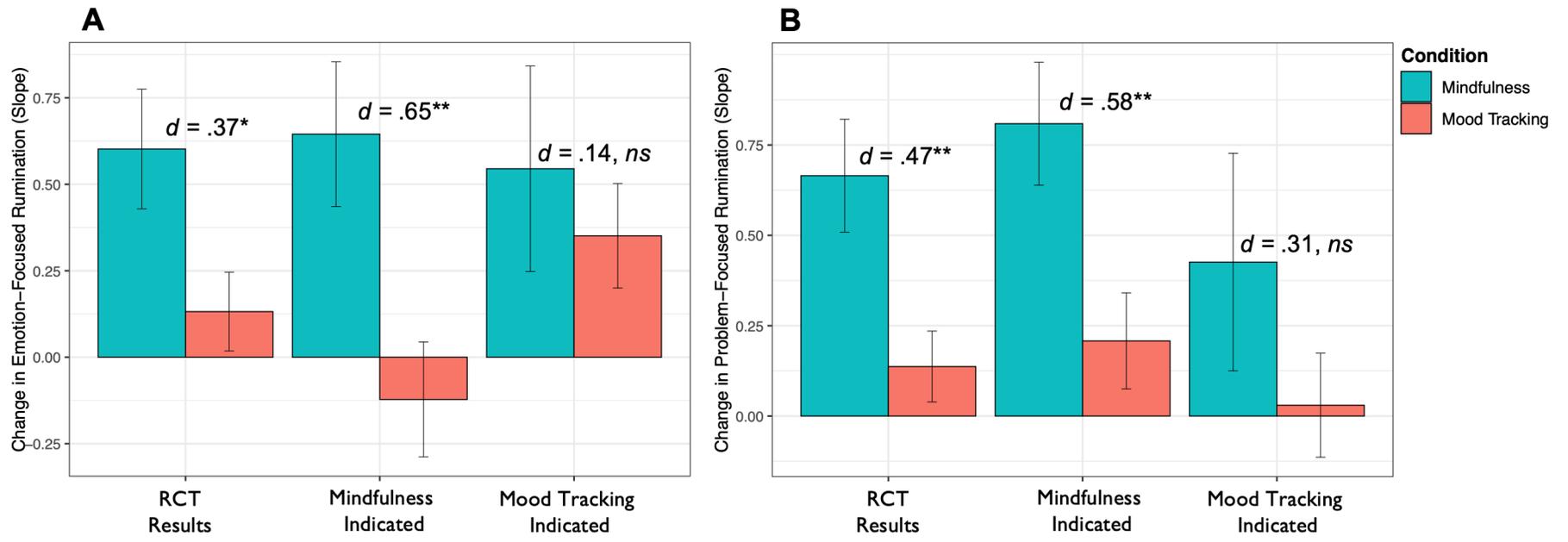


Figure 3



* $p < .05$, ** $p < .01$

Figure 4

