Testing the Efficacy of a Multicomponent, Self-Guided, Smartphone-Based Meditation App: Three-Armed Randomized Controlled Trial

Simon B Goldberg1,2, PhD; Theodore Imhoff-Smith1, BSc; Daniel M Bolt3, PhD; Christine D Wilson-Mendenhall1, PhD; Cortland J Dahl1,4, PhD; Richard J Davidson1,5,6, PhD; Melissa A Rosenkranz1,6, PhD

1Center for Healthy Minds, University of Wisconsin, Madison, WI, United States
2Department of Counseling Psychology, University of Wisconsin-Madison, Madison, WI, United States
3Department of Educational Psychology, University of Wisconsin-Madison, Madison, WI, United States
4Healthy Minds Innovations Inc, Madison, WI, United States
5Department of Psychology, University of Wisconsin-Madison, Madison, WI, United States
6Department of Psychiatry, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Simon B Goldberg, PhD
Center for Healthy Minds
University of Wisconsin
625 W Washington Ave
Madison, WI, 53703
United States
Phone: 1 6082658986
Fax: 1 6082654174
Email: sbgoldberg@wisc.edu

Abstract

Background: A growing number of randomized controlled trials (RCTs) suggest psychological benefits associated with meditation training delivered via mobile health. However, research in this area has primarily focused on mindfulness, only one of many meditative techniques.

Objective: This study aims to evaluate the efficacy of 2 versions of a self-guided, smartphone-based meditation app—the Healthy Minds Program (HMP)—which includes training in mindfulness (Awareness), along with practices designed to cultivate positive relationships (Connection) or insight into the nature of the self (Insight).

Methods: A three-arm, fully remote RCT compared 8 weeks of one of 2 HMP conditions (Awareness+Connection and Awareness+Insight) with a waitlist control. Adults (≥18 years) without extensive previous meditation experience were eligible. The primary outcome was psychological distress (depression, anxiety, and stress). Secondary outcomes were social connection, empathy, compassion, self-reflection, insight, rumination, defusion, and mindfulness. Measures were completed at pretest, midtreatment, and posttest between October 2019 and April 2020. Longitudinal data were analyzed using intention-to-treat principles with maximum likelihood.

Results: A total of 343 participants were randomized and 186 (54.2%) completed at least one posttest assessment. The majority (166/228, 72.8%) of those assigned to HMP conditions downloaded the app. The 2 HMP conditions did not differ from one another in terms of changes in any outcome. Relative to the waitlist control, the HMP conditions showed larger improvements in distress, social connectedness, mindfulness, and measures theoretically linked to insight training (d=–0.28 to 0.41; Ps≤.02), despite modest exposure to connection- and insight-related practice. The results were robust to some assumptions about nonrandom patterns of missing data. Improvements in distress were associated with days of use. Candidate mediators (social connection, insight, rumination, defusion, and mindfulness) and moderators (baseline rumination, defusion, and empathy) of changes in distress were identified.

Conclusions: This study provides initial evidence of efficacy for the HMP app in reducing distress and improving outcomes related to well-being, including social connectedness. Future studies should attempt to increase study retention and user engagement.
Introduction

Background

Mindfulness and meditation have become household words for many people in the United States and across the globe in the past 20 years. Derived from Buddhist and Hindu contemplative traditions [1], secularized meditative practices are being taught in schools, recommended by health care providers, and employed by businesses [2-4]. The use of meditation tripled in the United States between 2012 and 2017 (from 4.1% to 14.2%) [5]. Meta-analyses involving hundreds of randomized controlled trials (RCTs) suggest that meditation training can decrease psychological symptoms (eg, depression, anxiety, stress) and increase aspects of well-being and positive functioning (eg, meaning in life, compassion, prosocial behavior) [6-15].

To date, the vast majority of research on meditation has focused on interventions delivered in person. Standardized mindfulness-based interventions such as mindfulness-based stress reduction (MBSR) [16] and mindfulness-based cognitive therapy (MBCT) [17] were explicitly designed as group-based interventions delivered by trained instructors, and these interpersonal elements are viewed as central ingredients (eg, group format) [18]. Despite some meditation-based interventions (MBIs) being recommended as first-line treatments (eg, MBCT for depression relapse prevention) [19-21], their availability remains limited [22]. Barriers for the dissemination of MBIs are similar to those facing other evidence-based psychotherapies (eg, lack of available providers, cost, logistical challenges) [23-25].

Delivering interventions through mobile technology has been proposed as a solution for increasing access to psychological interventions, including MBIs [26,27]. Web- and smartphone-based interventions have obvious advantages over traditional in-person delivery in terms of cost and scalability. Furthermore, mobile health (mHealth) interventions can, in theory, do things that in-person interventions typically never do, such as providing access 24 hours a day or customizing content based on passively sensed data (eg, location) [28]. Among mHealth delivery platforms, smartphone-based interventions may be particularly promising, with these devices often kept within arm’s reach, charged, turned on, and being owned by the vast majority of the population [29].

There has been a dramatic increase in the past five years in RCTs testing smartphone-based interventions that include training in meditation [30]. These studies have begun to examine efficacy in various clinical and nonclinical populations [31-38]. Although preliminary, available evidence suggests that smartphone-based interventions that include training in meditation and mindfulness may provide psychological benefits that are similar to in-person MBIs (eg, decreased psychological symptoms, increased positive functioning), albeit smaller in magnitude [30,39-42].

Similar to the in-person MBI literature, RCTs testing the mobile delivery of MBIs have focused almost entirely on mindfulness. The term mindfulness is derived from the Pali word sati, which in Buddhism refers to the cultivation of receptive, present-moment awareness [43]. In the scientific literature, mindfulness can refer to a mental state, trait, or faculty amenable to training [44-47]. Mindfulness-based interventions commonly adopt the definition by Kabat-Zinn [48]: “paying attention in a particular way: on purpose, in the present moment, and nonjudgmentally.” Meta-analyses suggest that dispositional mindfulness along with both short-term (eg, mindfulness inductions) and long-term (eg, mindfulness-based interventions such as MBSR) training are associated with decreased psychiatric symptoms, negative affect, substance use, and neuroticism [9,49-53].

Importantly, mindfulness training represents the implementation of primarily one meditative approach drawn from rich contemplative traditions [54]. Although largely untested, it is possible that a variety of meditative techniques may serve as valuable complements or alternatives to mindfulness. Different meditation practices have shown distinct neural signatures [55,56] and can produce different psychological effects [57]. Dahl et al [54] provide a useful typology for situating mindfulness training within the broader contemplative practice landscape. Using a family resemblance approach, they describe attential, constructive, and deconstructive families. Mindfulness meditation, as implemented in MBSR, falls primarily within the attential family, with training focused on regulating attention. The constructive family includes practices designed to strengthen psychological habits conducive to psychosocial health. This includes connection-based practices that involve cultivating feelings of warmth and friendliness toward oneself and others (eg, gratitude, loving kindness, and compassion practices) [58,59]. Experimental evidence suggests that connection practices increase well-being and decrease psychological symptoms [8,60]. The deconstructive family includes practices designed to modify unhelpful cognitive patterns, particularly regarding one’s view of self and others. Practices in this family involve intentional self-inquiry into the dynamics of conscious experience and the nature of the self with the goal of generating an understanding of cognitive patterns (ie, insight). Deconstructive elements are present in MBCT and cognitive therapy more generally (eg, seeing thoughts as thoughts) [17,61]. However, research on deconstructive meditative practices has been limited.

Smartphone-based meditation interventions have almost exclusively focused on mindfulness training [30,62], although several studies have investigated internet-based interventions that include connection-related practices [63-65]. Although...
some interventions include connection-based practice as one of several guided practices within a general mindfulness framework [32], RCTs primarily examining constructive or deconstructive practices are rare (with some promising exceptions) [66,67].

**This Study**

This study sought to investigate the effects of a self-guided, smartphone-based meditation intervention that included explicit training in constructive and deconstructive families of practices [54]. In a three-arm RCT, we compared training in mindful awareness, paired with Connection or Insight practices, with a waitlist control. As both arms included the Awareness module first, we refer to them by their unique module (ie, Connection or Insight, rather than Awareness+Connection and Awareness+Insight). We included outcome measures designed to detect global effects (psychological distress) and practice-specific effects (eg, social connection, shift in relationship to one’s thoughts). Our primary hypothesis was that participants in both active conditions would show reduced psychological distress relative to the waitlist control. In addition, we expected those randomized to connection practices to show larger improvements in connection-related measures and those randomized to insight practice to show larger improvements in insight-related measures. We had several exploratory secondary hypotheses. We hypothesized that app usage would be positively associated with reduced distress. We hypothesized that improvements in connection- and insight-related measures would mediate effects on distress for those in the Connection and Insight arms, respectively. We hypothesized that those lower in mindfulness at baseline would show larger improvements in the active conditions and that those lower in connection- and insight-related measures would show larger improvements in the Connection and Insight arms, respectively. These hypotheses were preregistered at the Open Science Framework [68].

**Methods**

**Procedure**

We conducted an 8-week, fully remote RCT comparing 2 active smartphone-based meditation interventions with a waitlist control. Participants were recruited through emails sent to faculty, staff, and students at the University of Wisconsin-Madison and through a database of individuals who had previously expressed interest in research at the Center for Healthy Minds. All screening procedures and data collection were web-based and carried out using REDCap [69]. Participants completed a screening protocol to determine eligibility and received their group assignment via an automated email following the completion of baseline questionnaires. Randomization was achieved by automatically allocating participants to groups based on sequentially assigned participant identification numbers (ie, 1:1:1 randomization ratio). Participants were contacted by email to complete questionnaires at the conclusion of the study.

Progress through the material in the Healthy Minds Program (HMP) app was self-guided. There was minimal contact with the study staff. Participants were provided with a study email address to contact for technical support or study-related questions. All procedures were approved by the institutional review board. The study was registered at ClinicalTrials.gov (NCT04139005).

**Participants**

Eligible participants were aged ≥18 years, had access to a smartphone or other device capable of running the intervention app (Android or iOS), and did not have extensive previous meditation experience defined as meditation retreat experience, meditation practice weekly for >1 year or daily practice within the previous 6 months; or previous training under the instruction of a meditation teacher, other than an introductory course. Participants received US $25 for completing the assessments.

**Intervention**

Participants assigned to one of the 2 active intervention arms were instructed to download the HMP app through the Google Play or Apple App Store. The full HMP app includes 4 modules with practices designed to cultivate categories of mental and emotional skills linked to both hedonic and eudaimonic well-being [70,71]. These include the cultivation of mindful attention (Awareness), positive relationships with self and others (Connection), insight into the nature of self and internal experience (Insight), and purpose, values, and meaning in life (Purpose). In this study, the 2 active interventions included 4 weeks of Awareness training, followed by 4 weeks of either Connection or Insight training. This design was predicated on the view that training in the stabilization of attention is foundational to skills trained by Connection and Insight [72]. Each module included brief, podcast-style didactic material along with guided meditation practices. Didactic content included discussion of the scientific bases of the practices. Participants were encouraged to follow a prespecified sequence while going through the material. Participants could select the length of the guided practices (5-30 min) and a variety of practices were available in each module. For example, the Awareness module included practices focused on awareness of breathing and mindfulness of sound. The Connection module included gratitude and kindness practices. The Insight module included practices involving noticing the changing nature of the phenomenon (ie, impermanence) and examining how thoughts and emotions influence perception. Participants in the waitlist condition received access to the full HMP app (ie, all 4 modules) at the conclusion of the study.

**Measures**

A demographic questionnaire was completed at baseline. App usage was measured objectively using the HMP app. Additional information about the psychometric properties and theoretical relevance of the included measures is provided in Multimedia Appendix 1 [73-101].

**Psychological Distress**

A psychological distress composite score was created from measures of depression, anxiety, and stress. We computed the mean across scaled (z-transformed) scores for each measure. The 8-item Patient-Reported Outcome Measures Information System Depression and Anxiety Scales [73] assessed depression and anxiety. Items are rated on a 5-point scale (1=never; 5=always), with higher scores indicating greater severity in the
past 7 days. T scores ≥55 suggest mild or greater severity [102,103]. Internal consistency was high (=.93 to .94).

The 14-item Perceived Stress Scale [74] assessed psychological stress. Items are rated on a 5-point scale (0=never; 4=very often), with higher scores indicating greater stress in the past month. Internal consistency was high (=.89).

**Measures Related to the Connection Module**

The 20-item Social Connectedness Scale-Revised [75] assessed interpersonal connections. Items are rated on a 6-point scale (1=very much; 6=very true of me). Higher scores indicate greater interpersonal connectedness. Internal consistency was high (=.95).

The 28-item Interpersonal Reactivity Index (IRI) [76] assessed empathy. Items are rated on a 6-point scale (0=not well; 6=very well), with higher scores indicating greater empathy. Internal consistency was high (=.95).

The 21-item Compassionate Love Scale [77] assesses feelings of compassion. Items are rated on a 7-point scale (1=not at all true of me; 7=very true of me). Higher scores indicate greater feelings of compassion. Internal consistency was high (=.95).

**Measures Related to the Insight Module**

The 20-item Self-Reflection and Insight Scale (SRIS) [78] assessed participants’ tendency toward self-reflection (eg, “I frequently examine my feelings”) and self-understanding or insight (eg, “I usually know why I feel the way I do”). Items are rated on a 6-point scale (1=strongly disagree; 6=strongly agree) and yield subscales for self-reflection and insight, with higher scores indicating greater self-reflection or insight. Internal consistency was high (=.88 to .92).

The 15-item Perseverative Thinking Questionnaire (PTQ) [79] assessed rumination. For simplicity, we used the term “rumination” to refer to repetitive negative thinking as captured by the PTQ, although it captures both rumination and worry. Items are rated on a 6-point scale (0=not at all; 6=almost always), with higher scores indicating greater rumination. Internal consistency was high (=.96).

The 10-item Drexel Defusion Scale [80] assessed the ability to achieve psychological distance from internal experiences (ie, defusion). Items are rated on a 6-point scale (0=not at all; 5=very much), with higher scores indicating greater defusion. Internal consistency was high (=.89).

**Mindfulness**

The 39-item Five Facet Mindfulness Questionnaire [81] assessed mindfulness. Items are rated on a 5-point scale (1=never or very rarely true; 5=very often or always true), with higher scores indicating greater mindfulness. Internal consistency was high for the total score (=.94).

**Data Analysis**

Results from all preregistered primary and secondary measures are reported. For deviations made from the preregistered data analytic plan, see Multimedia Appendix 1.
adjustment [85] was applied to all analyses to control for inflation of a type I error.

Sample Size and Power

We planned to recruit 300 participants (100 per group), which would allow the detection of small-to-moderate differences between any 2 groups ($d=0.40$) and between the active and waitlist control conditions ($d=0.34$) at 80% power and $P=.05$. Power was estimated using the `pwr.test` and `pwr2n.test` functions in the `pwr` package in R [109].

Results

Recruitment and Participant Characteristics

A total of 954 potential participants were assessed for eligibility, of which 343 met the inclusion criteria and were randomized to Connection (n=121), Insight (n=107), or waitlist (n=115; Figure 1). Demographics are reported in Table 1. The sample was predominantly White (280/343, 81.6%), female (290/343, 84.5%), and with graduate-level education (190/343, 55.4%). Income was more variable (89/343, 25.9% earned US $50,000 or less). The mean age was 41.74 (SD 12.52) years.

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram. Mid=week 4 assessment; Post=week 8 assessment.
### Table 1. Sample demographics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall (n=343)</th>
<th>Connection (n=121)</th>
<th>Insight (n=107)</th>
<th>Waitlist (n=115)</th>
<th>( P ) value(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Race and ethnicity, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.91</td>
</tr>
<tr>
<td>White</td>
<td>280 (81.6)</td>
<td>99 (81.8)</td>
<td>86 (80.4)</td>
<td>95 (82.6)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>6 (1.7)</td>
<td>3 (2.5)</td>
<td>1 (0.9)</td>
<td>2 (1.7)</td>
<td></td>
</tr>
<tr>
<td>Latinx</td>
<td>4 (1.2)</td>
<td>0 (0.0)</td>
<td>4 (3.7)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>18 (5.2)</td>
<td>7 (5.8)</td>
<td>6 (5.6)</td>
<td>5 (4.3)</td>
<td></td>
</tr>
<tr>
<td>Multiracial</td>
<td>33 (9.6)</td>
<td>12 (9.9)</td>
<td>9 (8.4)</td>
<td>12 (10.4)</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>2 (0.6)</td>
<td>0 (0.0)</td>
<td>1 (0.9)</td>
<td>1 (0.9)</td>
<td></td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.93</td>
</tr>
<tr>
<td>Female</td>
<td>290 (84.5)</td>
<td>101 (83.5)</td>
<td>89 (83.2)</td>
<td>98 (85.2)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>51 (14.9)</td>
<td>20 (16.5)</td>
<td>15 (14.0)</td>
<td>16 (13.9)</td>
<td></td>
</tr>
<tr>
<td>Nonbinary</td>
<td>2 (0.6)</td>
<td>0 (0.0)</td>
<td>1 (0.9)</td>
<td>1 (0.9)</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>2 (0.6)</td>
<td>0 (0.0)</td>
<td>2 (1.9)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Income (US $), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.18</td>
</tr>
<tr>
<td>(\leq 50,000)</td>
<td>89 (25.9)</td>
<td>33 (27.3)</td>
<td>26 (24.3)</td>
<td>30 (26.1)</td>
<td></td>
</tr>
<tr>
<td>50,000-100,000</td>
<td>120 (35.0)</td>
<td>35 (28.9)</td>
<td>37 (34.6)</td>
<td>48 (41.7)</td>
<td></td>
</tr>
<tr>
<td>100,000-150,000</td>
<td>76 (22.2)</td>
<td>32 (26.4)</td>
<td>22 (20.6)</td>
<td>22 (19.1)</td>
<td></td>
</tr>
<tr>
<td>(&gt;150,000)</td>
<td>57 (16.6)</td>
<td>21 (17.4)</td>
<td>21 (19.6)</td>
<td>15 (13.0)</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>1 (0.3)</td>
<td>0 (0.0)</td>
<td>1 (0.9)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Education, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.45</td>
</tr>
<tr>
<td>Some high school</td>
<td>1 (0.3)</td>
<td>0 (0.0)</td>
<td>1 (0.9)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>High school graduate</td>
<td>6 (1.7)</td>
<td>3 (2.5)</td>
<td>0 (0.0)</td>
<td>3 (2.6)</td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>32 (9.3)</td>
<td>9 (7.4)</td>
<td>11 (10.3)</td>
<td>12 (10.4)</td>
<td></td>
</tr>
<tr>
<td>College graduate</td>
<td>114 (33.2)</td>
<td>37 (30.6)</td>
<td>36 (33.6)</td>
<td>41 (35.7)</td>
<td></td>
</tr>
<tr>
<td>Graduate school</td>
<td>190 (55.4)</td>
<td>72 (59.5)</td>
<td>59 (55.1)</td>
<td>59 (51.3)</td>
<td></td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>41.74 (12.52)</td>
<td>42.31 (12.8)</td>
<td>43.21 (12.39)</td>
<td>39.78 (12.2)</td>
<td>.10</td>
</tr>
<tr>
<td>Elevated symptoms(^b), n (%)</td>
<td>252 (73.5)</td>
<td>86 (71.1)</td>
<td>81 (75.7)</td>
<td>85 (73.9)</td>
<td>.73</td>
</tr>
</tbody>
</table>

\(^a\)\( P \) values based on a one-way analysis of variance with group (Connection, Insight, or waitlist) predicting demographics (White, female, high income [\(\geq US \$100,000\)], and graduate school).

\(^b\)Elevated symptoms: Patient-Reported Outcomes Measurement Information System (PROMIS) Depression or PROMIS Anxiety in the mild or higher range (T \(\geq 55\)).

### Utilization

Of those randomized to one of the 2 active conditions, 77.7% (94/121) of Connection participants and 67.3% (72/107) of Insight participants downloaded and used the HMP app at least once. By assigning values of zero to those who did not use the app, we found that average utilization was 10.52 days (SD 13.31; median 4), with 18.09 activities within the app (SD 23.30; median 7), 9.45 meditation practices (SD 13.34; median 3), and 102.16 total min of meditation practice (SD 187.74; median 26). All usage metrics were highly zero inflated (Multimedia Appendix 1). Days of use had the lowest skewness (1.34) and kurtosis (0.91), so a median split of days of use was used in the analyses. The median survival time (ie, time before last use) was 12 days. Group status (Connection vs Insight) was not associated with usage (\( P=15 \)), and survival time did not differ between groups (Multimedia Appendix 1; \( P=.24 \)). As Connection or Insight content was provided at week 5 of the program, 32.2% (39/121) of Connection and 23.4% (25/107) of Insight participants engaged with the unique content. This proportion did not differ between groups (OR 0.64, 95% CI 0.35-1.15; \( P=.14 \)).

### Attrition Analysis

We examined baseline demographic and outcome variables as predictors of attrition. We constructed logistic regression models predicting the presence of any follow-up data (ie, midtreatment or posttest). Participants were invited to complete the posttest measures even if they had not completed the midtreatment measures. The average completion of at least one follow-up assessment (mid- or posttest) was 54.2% (186/343). Waitlist participants were more likely to complete follow-up
assessments (77/115, 67.0% vs 109/228, 47.8%; waitlist n=77; Connection and Insight combined n=109; OR 2.21, 95% CI 1.39-3.56; P<.001). Completion of follow-up assessments did not differ between the Connection and Insight groups (OR 0.99, 95% CI 0.59-1.67; P=.97). However, participants who used the app at least once were more likely to complete the follow-up assessments (OR 3.66, 95% CI 1.95-7.16; P<.001). Completion of follow-up was not associated with demographics (White, female, high income [≥US $100,000], and graduate education) or outcome measures at baseline (Ps≥.15), with one exception. Participants with higher empathy scores (IRI) at baseline were more likely to complete follow-up assessments (OR 1.02, 95% CI 1.00-1.04; P=.04).

**Primary Analyses**

Correlations between outcomes are reported in Multimedia Appendix 1. The 3 groups did not differ in any demographic or outcome measures at baseline (Ps≥.10; Tables 1 and 2). Within- and between-group effect sizes (Cohen d) and P values from MLMs are reported in Table 3. The 2 active conditions did not differ from one another in terms of change over time for distress or any secondary outcomes (time×group, Ps≥.29). Therefore, all subsequent analyses combined the 2 active groups. When compared with the waitlist control, the active conditions showed greater decreases in distress (d=0.28) and rumination (d=0.18) and greater increases in social connectedness, self-reflection, insight, defusion, and mindfulness (d=0.13 to 0.41; FDR-adjusted P≤.02; Figure 2). The active conditions did not differ from the waitlist on changes in empathy (d=0.02) or compassion (d=0.12). Significance tests for time×group interactions did not change when restricting to those with elevated symptoms at baseline (Table 2), when excluding outliers (with the exception of Self-Reflection, P=.05; Multimedia Appendix 1), nor when each case was excluded sequentially.

A larger proportion of participants in the active conditions showed a minimally important decrease in distress (d≤–0.30) [86] relative to the waitlist condition (70% vs 49%; Connection and Insight combined n=64/91; waitlist n=33/67; OR 2.44 [95% CI 1.27-4.75]; P=.008). A smaller proportion in the active condition showed a minimally important increase in distress (ie, deterioration, d≥0.30) relative to the waitlist condition (3% vs 16%; Connection and Insight combined n=3/91; waitlist n=11/67; OR 0.17 [0.04, 0.58]; P=.009).
Table 2. Descriptive statistics for repeated measures by group and timepoint.

<table>
<thead>
<tr>
<th>Group and outcome</th>
<th>Pretest</th>
<th>Midtreatment</th>
<th>Posttest</th>
<th>( P ) value(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean (SD)</td>
<td>n</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td><strong>CO(^b)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological distress(^c)</td>
<td>121</td>
<td>0.08 (0.90)</td>
<td>49</td>
<td>0.74 (0.80)</td>
</tr>
<tr>
<td>Social connection(^d)</td>
<td>121</td>
<td>83.41 (21.43)</td>
<td>48</td>
<td>90.62 (17.29)</td>
</tr>
<tr>
<td>Empathy(^e)</td>
<td>121</td>
<td>67.67 (12.33)</td>
<td>48</td>
<td>67.42 (12.55)</td>
</tr>
<tr>
<td>Compassion(^f)</td>
<td>121</td>
<td>99.18 (22.19)</td>
<td>48</td>
<td>102.52 (21.71)</td>
</tr>
<tr>
<td>Self-reflection subscale(^g)</td>
<td>121</td>
<td>56.31 (11.13)</td>
<td>48</td>
<td>58.45 (10.29)</td>
</tr>
<tr>
<td>Insight subscale(^h)</td>
<td>121</td>
<td>34.12 (7.46)</td>
<td>48</td>
<td>37.06 (6.17)</td>
</tr>
<tr>
<td>Ruminatin(^b)</td>
<td>121</td>
<td>31.12 (12.65)</td>
<td>46</td>
<td>25.31 (9.63)</td>
</tr>
<tr>
<td>Defusion(^i)</td>
<td>121</td>
<td>22.95 (9.37)</td>
<td>48</td>
<td>28.54 (8.37)</td>
</tr>
<tr>
<td>Mindfulness(^j)</td>
<td>121</td>
<td>122.16 (20.41)</td>
<td>49</td>
<td>136.28 (16.37)</td>
</tr>
<tr>
<td><strong>IN(^k)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological distress</td>
<td>107</td>
<td>0.04 (0.94)</td>
<td>43</td>
<td>0.52 (0.88)</td>
</tr>
<tr>
<td>Social connection</td>
<td>107</td>
<td>83.28 (19.04)</td>
<td>41</td>
<td>87.57 (19.28)</td>
</tr>
<tr>
<td>Empathy</td>
<td>107</td>
<td>69.77 (12.45)</td>
<td>41</td>
<td>70.24 (11.76)</td>
</tr>
<tr>
<td>Compassion</td>
<td>107</td>
<td>101.57 (21.44)</td>
<td>41</td>
<td>101.07 (24.52)</td>
</tr>
<tr>
<td>Self-reflection subscale</td>
<td>107</td>
<td>55.99 (10.40)</td>
<td>41</td>
<td>55.90 (9.00)</td>
</tr>
<tr>
<td>Insight subscale</td>
<td>107</td>
<td>32.84 (8.42)</td>
<td>41</td>
<td>33.67 (8.28)</td>
</tr>
<tr>
<td>Ruminatin</td>
<td>107</td>
<td>30.49 (11.71)</td>
<td>41</td>
<td>27.21 (10.79)</td>
</tr>
<tr>
<td>Defusion</td>
<td>107</td>
<td>23.62 (10.36)</td>
<td>41</td>
<td>27.36 (10.01)</td>
</tr>
<tr>
<td>Mindfulness</td>
<td>107</td>
<td>121.55 (24.96)</td>
<td>43</td>
<td>128.56 (22.96)</td>
</tr>
<tr>
<td><strong>WL(^l)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychological distress</td>
<td>115</td>
<td>0.05 (0.88)</td>
<td>64</td>
<td>0.23 (1.01)</td>
</tr>
<tr>
<td>Social connection</td>
<td>115</td>
<td>81.53 (19.68)</td>
<td>60</td>
<td>82.78 (21.10)</td>
</tr>
<tr>
<td>Empathy</td>
<td>115</td>
<td>69.31 (12.85)</td>
<td>60</td>
<td>71.09 (13.49)</td>
</tr>
<tr>
<td>Compassion</td>
<td>115</td>
<td>98.62 (22.39)</td>
<td>58</td>
<td>99.53 (22.64)</td>
</tr>
<tr>
<td>Self-reflection subscale</td>
<td>115</td>
<td>56.34 (10.78)</td>
<td>57</td>
<td>56.87 (11.14)</td>
</tr>
<tr>
<td>Insight subscale</td>
<td>115</td>
<td>33.22 (8.02)</td>
<td>57</td>
<td>35.40 (8.29)</td>
</tr>
<tr>
<td>Ruminatin</td>
<td>115</td>
<td>31.14 (11.56)</td>
<td>56</td>
<td>29.51 (13.04)</td>
</tr>
<tr>
<td>Defusion</td>
<td>115</td>
<td>23.49 (9.58)</td>
<td>60</td>
<td>24.88 (9.95)</td>
</tr>
<tr>
<td>Mindfulness</td>
<td>115</td>
<td>120.16 (18.93)</td>
<td>62</td>
<td>125.15 (20.21)</td>
</tr>
</tbody>
</table>

\(^a\)\(P\) value from a one-way analysis of variance predicting baseline values for outcome measures by group status.

\(^b\)CO: Awareness+Connection.

\(^c\)Composite of Patient-Reported Outcomes Measurement Information System (PROMIS) Depression, PROMIS Anxiety, and Perceived Stress Scale.

\(^d\)Social Connectedness Scale.

\(^e\)Interpersonal Reactivity Index.

\(^f\)Compassionate Love Scale.

\(^g\)Subscales of the Self-Reflection and Insight Scale.

\(^h\)Perseverative Thinking Questionnaire.

\(^i\)Drexel Defusion Scale.

\(^j\)Total score of Five Facet Mindfulness Questionnaire.
Table 3. Results of multilevel models assessing differential change over time.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>CO^a versus IN^b</th>
<th>Active versus WL^c</th>
<th>d_CO</th>
<th>d_IN</th>
<th>d_diff</th>
<th>P value^d</th>
<th>P_FDR</th>
<th>d_Active</th>
<th>d_WL</th>
<th>d_diff</th>
<th>P value</th>
<th>P_FDR</th>
<th>Elev P^e</th>
<th>Elev P_FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychological distress^j</td>
<td></td>
<td></td>
<td>-0.77</td>
<td>-0.70</td>
<td>-0.07</td>
<td>.06</td>
<td>.97</td>
<td>-0.74</td>
<td>-0.46</td>
<td>-0.28</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Social connection^l</td>
<td></td>
<td></td>
<td>0.42</td>
<td>0.36</td>
<td>0.06</td>
<td>0.54</td>
<td>0.82</td>
<td>0.39</td>
<td>0.16</td>
<td>0.23</td>
<td>.003</td>
<td>.007</td>
<td>.01</td>
<td>.02</td>
</tr>
<tr>
<td>Empathy^k</td>
<td>-0.14</td>
<td>-0.02</td>
<td>-0.12</td>
<td>.37</td>
<td>.82</td>
<td>-0.08</td>
<td>-0.10</td>
<td>0.02</td>
<td>.63</td>
<td>.48</td>
<td>.48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Compassion^l</td>
<td>0.11</td>
<td>0.34</td>
<td>-0.23</td>
<td>.29</td>
<td>.82</td>
<td>0.22</td>
<td>0.10</td>
<td>0.12</td>
<td>.14</td>
<td>.18</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-reflection subscale^m</td>
<td>0.11</td>
<td>0.18</td>
<td>-0.07</td>
<td>.51</td>
<td>.82</td>
<td>0.14</td>
<td>-0.01</td>
<td>0.15</td>
<td>.007</td>
<td>.02</td>
<td>.03</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insight subscale^m</td>
<td>0.46</td>
<td>0.39</td>
<td>0.07</td>
<td>.98</td>
<td>.98</td>
<td>0.42</td>
<td>0.29</td>
<td>0.13</td>
<td>.02</td>
<td>.02</td>
<td>.001</td>
<td>.001</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Ruminations^b</td>
<td>-0.45</td>
<td>-0.56</td>
<td>0.11</td>
<td>.32</td>
<td>.82</td>
<td>-0.5</td>
<td>-0.32</td>
<td>-0.18</td>
<td>.01</td>
<td>.02</td>
<td>.007</td>
<td>.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defusion^o</td>
<td>0.83</td>
<td>0.66</td>
<td>0.17</td>
<td>.78</td>
<td>.97</td>
<td>0.75</td>
<td>0.34</td>
<td>0.41</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mindfulness^p</td>
<td>0.86</td>
<td>0.70</td>
<td>0.16</td>
<td>.55</td>
<td>.82</td>
<td>0.77</td>
<td>0.46</td>
<td>0.31</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^aCO: Awareness+Connection.
^bIN: Awareness+Insight.
^cWL: waitlist.
^dCohen d calculated as pre-post for within-group effects and the difference between within-group effects (Connection–Insight, active–waitlist) for d_diff. For within-group, subscripted CO (ie, d_CO), IN, Active, and WL refer to subgroups noted.
^eP value from time×group interaction from multilevel models.
^fFDR: false discovery rate adjusted P values.
^gCombined Awareness+Connection and Awareness+Insight.
^hActive versus waitlist time×group interaction restricted to sample with elevated depression or anxiety at baseline (T≥55).
^iComposite of Patient-Reported Outcomes Measurement Information System (PROMIS) Depression, PROMIS Anxiety, and Perceived Stress Scale.
^jSocial connection: Social Connectedness Scale.
^kInterpersonal Reactivity Index.
^lCompassionate Love Scale.
^mSubscales of the Self-Reflection and Insight Scale.
^nPerseverative Thinking Questionnaire.
^oDrexel Defusion Scale.
^pTotal score of Five Facet Mindfulness Questionnaire.
Robustness Check: Sensitivity Analyses

Although maximum likelihood is robust to data MAR [83], subsequent analyses evaluated treatment effects based on varying assumptions under MNAR conditions. Using the completer sample, a Wilcoxon rank sum test on the residualized gain score mirrored the MLM results, with larger improvements in the active conditions relative to the waitlist on several outcomes (FDR-adjusted \( P \leq 0.047 \); Multimedia Appendix 1). In the worst-case scenario model in which missing reflects the worst possible outcome across both active and waitlist groups, the groups did not differ, although the direction of the mean rank favored the waitlist group for all outcomes. Thus, we examined the results in between these extreme conditions to understand where significance goes away and where the direction of intervention effect reverses. When we assumed that missing values are on average 0.25 SD above the mean (implying worse than average outcomes for the missing observations), the results continued to favor the active conditions for changes in distress, social connectedness, defusion, and mindfulness (FDR-adjusted \( P \leq 0.04 \); Multimedia Appendix 1). When we assumed that missing values were on average 0.50 SD above the mean, the differences between groups were not statistically significant for any outcome. The difference remained nonsignificant when we assumed that missing values are on average 0.75 SD above the mean. Thus, it appears that our results are robust to MNAR up to a point, specifically that missing outcomes are no more than 0.25 SD above the mean on average, under the assumption that missingness implies comparable outcomes for both the active and waitlist groups.

Secondary Analyses

The results of the usage analyses are reported in Table 4. HMP use above the median number of days was associated with larger improvements in distress, insight, defusion, and mindfulness (FDR-adjusted \( P \leq 0.03 \); Figure 3).

The results of the mediation analyses are reported in Table 5. Changes in 5 candidate mechanisms showed a significant average causal mediation effect (FDR-adjusted \( P \leq 0.04 \)) in the expected direction (ie, improvements in social connection, insight, rumination, defusion, and mindfulness mediated improvements in distress). Changes in mindfulness were associated with the largest proportion mediated (0.45).
Table 4. Results of multilevel models predicting changes in outcomes from Healthy Minds Program app usage (n=228).

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time×usage $B^a$</th>
<th>$t$ test $(df)^b$</th>
<th>$P$ value$^c$</th>
<th>$P_{FDR}^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychological distress$^e$</td>
<td>-0.17</td>
<td>-2.46 (207)</td>
<td>.02</td>
<td>.03</td>
</tr>
<tr>
<td>Social connection$^f$</td>
<td>2.42</td>
<td>1.61 (202)</td>
<td>.11</td>
<td>.17</td>
</tr>
<tr>
<td>Empathy$^g$</td>
<td>-0.47</td>
<td>-0.58 (200)</td>
<td>.56</td>
<td>.56</td>
</tr>
<tr>
<td>Compassion$^h$</td>
<td>-1.16</td>
<td>-0.76 (201)</td>
<td>.45</td>
<td>.50</td>
</tr>
<tr>
<td>Self-reflection subscale$^i$</td>
<td>0.80</td>
<td>0.87 (218)</td>
<td>.38</td>
<td>.49</td>
</tr>
<tr>
<td>Insight subscale$^i$</td>
<td>1.55</td>
<td>2.46 (211)</td>
<td>.02</td>
<td>.03</td>
</tr>
<tr>
<td>Rumination$^j$</td>
<td>-1.85</td>
<td>-2.13 (197)</td>
<td>.03</td>
<td>.06</td>
</tr>
<tr>
<td>Defusion$^k$</td>
<td>2.40</td>
<td>2.64 (229)</td>
<td>.009</td>
<td>.03</td>
</tr>
<tr>
<td>Mindfulness$^l$</td>
<td>5.17</td>
<td>2.93 (212)</td>
<td>.004</td>
<td>.03</td>
</tr>
</tbody>
</table>

$^a$Multilevel model regression coefficient. Usage: days of use split into high (median or above) and low (below median) groups.

$^b$ $t$ statistic for time×usage interaction with associated degrees of freedom $(df)$.

$^c$ $P$ value for time×usage interaction.

$^d$ FDR: false discovery rate adjusted $P$ values.

$^e$ Composite of Patient-Reported Outcomes Measurement Information System (PROMIS) Depression, PROMIS Anxiety, and Perceived Stress Scale.

$^f$ Social Connectedness Scale.

$^g$ Interpersonal Reactivity Index.

$^h$ Compassionate Love Scale.

$^i$ Subscales of the Self-Reflection and Insight Scale.

$^j$ Perseverative Thinking Questionnaire.

$^k$ Drexel Defusion Scale.

$^l$ Total score of Five Facet Mindfulness Questionnaire.

Figure 3. Healthy Minds Program app usage predicting longitudinal changes in psychological distress, defusion, insight, and mindfulness in active conditions (FDR-corrected $P$≤.04). Usage=median split of days of use (n=228). HMP: Healthy Minds Program; WL: waitlist.
Table 5. Results of mediation analyses predicting changes in psychological distress.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>ACME&lt;sup&gt;a&lt;/sup&gt;</th>
<th>ADE&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Prop mediated&lt;sup&gt;c&lt;/sup&gt;</th>
<th>P value&lt;sup&gt;d&lt;/sup&gt;</th>
<th>P&lt;sub&gt;FDR&lt;/sub&gt;&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social connection&lt;sup&gt;f&lt;/sup&gt;</td>
<td>−0.10</td>
<td>−0.32</td>
<td>0.24</td>
<td>.002</td>
<td>.005</td>
</tr>
<tr>
<td>Empathy&lt;sup&gt;g&lt;/sup&gt;</td>
<td>0.02</td>
<td>−0.44</td>
<td>−0.03</td>
<td>.40</td>
<td>.47</td>
</tr>
<tr>
<td>Compassion&lt;sup&gt;h&lt;/sup&gt;</td>
<td>0.01</td>
<td>−0.43</td>
<td>−0.01</td>
<td>.59</td>
<td>.59</td>
</tr>
<tr>
<td>Self-reflection subscale&lt;sup&gt;i&lt;/sup&gt;</td>
<td>0.02</td>
<td>−0.44</td>
<td>−0.04</td>
<td>.41</td>
<td>.47</td>
</tr>
<tr>
<td>Insight subscale&lt;sup&gt;i&lt;/sup&gt;</td>
<td>−0.06</td>
<td>−0.37</td>
<td>0.13</td>
<td>.03</td>
<td>.04</td>
</tr>
<tr>
<td>Rumination&lt;sup&gt;i&lt;/sup&gt;</td>
<td>−0.10</td>
<td>−0.32</td>
<td>0.23</td>
<td>.02</td>
<td>.04</td>
</tr>
<tr>
<td>Defusion&lt;sup&gt;k&lt;/sup&gt;</td>
<td>−0.10</td>
<td>−0.32</td>
<td>0.23</td>
<td>.002</td>
<td>.005</td>
</tr>
<tr>
<td>Mindfulness&lt;sup&gt;l&lt;/sup&gt;</td>
<td>−0.18</td>
<td>−0.22</td>
<td>0.45</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

<sup>a</sup>ACME: average causal mediation effect (ie, indirect effect).

<sup>b</sup>ADE: average direct effect (ie, from active to posttreatment distress controlling for pretreatment distress, when active=1 and waitlist=0).

<sup>c</sup>Proportion mediated computed as indirect effect (ie, ACME) divided by total effect [108].

<sup>d</sup>P value based on quasi-Bayesian CIs.

<sup>e</sup>FDR: false discovery rate adjusted P values. Models examining pre-post change in constructs related to Awareness, Connection, and Insight modules as mediators of pre-post change in (composite of PROMIS Depression, PROMIS Anxiety, and Perceived Stress Scale). Proportion mediated can be negative in instances where direct effect and indirect effect have opposite signs.

<sup>f</sup>Social connection: Social Connectedness Scale.

<sup>g</sup>Interpersonal Reactivity Index.

<sup>h</sup>Compassionate Love Scale.

<sup>i</sup>Subscales of the Self-Reflection and Insight Scale.

<sup>j</sup>Perseverative Thinking Questionnaire.

<sup>k</sup>Drexel Defusion Scale.

<sup>l</sup>Total score of Five Facet Mindfulness Questionnaire.

The results of the baseline moderation analyses are reported in Table 6. A total of 3 baseline variables showed significant time×group×baseline interactions after FDR adjustment. Psychological vulnerability, as indicated by 2 outcomes (higher rumination and lower defusion) at baseline, was associated with significant improvements in distress in the HMP conditions relative to the waitlist condition. Baseline empathy showed the opposite pattern, with those higher at baseline showing significant improvements in distress in HMP relative to the waitlist condition (Multimedia Appendix 1).
Table 6. Baseline outcomes as moderators of longitudinal changes in psychological distress.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Time×group B (^a)</th>
<th>Time×group×baseline B</th>
<th>(t) test (df) (^b)</th>
<th>(P) value (^c)</th>
<th>(P) _FDR (^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social connection (^e)</td>
<td>−0.64</td>
<td>0.01</td>
<td>2.16 (363)</td>
<td>.03</td>
<td>.06</td>
</tr>
<tr>
<td>Empathy (^f)</td>
<td>0.67</td>
<td>−0.01</td>
<td>−3.23 (343)</td>
<td>.001</td>
<td>.008</td>
</tr>
<tr>
<td>Compassion (^g)</td>
<td>0.07</td>
<td>0.00</td>
<td>−1.20 (343)</td>
<td>.23</td>
<td>.29</td>
</tr>
<tr>
<td>Self-reflection subscale (^h)</td>
<td>−0.59</td>
<td>0.01</td>
<td>1.55 (339)</td>
<td>.12</td>
<td>.19</td>
</tr>
<tr>
<td>Insight subscale (^h)</td>
<td>−0.44</td>
<td>0.01</td>
<td>1.14 (359)</td>
<td>.25</td>
<td>.29</td>
</tr>
<tr>
<td>Ruminiation (^i)</td>
<td>0.21</td>
<td>−0.01</td>
<td>−3.15 (390)</td>
<td>.002</td>
<td>.008</td>
</tr>
<tr>
<td>Defusion (^j)</td>
<td>−0.54</td>
<td>0.01</td>
<td>2.86 (361)</td>
<td>.004</td>
<td>.01</td>
</tr>
<tr>
<td>Mindfulness (^k)</td>
<td>−0.44</td>
<td>0.00</td>
<td>0.82 (372)</td>
<td>.41</td>
<td>.41</td>
</tr>
</tbody>
</table>

\(^a\) Multilevel model regression coefficient.
\(^b\) \(t\) test: \(t\) statistic for time×group×baseline (with group coded as active=1, waitlist=0) with associated degrees of freedom (df).
\(^c\) \(P\) value for time×group×baseline.
\(^d\) FDR: false discovery rate adjusted \(P\) values.
\(^e\) Social Connectedness Scale.
\(^f\) Interpersonal Reactivity Index.
\(^g\) Compassionate Love Scale.
\(^h\) Subscales of the Self-Reflection and Insight Scale.
\(^i\) Perseverative Thinking Questionnaire.
\(^j\) Drexel Defusion Scale.
\(^k\) Total score of Five Facet Mindfulness Questionnaire.

Discussion

Principal Findings

This study sought to expand the scientific understanding of the impact of smartphone-delivered meditation training beyond mindfulness. To do so, we evaluated the effects of mindfulness training (Awareness) paired with practices designed to cultivate kindness toward oneself and others (Connection) or insight into the nature of self and internal experience (Insight). We assessed the effects on psychological distress and constructs theoretically linked to connection- and insight-based training [54].

Contrary to our expectations, there was no indication that training in connection produced differential effects relative to insight-related practices. There are several potential reasons for this. One likely explanation is that the actual content completed by each group was largely overlapping. Both groups began with foundational mindfulness training. Given the modest engagement (a perennial concern in mHealth interventions) [110,111], most participants did not engage with the unique Connection or Insight modules. It is also possible that meditation training produces similar effects for novices, regardless of the specific type of training. Novices may spend much of their initial meditation practice simply regathering a wandering attention, regardless of the actual practice instructions. Indeed, studies showing distinct neural signatures associated with various forms of meditation practice have primarily been conducted with long-term practitioners with thousands of hours of experience [55]. A third possibility is that various forms of meditation training contain common ingredients (eg, acceptance, curiosity) that may, especially early in training, be more potent than style-specific ingredients.

Despite the absence of differential effects, the results suggest that meditation delivered via smartphones produced small reductions in psychological distress (\(d=−0.28\)) and improvements in several candidate mechanisms relative to a waitlist control (\(d=−0.18\) to 0.41). These results are generally consistent with meta-analyses of the broader mHealth and mHealth MBI literature, which has shown small benefits of self-guided smartphone apps on depression and anxiety symptoms (\(g=0.21\) to 0.23) and measures of mindfulness and acceptance (\(g=0.27\)) [30]. These effects are considerably smaller than those produced by in-person MBIs (eg, \(d=0.55\) vs waitlist) [9]. It is likely that mHealth MBIs may be less potent than in-person interventions, indicating trade-offs between scalability, cost, and potency. On the basis of those completing posttreatment measures, HMP appears safe in that the rates of clinically significant increases in distress were rare (3%) and were less common than the rates in the control condition (16%). This finding is consistent with a recent large-scale evaluation of the deterioration in MBSR [87].

One important caveat for interpreting our findings is high attrition, particularly within the active conditions. Both high attrition and differential attrition are common in mHealth research [107,112]. Our overall attrition rate was almost identical to that typically found in RCTs testing smartphone interventions without telephone or in-person enrollment (45.8% in this study and 43.4% in the meta-analysis) [107]. In addition to employing maximum likelihood estimation in all MLMs (which is robust to MAR) [83], we conducted a series of
sensitivity analyses to assess the effects of various MNAR assumptions. Most effects were robust to noncompleters having outcomes slightly worse than completers (SD 0.25). However, the effects did not persist with larger deviations (SD ≥0.50). It is impossible to directly test which of these scenarios is most likely (as is the case for other MNAR approaches) [113]. Future studies should include items specifically to predict missingness (eg, “how likely are you to drop out of this study?”) [114]. Responses can then be included as auxiliary variables to improve the performance of MAR methods (effectively converting MNAR to MAR) [114].

In light of the degree of attrition, secondary analyses should be interpreted as exploratory. However, these models provide tentative possibilities to be examined further. We found evidence that higher usage (median or above days of use) was associated with larger improvements in distress and several other outcomes. This mirrors the dosage-outcome associations seen in the in-person MBI literature [115]. Mediation analyses suggest candidate mechanisms theoretically linked to each HMP module that may indirectly contribute to decreased distress (ie, mindfulness, social connectedness, defusion, rumination). This also mirrors reviews of the in-person MBI literature that have found changes in mindfulness and rumination mediate effects [116]. The possibility that social connectedness also plays a role should be explored further, particularly as it has been associated with numerous psychological and physical health outcomes [117,118]. Moderation analyses indicated larger improvements among those showing higher rumination or empathy and lower defusion at baseline. These are somewhat conflicting findings, with the rumination and defusion associations suggesting that HMP may be most effective for those with deficits at baseline, whereas the association with empathy suggesting a higher baseline level may be necessary to benefit most. Given that mHealth interventions could, in theory, be easily adapted to participant characteristics (eg, participants routed to receive a particular version based on baseline questionnaires), future experimental work can specifically examine who is likely to benefit from which kind of training (eg, randomizing to adapted vs nonadapted versions). The scalability of mHealth RCTs may allow recruitment of the sample sizes necessary for adequately powered tests of moderation [119].

Limitations and Future Directions

As noted, high and differential attrition are limitations of this study. Although attrition is commonplace in MBI research [110], future studies should include study design features that decrease attrition (eg, telephone enrollment) [107]. Relatedly, engagement with the HMP app was relatively modest. Similar to attrition, low engagement is a rule rather than an exception for mHealth [120]. Presumably low engagement diminished any potential effects of the HMP app, making estimates of efficacy more ambiguous. Future studies could evaluate engagement strategies by randomizing participants to receive approaches found to improve adherence to medical regimens (eg, modifying dosage recommendations, providing reminders) [121]. Microrandomized trials could assess the impact of a variety of small manipulations on short-term engagement outcomes [122].

The lack of a follow-up assessment is another limitation, making it unclear whether any benefits were sustained. Furthermore, the lack of active control conditions makes it impossible to rule out effects due to a digital placebo [123]. Similarly, the included self-report measures are vulnerable to social desirability bias, although this may be less of an issue within a fully remote RCT [124]. Sampling procedures and sample demographics raise questions regarding generalizability, especially to racial/ethnic minority populations and those with lower levels of education. Participants in the Center for Healthy Minds database may have been particularly amenable to the HMP app (although those with prior meditation experience would have been excluded).

An obvious future study would assign participants to receive only Connection or Insight module content. This could clarify the unique effects of these types of practices. As we observed effects on distress that persisted when restricted to those with elevated symptoms at baseline, it would be worthwhile replicating this study within a clinical sample. For this, it could be important or even necessary, for safety reasons, to include some amount of professional guidance [41], perhaps telephone or text-based support [125]. Other extensions of this study could include the use of non–self-report measures, both to rule out social desirability as well as to clarify underlying mechanistic processes. Numerous biological and behavioral mechanisms have been assessed for in-person MBIs (eg, changes in attention, body awareness, stress physiology) [126-128] and may be operant within mHealth MBIs. A wide variety of extensions could also be made to the HMP app itself. For example, intervention components could be provided in response to passively assessed mood states (eg, using data streams routinely gathered through phone sensors). This would require not only the validation of passive measures [129,130] but also studies that clarify the optimal pairing of intervention components to mood. Microrandomized trials may be an ideal design for this purpose. The amount of engagement necessary for clinical benefits (ie, recommended dosage) should be clarified in future studies (eg, by random assignment to dosage conditions). RCTs using active control conditions can help clarify the role that nonspecific factors may play in mHealth MBIs. On the basis of the robust association between working alliances and outcomes within in-person interventions [131], future studies should assess its digital corollary [132]. Finally, a critical future direction is investigating the efficacy of mHealth MBIs specifically (and ideally tailored for) [133,134] racial/ethnic minorities. Racial/ethnic minority populations are at increased risk for racism-related negative psychological and physical health consequences [135] and have been historically underrepresented in research on mindfulness [136,137].

Conclusions

Research into the mobile delivery of meditation training has expanded rapidly in the past five years. However, the vast majority of this work has focused exclusively on training mindfulness. We found preliminary evidence that 2 versions of a novel smartphone app that includes training in mindfulness, in addition to skills specifically targeted to improve social connection and intrapersonal insight, are associated with reduced psychological distress, increased social connectedness, and improvements in candidate mechanisms believed to underlie...
Conflicts of Interest

RD is the founder, president, and serves on the board of directors for the nonprofit organization, Healthy Minds Innovations, Inc. CD is Chief Contemplative Officer for Healthy Minds Innovations, Inc. CW has served as a paid consultant and content contributor for Healthy Minds Innovations, Inc.

Multimedia Appendix 1

Healthy Minds Program randomized controlled trial supplemental materials.
[DOCX File, 113 KB - Multimedia Appendix 1]

Multimedia Appendix 2

CONSORT-eHealth (V 1.6.1).
[PDF File (Adobe PDF File), 3037 KB - Multimedia Appendix 2]

References


Abbreviations

- **FDR**: false discovery rate
- **HMP**: Healthy Minds Program
- **IRI**: Interpersonal Reactivity Index
- **MAR**: missing at random
- **MBCT**: mindfulness-based cognitive therapy
- **MBI**: meditation-based intervention
- **MBSR**: mindfulness-based stress reduction
- **mHealth**: mobile health
- **MLM**: multilevel model
- **MNAR**: missing not at random
- **PTQ**: Perseverative Thinking Questionnaire
- **RCT**: randomized controlled trial

Please cite as:

Testing the Efficacy of a Multicomponent, Self-Guided, Smartphone-Based Meditation App: Three-Armed Randomized Controlled Trial
JMIR Ment Health 2020;7(11):e23825
URL: http://mental.jmir.org/2020/11/e23825/
doi: 10.2196/23825
PMID: 33245288

©Simon B Goldberg, Theodore Imhoff-Smith, Daniel M Bolt, Christine D Wilson-Mendenhall, Cortland J Dahl, Richard J Davidson, Melissa A Rosenkranz. Originally published in JMIR Mental Health (http://mental.jmir.org), 27.11.2020. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.