Revealing subgroup-specific mechanisms of change via moderated mediation:
A meditation intervention example

Christian A. Webb, PhD1,2, Matthew J. Hirshberg, PhD3, Oscar Gonzalez, PhD4,
Richard J. Davidson, PhD3,5,6, Simon B. Goldberg, PhD3,7

1Harvard Medical School, Department of Psychiatry, Boston, MA
2McLean Hospital, Center for Depression, Anxiety & Stress Research, Belmont, MA
3Center for Healthy Minds, University of Wisconsin – Madison, Madison, WI, USA
4University of North Carolina at Chapel Hill, Department of Psychology, Chapel Hill, NC
5Department of Psychology, University of Wisconsin – Madison, Madison, WI, USA
6Department of Psychiatry, University of Wisconsin – Madison, Madison, WI, USA
7Department of Counseling Psychology, University of Wisconsin – Madison, Madison, WI, USA

Corresponding authors:
Christian A. Webb, Ph.D. Simon B. Goldberg, Ph.D.
Harvard Medical School & McLean Hospital University of Wisconsin
115 Mill Street, Belmont, MA 02478 335 Education Building, Madison, WI 53706
Email: cwebb@mclean.harvard.edu Email: sbgoldberg@wisc.edu

Funding: This research was supported by NCCIH Grant K23AT010879 (SBG), NIMH Grant R01MH43454 (RJD), Chan Zuckerberg Initiative Grant 2020-218037 (RJD), a National Academy of Education/Spencer Postdoctoral Fellowship (MJH), and with funding from the Wisconsin Center for Education Research (SBG). The first author (CAW) was partially supported by NIMH R01MH116969, NCCIH R01AT011002, the Tommy Fuss Fund and a Young Investigator Grant from the Brain & Behavior Research Foundation.

Disclosure: RJD is the founder, president, and serves on the board of directors for the nonprofit organization, Healthy Minds Innovations, Inc. MJH has been a paid consultant at Healthy Minds Innovations, Inc. for work unrelated to this research.
Abstract

Objective: Effective psychosocial interventions exist for numerous mental health conditions. However, despite decades of research, limited progress has been made on clarifying the mechanisms that account for their beneficial effects. We know that many treatments work, but we know relatively little about why they work. Mechanisms of change may be obscured due to prior research collapsing across heterogeneous subgroups of patients with differing underlying mechanisms of response. Studies identifying baseline individual characteristics that predict differential response (i.e., moderation) may inform research on why (i.e., mediation) a particular subgroup has better outcomes to an intervention via tests of moderated mediation.

Method: In a recent randomized controlled trial comparing a 4-week meditation app with a control condition in school system employees (N=662), we previously developed a “Personalized Advantage Index” (PAI) using baseline characteristics, which identified a subgroup of individuals who derived relatively greater benefit from meditation training. Here, we tested whether the effect of mindfulness acquisition in mediating group differences in outcome was moderated by PAI scores.

Results: A significant index of moderated mediation (IMM=1.22, 95% CI: 0.30, 2.33) revealed that the effect of mindfulness acquisition in mediating group differences in outcome was only significant among those individuals with PAI scores predicting relatively greater benefit from the meditation app.

Conclusions: Subgroups of individuals may differ meaningfully in the mechanisms that mediate their response to an intervention. Considering subgroup-specific mediators may accelerate progress on clarifying mechanisms of change underlying psychosocial interventions and may help inform which specific interventions are most beneficial for whom.
Keywords: moderation; mediation; moderated mediation; mechanism; meditation; mobile health

Public Health Significance Statement

Individuals receiving psychosocial interventions likely differ substantially in which intervention elements they respond to and in their causal pathways of change. This study demonstrates the utility of considering subgroup-specific mediators, which may accelerate progress on clarifying mechanisms of change underlying psychosocial interventions and may help inform which specific interventions are most beneficial for whom.
Introduction

An array of empirically supported psychosocial interventions (e.g., cognitive behavioral therapy, interpersonal therapy, behavioral activation) exist for a broad range of mental health conditions in children, adolescents, and adults (Barlow, 2021; Tolin et al., 2015). Despite decades of research, however, minimal progress has been made in identifying the mechanisms that account for why patients improve in these treatments (Cuijpers et al., 2019; Kazdin, 2009; Lorenzo-Luaces et al., 2015; Webb et al., 2010; Zilcha-Mano, 2021). After over half a century of research on the mechanisms of change and 37 years since the introduction of mediation analysis (Baron & Kenny, 1986), we know that many treatments work, yet we still know strikingly little about why or how they work.

Mechanisms of change may be obscured due to prior intervention research collapsing across heterogeneous subsamples of patients who in fact have differing underlying mechanisms of response (Hollon, 2019, 2020; Huibers et al., 2021; Webb, Murray, et al., 2022; Zilcha-Mano et al., 2021; Zilcha-Mano & Webb, 2021). In all likelihood, causal mechanisms of change differ across patients (Cuijpers et al., 2019). For example, individuals receiving cognitive behavioral therapy (CBT) for depression may vary substantially in the mechanisms that account for their symptom improvement. It may be that only a subset of these individuals respond due to the core cognitive or behavioral change strategies encouraged in CBT (Forand et al., 2017; Lorenzo-Luaces et al., 2015; Sasso et al., 2015; Webb et al., 2019). In contrast, other subgroups of patients may improve for other reasons, including other therapy components (e.g., problem solving skills, assertiveness), non-specific/common factors (e.g., the passage of time, regression to the mean, placebo-related expectancies, the therapeutic alliance) or extra-therapeutic causes (e.g., positive life events outside of therapy). As a result, the theory-specified mechanism of
change associated with CBT may only account for response in a subgroup of individuals.

Mindfulness meditation-based interventions are a family of evidence-based mental health promotion strategies that have surged in popularity in recent years (Creswell, 2017). For example, meditation apps are now the most commonly downloaded and widely used apps for depression and anxiety (Wasil et al., 2020). There is a large body of literature supporting the overall efficacy of mindfulness interventions in alleviating stress, anxiety, and depressive symptoms (Baminiwatta & Solangaarachchi, 2021; Galante et al., 2021; Goldberg et al., 2018; Goldberg, Riordan, et al., 2022; Khoury et al., 2015). However, the mechanisms that account for the benefits of mindfulness training remain elusive (Goldberg, 2022). There is preliminary evidence for a range mindfulness-specific psychological mechanisms (e.g., beneficial changes in mindful awareness, acceptance, self-compassion, reappraisal, rumination, and psychological flexibility)(Goldberg, 2022; Gu et al., 2015; van der Velden et al., 2015). It may be that only a subset of individuals in fact acquire and benefit from the core mindfulness skills taught in these programs; whereas others may benefit for other reasons, including non-specific/common factors (e.g., placebo-related expectancies, therapeutic alliance with a mindfulness instructor, supportive interpersonal context of group meditation practice)(Goldberg, 2022). In summary, individuals likely differ substantially in which intervention elements they respond to and in their causal pathways of change (Cuypers et al., 2019; Hollon, 2019). It is also important to note that there is growing recognition in the literature that for some individuals, meditation or mindfulness training be may associated with adverse experiences or worsening of symptoms (Goldberg, Lam,

---

1 Of relevance, there is also a growing body of research on mechanisms of change in digital interventions (including mobile-based interventions) for depression and anxiety, which provide early evidence for a range of plausible mediators including cognitive variables (e.g., perceived control, rumination, or interpretation bias) and other potential mediators (e.g., mindfulness, acceptance, and behavioral activation).
et al., 2022; Hirshberg, Goldberg, et al., 2022; Montero-Marin et al., 2022).

Precision medicine, which seeks to identify subgroups of individuals with superior outcomes to specific interventions, provides an opportunity for progress on identifying mechanisms of change (Cohen & DeRubeis, 2018; Hollon, 2019). Specifically, a subgroup of individuals found to preferentially benefit from a specific treatment may be particularly responsive to the theory-specified “active ingredients” of that intervention (Hollon, 2019). For example, DeRubeis et al. (2014) developed a predictive algorithm (“Personalized Advantage Index”; PAI), based on pre-treatment patient characteristics, which identified subgroups of individuals who derived greater benefit from CBT vs. an antidepressant medication. Subsequent studies have successfully adapted the PAI to predict differential response to CBT versus interpersonal therapy (Huibers et al., 2015), CBT versus psychodynamic therapy (Cohen et al., 2019; Schwartz et al., 2020), an antidepressant medication versus placebo (Webb et al., 2018), and a mindfulness app vs. a mood monitoring app (Webb et al., 2021; Webb, Swords, et al., 2022; also see, Webb, Hirshberg, et al., 2022). In contrast to studies identifying general “prognostic” (i.e., intervention non-specific) predictors of outcome (i.e., baseline patient characteristics that predict better or worse outcomes across interventions), the above studies using the PAI demonstrate that baseline characteristics can identify subgroups of individuals who show better (i.e., specific) response to one intervention relative to another. In statistical terms, these studies demonstrate that certain baseline characteristics moderate treatment group differences in outcome. Importantly, these moderation findings may also suggest differential mediation. That is, if a defined subgroup of patients exhibits significantly better outcomes to Treatment A than Treatment B, this suggests that these individuals are particularly responsive to the unique (i.e., not shared between treatments) ingredients of the former treatment. In summary,
studies identifying individual baseline characteristics that predict differential response (i.e., moderation) may inform research on why (i.e., mediation) a particular subgroup has better outcomes to a given intervention (Hollon, 2019). This can be evaluated empirically via tests of moderated mediation (Fairchild & MacKinnon, 2009; Hayes, 2022).

Recent PAI studies identifying subgroups of individuals who show significantly greater benefit from app-based mindfulness training relative to a control condition (Webb, Hirshberg, et al., 2022; Webb, Swords, et al., 2022) provide a unique opportunity to test differential mediation via moderated mediation. Specifically, it may be that the PAI is identifying a subgroup of individuals who are more likely to (1) acquire and/or (2) benefit from the specific mindfulness skills taught in these mindfulness apps. Put in moderated mediation terms, PAI scores may moderate: (1) between-group (i.e., meditation vs. control condition) differences in the acquisition of mindfulness skills (Path a in a mediation model) and/or (2) the relation between mindfulness skills and outcome (Path b). An important benefit of conducting moderated mediation analyses using data derived from clinical trials of app-delivered interventions is that relative to conventional in-person interventions, smartphone app-delivered interventions are highly scalable and make it quite feasible to collect a large enough sample for adequately powered analyses. Accordingly, using data from a recent randomized controlled trial comparing a 4-week meditation app with a waitlist control condition (N = 662), we used moderated mediation to test whether the effect of mindfulness acquisition in mediating group differences in outcome was moderated by PAI scores. The RCT from which these data were drawn was preregistered on clinicaltrials.gov ((https://clinicaltrials.gov/ct2/show/NCT04426318), and through the Open Science Framework (https://osf.io/eqgt7). However, the moderated mediation analyses reported here were not preregistered.
Methods

Participants and Procedure

Data were drawn from a recently completed RCT testing the effects of the Healthy Minds Program (HMP) (Dahl et al., 2020; Goldberg et al., 2020) app for 662 predominantly distressed (79.9% reporting clinically elevated depression and/or anxiety symptoms at baseline) school district employees during the early months of the COVID-19 pandemic (for additional study details and primary outcome results, see [omitted for blind review]). This manuscript presents secondary analysis from the parent RCT. The publication of the parent RCT (omitted for blind review) reported on group differences in improvement in psychological distress (pre-registered primary outcome). In contrast, here we tested a moderated mediation using the latter dataset. In this RCT, participants were randomized to receive the HMP app (https://hminnovations.org/meditation-app) immediately \((n = 344)\) or at the conclusion of the study (i.e., waitlist control; \(n = 318\)). The RCT had a target sample size of 400, estimated to provide 80% power to detect small-to-moderate between-group differences \((d \geq 0.38)\), assuming 43.4% attrition (Linardon & Fuller-Tyszkiewicz, 2020) and \(\alpha = .050\). This magnitude of effect is similar to that observed in recent meta-analyses of app-based meditation training (Gál et al., 2021). The preregistration noted that a larger sample size may be recruited if additional funding was secured.

To be eligible, participants had to report no or minimal prior meditation experience and depressive symptoms below the severe range on the Patient-Reported Outcomes Monitoring Information System [PROMIS] Depression scale (i.e., T-score \(\leq 70\)) (Pilkonis et al., 2011). Psychological distress was preregistered as the primary outcome for the RCT. Psychological distress was operationalized as the composite of PROMIS Depression (T-score), PROMIS (T-
score) Anxiety, and Perceived Stress Scale. Participants completed measures of psychological distress, secondary outcomes, and candidate mechanisms at baseline, weekly during the 4-week intervention period, and at a 3-month follow-up timepoint. The current study included all 662 participants. Sample demographics are reported in Supplemental Table 1.

Participants randomized to the HMP app condition were encouraged to use the app during the 4-week intervention period. They also maintained access to the app between post-test and follow-up timepoints. The HMP app includes training in four aspects of well-being: Awareness, Connection, Insight, and Purpose (ACIP) (Dahl et al., 2020). The Awareness module emphasizes attention regulation (e.g., focused attention; meta-awareness of internal experience). The Connection module trains capacities designed to support positive relations with oneself and others (e.g., gratitude, compassion). The Insight module includes practices designed to clarify the nature of self and internal experience (e.g., seeing thoughts as only thoughts). The Purpose module involves clarifying value and purpose and expressing these in daily life activities. HMP includes both didactic content discussing the science of well-being as well as guided meditation practices designed to support cultivation of ACIP-relevant skills. There were not human-supported or interactive components within the app. However, there is evidence that participants do experience some digital corollary to the therapeutic alliance even in an unguided context (Goldberg, Baldwin, et al., 2022; Henson et al., 2019). For additional intervention details, see Hirshberg, Frye et al. (2022). The study procedures were approved by the University of Wisconsin – Madison Institutional Review Board (number 2020-0533).

**Measures**

*Psychological Distress*
Consistent with our pre-registration, the computer adaptive PROMIS Depression and PROMIS Anxiety scales (v1.0) (Pilkonis et al., 2011) and the 10-item Perceived Stress Scale (PSS) (Roberti et al., 2006) were used to assess psychological distress. All three measures are widely used and have shown strong psychometric properties. The psychological distress composite was computed by taking the z-score of each measure (standardized to the baseline observation) and averaging across the three scales. This was done based on prior work showing high inter-correlations between these three measures (Goldberg et al., 2020) and evidence that psychopathology generally loads on a single factor (Caspi et al., 2014). Accordingly, internal consistency for the three-item composite scale was acceptable in the current sample ($\alpha = .87$).

The PROMIS Depression and PROMIS Anxiety scales have shown convergent validity with legacy measures (Choi et al., 2014; Schalet et al., 2014). Sample items include “I felt worthless” (depression) and “I felt fearful” (anxiety). Participants rate their symptoms in the past 7 days on a 5-point Likert-type scale ranging from 1 (never) to 5 (always). The computer adaptive PROMIS measures yield T-scores (i.e., mean = 50, $SD = 10$), with T-score $\geq 55$ indicating clinical elevations. Internal consistency cannot be computed for the computer adaptive versions. However, the fixed form versions of the PROMIS Depression and PROMIS Anxiety scales have shown adequate internal consistency reliability ($\alpha \geq .90$) (Pilkonis et al., 2011).

The PSS (S. Cohen & Williamson, 1988) assesses perceived stress in the past month. Sample items include, “How often have you felt that you were unable to control the important things in your life?” Participants rate their stress level on a 5-point Likert-type scale ranging from 1 (never) to 5 (very often). The 10-item version of the PSS has shown strong convergent and discriminant validity (Roberti et al., 2006). Internal consistency was adequate in the current sample ($\alpha = .85$).
**Mindful Awareness**

Mindful awareness was assessed with the Acting with Awareness subscale of the Five Facet Mindfulness Questionnaire (FFMQ; Baer et al., 2008). This widely used measure evaluates participants’ perceived ability to attend to present moment experience. Mindful awareness is a core purported mechanism within the HMP app (Dahl et al., 2020) and was a preregistered candidate mechanism in the RCT (Hirshberg, Frye et al., 2022). Sample items include, “I find it difficult to stay focused on what’s happening in the present” (reverse scored). Participants rate their agreement with each item on a 5-point Likert-type scale ranging from 1 (*never or very rarely true*) to 5 (*very often or always true*). Higher scores represent higher mindful awareness. The Acting with Awareness subscale has been shown to differentiate between experienced meditators and community adults (Baer et al., 2008) and to increase in response to mindfulness training (e.g., Goldberg et al., 2016; Quaglia et al., 2016). Internal consistency was adequate in the current study (*α* = .91). To limit participant burden and tie our measurement most closely to the model of well-being in the HMP app (as described above), we included only the Acting with Awareness subscale of the FFMQ in the parent RCT.

**PAI**

Several additional self-report measures were administered at baseline and used to estimate participants’ baseline PAI scores (see Supplement for details and Hirshberg, Frye et al., 2022). These included the 15-item Perseverative Thinking Questionnaire (Ehring et al., 2011; *α* = .95) which assesses rumination and worry; the 5-item World Health Organization (Topp et al., 2015; *α* = .85) which assesses global well-being; the 5-item National Institutes of Health Toolbox Loneliness Questionnaire (Cyranowski et al., 2013; *α* = .90) which assesses perceived social disconnection; the 12-item Self-Compassion Scale Short Form (Raes et al., 2011; *α* = .86)
which assesses feelings of kindness toward oneself during difficult experience; the Drexel Defusion Questionnaire (Forman et al., 2012; $\alpha = .84$) which assesses the ability to maintain healthy psychological distance from internal experiences; the 10-item Meaning in Life Questionnaire (Steger et al., 2006) which assess presence ($\alpha = .91$) and search for meaning ($\alpha = .93$). Demographic variables assessed at baseline and included in the PAI model were age, gender identity, race, marital status, and income.

**Analytic Approach**

Moderated mediation is present when the effect of X (e.g., treatment condition) on Y (e.g., change on primary outcome) through M (mediator) is conditional (i.e., is moderated by another variable) (Fairchild & MacKinnon, 2009; Hayes, 2022). In the present study, we tested whether the role of change in mindfulness during the intervention period in mediating between-group (i.e., meditation vs. control) differences in outcome (3-month post-intervention follow-up distress scores residualized on pre-treatment distress) was moderated by baseline PAI scores. To characterize change in mindfulness, we fit multilevel models assessing change in FFMQ Acting with Awareness during the 4-week intervention period. Models were constructed using the ‘lme4’ package (Bates et al., 2015) in R (R Core Team, 2022). Person-specific random slopes were extracted from these models as measures of change in mindful awareness. To evaluate moderated mediation, we tested whether individuals with more negative PAI scores (i.e., predicting better outcomes [reduced distress] in response to meditation training) may be more likely to acquire and benefit from the mindfulness skills taught in the meditation training. As illustrated in Figure 1, PAI scores may moderate: (1) between-group (i.e., meditation vs. control condition) differences in the acquisition of mindfulness skills (Path $a$) and/or (2) the relation between mindfulness skills and outcome (Path $b$). We tested whether PAI scores (mean-
centered) moderate these two effects using a moderated mediation model (PROCESS [version 4.1.1] macro in R) which yields an index of moderated mediation with Monte Carlo bootstrapped 95% confidence intervals (CIs) (Hayes, 2022). A separate PROCESS model was run for the moderation of Path $a$ and Path $b$ (Models 8 and 15 in PROCESS, respectively) given that when both are included simultaneously (and the moderator is a continuous variable) the indirect effect is a nonlinear function of the moderator and an index of moderated mediation cannot be computed (Hayes, 2022; Hayes & Rockwood, 2020). For details of these models, see Hayes (2022).

Given that the relationship between the mediator (M) and outcome (Y) may vary across groups (X), the mediation literature has increasingly emphasized the importance of testing the interaction between X and M as a predictor of outcome prior to preceding with a test mediation (Gonzalez & Valente, 2022; Hesser, 2022; Kraemer et al., 2008; MacKinnon, 2008; MacKinnon et al., 2020; Morgan-Lopez & MacKinnon, 2006; Rijnhart et al., 2021). For example, mindfulness skills are only taught in the meditation condition and not in the control group. Thus, the relation between mindfulness skills and outcome may be attenuated (or non-existent) in the control condition due to restricted variance (i.e., little change in mindfulness scores) or due to the fact that any observed increases in mindfulness scores among control participants may not actually reflect “true” acquisition of mindfulness (e.g., instead may be due to the influence of study demand characteristics, a consequence of reductions in distress)(Baer et al., 2019; Goldberg et al., 2019). Thus, for these reasons (also see MacKinnon et al., 2020), prior to running our moderated mediation analyses, we first tested a group x change in mindfulness scores.

---

2 The XM interaction provides a link between the traditional mediation analysis and newer causal mediation methods (For a detailed discussion, see MacKinnon et al., 2020).
interaction in predicting change in distress (via the ‘test.TMint’ in the Mediation package in R) (Tingley et al., 2014). If significant, the interaction would be incorporated into the equation predicting change in distress. If not significant, the group term would be omitted from this equation. Primary analyses focused on a completer dataset (i.e., only retaining subjects without missing data on the included variables) \( n = 566 \), but sensitivity analyses were conducted on a dataset in which missing data were imputed \( n = 662 \) (see Supplement for random forest-based multiple imputation procedure). R code and data are available through the Open Science Framework (OSF): https://osf.io/xr8z3/?view_only=9101030ca2d147e3a29458513605d89c.

Results

Sample Demographics

Mean age was 42.58 years old (SD = 10.67) and 87.98% of participants reported their gender as female. Racial/ethnic composition was as follows: 89.36% White, 4.3% Hispanic, 4.07% Black, 2.03% Asian/Pacific Islander, and 1.0% American Indian/Native Alaskan. See Supplemental Table 1 for detailed demographic characteristics by group. The sample was more female than educators statewide, but representative in terms of race (Wisconsin Department of Public Instruction, 2023).

Moderated Mediation

The group x mindfulness interaction in predicting change in distress was not significant \( p = .360 \) and thus was not included in the below moderated mediation analysis.

To formally test for moderated mediation, we first specified PAI as a moderator of Path \( a \). The group x PAI interaction was significant in predicting changes in mindfulness \( b = -3.81, t(562) = -2.80, p = .005 \); see Table 1 and Figure 2). To help interpret this effect, we decomposed this interaction by testing the relation between PAI scores and change in mindfulness for each
group separately. In the HMP group (blue line in Figure 2), decreasing PAI scores (i.e., reflecting relatively stronger recommendations for the HMP app) were associated with greater increases in mindfulness ($b = -1.97$, $t(278) = -2.03$, $p = .043$). Conversely, in the control group (red line in Figure 2), there was a nonsignificant trend for decreasing PAI scores being associated with less acquisition of mindfulness ($b = 1.85$, $t(284) = 1.93$, $p = .055$).

The 95% bias-corrected bootstrap CI for the index of moderated mediation (IMM = 1.22, 95% CI: 0.30, 2.33) did not contain zero, indicating significant moderated mediation. As seen in Table 2, the role of changes in mindfulness in mediating between-group (i.e., meditation vs. control) differences in outcome was moderated by PAI scores. Namely, mediation was only observed at lower PAI scores and not present at higher PAI levels (i.e., confidence intervals for the indirect effect include 0 at a mean-centered PAI score > 0.01). When converted back to the original (i.e., not centered) PAI scoring, the confidence intervals for the indirect effect include 0 at a baseline PAI (raw) score > -0.49 (or .44 SD above the PAI mean).

Next, we reran the model with PAI moderating Path $b$. This model did not reveal moderated mediation given that the bias-corrected bootstrap CI included zero (IMM = 0.20, 95% CI: -0.48, 0.92).

The above moderated mediation analyses were re-run on a dataset in which missing data were imputed. This resulted in the same pattern of findings (see Supplement).

**Discussion**

Identifying mechanisms that account for the benefits of psychosocial interventions remains elusive. One possible reason for the lack of progress is that prior intervention research collapses across heterogeneous samples of patients with different underlying mechanisms of response. Subgroups of individuals found to preferentially benefit from a specific treatment
(relative to a control condition) may be particularly responsive to the theory-specified “active ingredients” of that intervention and thus help reveal subgroup-specific mediating pathways (Hollon, 2019). In a sample of 662 participants randomly assigned to a meditation app (HMP) vs. an assessment-only control condition, we tested whether the effect of mindfulness acquisition in mediating group differences in outcome was moderated by a “Personalized Advantage Index” (PAI) (Webb, Hirshberg et al., 2022), which identified a subgroup of individuals who preferentially benefitted from the HMP app.

As displayed in Figure 1, a moderator variable (in this case, baseline PAI scores) may moderate either Path $a$ (group differences in changes in the mediator) and/or Path $b$ (the relation between the mediator and outcome) in a mediation model. Results revealed that PAI scores significantly moderated Path $a$, but not Path $b$. Specifically, as displayed in Figure 2, decreasing baseline PAI scores (i.e., reflecting relatively stronger recommendations for the HMP app) were associated with larger group differences in the acquisition of mindfulness, favoring HMP. As previously reported (Webb, Hirshberg et al., 2022), individuals with more negative PAI scores (i.e., further left on the x-axis in Figure 2) are characterized by higher baseline levels of repetitive negative thinking (on the PTQ measure). Thus, individuals randomly assigned to the HMP app who reported greater baseline perseverative negative thinking were more likely to exhibit increases in mindfulness over time relative to those with lower levels of repetitive negative thinking (see the blue line in Figure 2). Importantly, this pattern was not observed in the control condition (red line in Figure 2). One possible interpretation of this pattern of findings is that participants in the HMP group with greater baseline perseverative negative thinking (i.e., lower PAI scores) may have had more room for improvement in mindfulness skills focused on cultivating attention to the present moment (in contrast to negative perseverative thoughts which
are often focused on the past [e.g., rumination] or future [e.g., worries] (Nolen-Hoeksema et al., 2008; Webb, Tierney, et al., 2022)). Thus, exposure to mindfulness training through HMP may have allowed those predicted to be most responsive to acquire beneficial mindfulness skills. With regards to the pattern in the control group, it is not entirely clear why the association was in the opposite direction (i.e., decreasing PAI score associated with less acquisition of mindfulness, albeit at the level of a nonsignificant trend, \( p < .10 \)), as opposed to their simply being no relation between these two variables. It may be that individuals with greater baseline perseverative negative thinking (and who struggle with associated internalizing symptoms) may be less likely to show overall improvement (including in mindfulness awareness) over time when they receive no intervention (i.e., randomized to waitlist control) relative to those with lower levels of repetitive negative thinking. Indeed, although some patients may improve in wait-list control conditions (e.g., due to spontaneous remission, or positive expectancies related to being enrolled in treatment after the waitlist period), others may not or even deteriorate (Mekonen et al., 2022; Rutherford et al., 2012).

The moderated mediation analysis indicated that the effect of increases in mindfulness in mediating group differences in outcome (i.e., reduction in distress) was conditional (i.e., moderated) on PAI scores. Increases in mindfulness significantly mediated group differences in distress only among the subgroup of individuals with PAI scores approximately half (0.44) a standard deviation above the mean and below (i.e., those participants falling to the left of the dotted vertical red line in Figure 2, representing 68% of the sample). Conversely, for individuals with PAI scores above this point (representing 32% of the sample), increases in mindfulness scores did not significantly mediate group differences in outcome. It is important to note the latter subgroup did not simply consist of participants who failed to show positive outcomes in
response to HMP. There were no significant differences in HMP outcomes between the two groups (i.e., the subgroup for whom the indirect (mediation) effect was statistically significant vs. the subgroup for whom it was not significant; $p = .070$). The percentage of participants with reductions in distress to HMP in these two subgroups were 61% and 62%, respectively. Thus, there may be other unmeasured variables (e.g., common/non-specific factors) mediating improvement in the latter group.

There are both research and potential clinical implications for these findings. First, results suggest that there may be subgroups with different underlying mechanisms of response to the same intervention. Considering individual baseline characteristics that moderate indirect (mediation) effects may be one fruitful strategy to accelerate progress on clarifying the causal mechanisms of symptom change for psychosocial interventions. Focusing on individual characteristics that are known to predict outcomes differentially between groups is likely a promising initial place to look for candidate moderators of mediational effects. The PAI may be particularly promising for this purpose, given it can aggregate information from a large number of baseline characteristics. In addition, the lack of moderation of Path $b$ suggests that changes in mindfulness skills are equally linked with long-term reductions in distress, irrespective of how one is predicted to respond to the intervention. Thus, the responsiveness captured by the PAI seems to be more about the likelihood of acquiring relevant skills rather than likelihood that these skills are associated with beneficial effects on distress.

Clinically, this line of research may also help inform which specific interventions are most likely to benefit different subgroups of individuals, identified by their baseline characteristics. For example, if replicated, the current pattern of findings suggests that training in
mindful awareness may be particularly beneficial for those individuals with relatively high levels of repetitive negative thinking since they seem to be more likely to acquire mindfulness skills.

In terms of statistical approaches, we considered two additional modeling approaches for our moderated mediation models. First, we fit the linear regression models using structural equation modeling (SEM), and we obtained similar results to what we obtained using PROCESS. Given the indices of moderated mediation are provided by PROCESS and not in standard SEM software, we decided to present the results from PROCESS. Furthermore, we considered modeling the repeated measures of the mediator using latent growth curves and include them in the mediation model (Cheong et al., 2003). However, it was unclear to us what is the best approach to accommodate the interaction between the slope factor (which is a latent variable) and PAI (Gonzalez & Valente, 2022), which is a future direction of this research.

The study had several limitations. First, we relied on an assessment-only control group, which controlled for the effect of time (regression to the mean, history) and repeated assessments of our mediator (mindfulness) and outcome (distress), but did not control for treatment non-specific or common factors shared between bona fide interventions (e.g., expectations of symptom improvement, working alliance)(Goldberg, Baldwin, et al., 2022). A future study could consider including an active comparison condition (e.g., CBT app, behavioral activation app, or compare two apps which provide training in different forms of meditation such as those cultivating focused attention on the breath, open monitoring, or loving-kindness practices). For example, a study could randomly assign participants to a mindfulness app vs. a CBT app to test (1) whether a PAI approach – informed by baseline individual characteristics – can identify a subgroup of individuals who derive significantly greater benefit from the mindfulness app relative to the CBT app (and vice versa) and (2) whether subgroup-specific mediators can be
identified via moderated mediation. Second, we only included a single (theory-informed) mediator. Multiple mediation models (Hayes, 2017; Hayes & Rockwood, 2017) exist which can accommodate more than one mediator simultaneously. To return to the above example, a study comparing a CBT app vs. a mindfulness app could simultaneously include multiple mediators relevant to each intervention (e.g., measures of cognitive and behavioral change vs. changes in different facets of mindfulness such as mindful awareness of the present and “attitudinal” facets of mindfulness such as non-judgment and curiosity towards internal experience). We only included one facet/subscale of the FFMQ and future studies should examine the other facets (e.g., simultaneously in a multiple mediation model). It will also be important for future studies to move beyond a sole reliance on self-report measures of putative mediators and consider meditation-relevant behavioral (e.g., attentional control tasks) and/or biological measures (e.g., dynamic resting state functional connectivity in attentional networks) (Lim et al., 2018; MacLean et al., 2010). Third, mindfulness and distress were assessed every week during the intervention phase, along with a 3-month follow-up. A denser assessment schedule (e.g., repeated, daily smartphone-delivered ecological momentary assessment surveys) would provide a more fine-grained assessment of change in different facets of mindfulness and outcome, and their temporal interrelationship (Webb, Forgeard, et al., 2021; Webb, Swords, et al., 2022). It is possible that different patterns of association would appear when examining a finer-grain time scale. A denser assessment would also allow to examine the time course of change (e.g., identifying rapid responders). Fourth, our sample was predominantly non-Latinx White. Finally, studies are needed to test whether a similar pattern of findings emerge in other meditation training contexts, such as conventional group-based mindfulness interventions or more intensive training settings (Zanesco et al., 2021).
These limitations notwithstanding, the current study illustrates the possibility that tests of moderated mediation can be used to identify candidate subgroup-specific mechanisms within interventions. Future studies including active controls and measures of competing intervention-specific and non-specific factors may help clarify what works for whom and why.
References:


Self-Guided, Smartphone-Based Meditation App: Three-Armed Randomized Controlled Trial. *JMIR Mental Health*, 7(11), e23825. https://doi.org/10.2196/23825


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

Webb, C. A. (2023, January 5). Supplemental materials for preprint: Revealing subgroup-
specific mechanisms of change via moderated mediation: A meditation intervention
example. Retrieved from osf.io/xr8z3

Meditation training modulates brain electric microstates and felt states of awareness.

Zilcha-Mano, S. (2021). Toward personalized psychotherapy: The importance of the trait-
like/state-like distinction for understanding therapeutic change. American Psychologist,
76, 516–528. https://doi.org/10.1037/amp0000629

(2021). Investigating patient-specific mechanisms of change in SET vs. EFT for
depression: Study protocol for a mechanistic randomized controlled trial. BMC

Like Within-Individual Effects in Studying the Mechanisms of Change in CBT. Frontiers
Table 1

Bootstrap results for regression model parameters

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>95% Bootstrap Confidence Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-0.071</td>
<td>0.032</td>
<td>-0.135 – -0.008</td>
</tr>
<tr>
<td>Group</td>
<td>0.154</td>
<td>0.046</td>
<td>0.065 – 0.245</td>
</tr>
<tr>
<td>PAI</td>
<td>1.849</td>
<td>1.038</td>
<td>-0.159 – 3.905</td>
</tr>
<tr>
<td>Group x PAI</td>
<td>-3.81</td>
<td>1.434</td>
<td>-6.632 – -0.993</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>95% Bootstrap Confidence Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.126</td>
<td>0.044</td>
<td>0.041 – 0.211</td>
</tr>
<tr>
<td>Group</td>
<td>-0.253</td>
<td>0.058</td>
<td>-0.366 – -0.138</td>
</tr>
<tr>
<td>Change in Mindfulness</td>
<td>-0.319</td>
<td>0.065</td>
<td>-0.445 – -0.186</td>
</tr>
<tr>
<td>PAI</td>
<td>-2.556</td>
<td>1.420</td>
<td>-5.442 – 0.196</td>
</tr>
<tr>
<td>Group x PAI</td>
<td>1.634</td>
<td>1.689</td>
<td>-1.681 – 5.020</td>
</tr>
</tbody>
</table>

Note. PAI = Personalized Advantage Index. Number of bootstraps for confidence intervals = 10,000 (bold values are statistically significant as the 95% confidence interval does not include 0).
Table 2. Conditional Indirect Effect of Changes in Mindfulness Mediating Group Differences in Outcome for Various Baseline Personalized Advantage Index (PAI) Values.

<table>
<thead>
<tr>
<th>Baseline PAI Value</th>
<th>Indirect Effect</th>
<th>Standard Error</th>
<th>95% Bootstrap Confidence Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.04</td>
<td>-0.098</td>
<td>0.031</td>
<td>-0.163 – -0.044</td>
</tr>
<tr>
<td>-0.03</td>
<td>-0.086</td>
<td>0.027</td>
<td>-0.143 – -0.039</td>
</tr>
<tr>
<td>-0.02</td>
<td>-0.073</td>
<td>0.023</td>
<td>-0.122 – -0.033</td>
</tr>
<tr>
<td>-0.01</td>
<td>-0.061</td>
<td>0.020</td>
<td>-0.103 – -0.026</td>
</tr>
<tr>
<td>0</td>
<td>-0.049</td>
<td>0.018</td>
<td>-0.087 – -0.018</td>
</tr>
<tr>
<td>0.01</td>
<td>-0.039</td>
<td>0.017</td>
<td>-0.074 – -0.007</td>
</tr>
<tr>
<td>0.02</td>
<td>-0.025</td>
<td>0.018</td>
<td>-0.062 – 0.008</td>
</tr>
<tr>
<td>0.03</td>
<td>-0.013</td>
<td>0.020</td>
<td>-0.054 – 0.026</td>
</tr>
<tr>
<td>0.04</td>
<td>-0.00</td>
<td>0.023</td>
<td>-0.047 – 0.046</td>
</tr>
</tbody>
</table>

Note. PAI = Personalized Advantage Index. Number of bootstraps for confidence intervals = 10,000 (bold values are statistically significant as the 95% confidence interval does not include 0).
**Figure Captions**

**Figure 1.** Moderated mediation model. Baseline Personalized Advantage Index (PAI) scores may moderate (1) between-group (i.e., meditation vs. control condition) differences in the acquisition of mindfulness skills (Path $a$) and/or (2) the relation between mindfulness skills and outcome (Path $b$).

**Figure 2.** Group x Personalized Advantage Index (PAI) interaction in predicting change in mindfulness over the 4-week intervention period. Decreasing PAI scores (i.e., reflecting relatively stronger recommendations for the Healthy Minds Program [HMP] app) are associated with greater increases (i.e., positive slope) in mindfulness in the HMP (but not the control) group. The dotted vertical red line represents the value of the moderator (PAI) at which the indirect (mediation) effect becomes significant (i.e., increases in mindfulness significantly mediate group differences in outcome among individuals to the left of the dotted red line).
Data Transparency: The original pre-registered randomized clinical trial (RCT) reporting group differences in overall outcomes is published Hirshberg et al. (2022). In addition, a previous study tested baseline predictors of symptom change in this same trial (Webb et al., 2022). Critically, neither of these studies tested moderated mediation, which is the focus of the present submission.