



# The effects of app-based mindfulness practice on the well-being of university students and staff

Oskari Lahtinen<sup>1</sup> · Jenni Aaltonen<sup>1</sup> · Johanna Kaakinen<sup>1</sup> · Lena Franklin<sup>2</sup> · Jukka Hyönä<sup>1</sup>

Accepted: 19 April 2021

© The Author(s) 2021

## Abstract

Mental health problems like anxiety, depression, and stress have been increasing in many countries and the 2020 COVID-19 pandemic has further exacerbated their toll. Mindfulness-based interventions have been shown to provide evidence-based treatments for anxiety and depression, and accumulating evidence is emerging in support of using mindfulness apps yielding small-to-moderate treatment effects. The study was a 4-week randomized controlled trial with 561 university students and staff as participants, divided into a treatment group (mindfulness app) and an active control group (psychoeducational online content). Depression, anxiety, and stress were evaluated as primary study outcomes. Saliva cortisol samples were also collected from a subgroup of the treatment arm ( $n = 29$ ). Using the mindfulness app for four weeks resulted in small reductions in stress ( $d = .16$ ), and depression ( $d = .16$ ). Attrition was 28.0%. Subjects who practiced more did not experience additional improvement in wellbeing. Mindfulness apps offer modest but clear benefits to users in terms of improved mental health. They present a promising supplement to traditional mental health services.

**Keywords** Mindfulness · Mindfulness-app · Anxiety · Depression · University staff · Students

Before the COVID-19 pandemic in 2020, around a third of university students reported suffering from mental health problems (Finnish sample; Kunttu et al., 2016). Instead of seeking help, students often self-medicate with alcohol (39%), cigarettes (36%), or drugs (15%; Pierceall & Keim, 2007). At the time of writing, the 2020 COVID-19 pandemic has caused around the world common mental health problems like anxiety, depression, and stress to skyrocket. More than 30% of Americans have reported symptoms of generalized anxiety disorder (peaking in the youngest age group, 18–29-year-olds with 40+ % prevalence). The numbers for depression are nearly as dire with 23.5% for the entire US population and 32.7% for 18–29-year-olds (CDC, 2020). The situation presents a substantial stress test to mental health services around the world. As a result, cost-effective and accessible mental health solutions are in high demand.

Mindfulness meditation has been used to alleviate anxiety, depression, and stress. Meta-analyses support its efficacy for anxiety and depression, while the jury is still partially out on stress (Goleman & Davidson, 2017; Goyal et al., 2014). Strongest evidence is for the efficacy of Mindfulness-Based Cognitive Therapy in reducing the risk of depressive relapse (Kuyken et al., 2015; Teasdale et al., 2000). Another well-researched, and moderately effective, mindfulness-based intervention (MBI) is Mindfulness-Based Stress Reduction (MBSR; de Vibe et al., 2017; Kabat-Zinn, 1990).

Mindfulness meditation is practiced by focusing attention on present moment experience in a nonjudgmental way. Practice usually begins with paying attention to body sensations like the breath. When the mind wanders to thoughts, attention is brought back to the erstwhile focus. Once competence builds, practice can open up to noticing anything in experience and maintaining attention on it. Practice time in hours correlates with how much participants benefit from it (Goleman & Davidson, 2017; Parsons et al., 2017).

With the increased popularity of smartphones, we have experienced a significant upsurge of smartphone applications geared towards enhancing subjective wellbeing and mental health. They are a promising means to enhance human wellbeing, as they are readily available for use by anyone anytime and anywhere. The 2010s also saw a spike in

---

✉ Oskari Lahtinen  
polaht@utu.fi

<sup>1</sup> Department of Psychology and Speech-Language Pathology, University of Turku, FI-20014 Turku, Finland

<sup>2</sup> Atlanta, GA, USA

studies examining the effectiveness of smartphone applications in promoting mental health. Unfortunately, however, many of them were found to be ineffective or even harmful. In their recent systematic review, Wang et al. (2018) assessed 100 smartphone applications designed to promote mental health. Only 14 of them were deemed to demonstrate evidence-based effectiveness.

In addition to general mental health apps, the previous decade also saw an explosion in the availability of mindfulness apps in particular. A 2015 survey of mindfulness apps concluded that most did not in fact offer mindfulness and only one out of hundreds was supported by research evidence (Mani et al., 2015). After 2015, studies evaluating apps began to crop up. Mindfulness apps usually offer 10 or 20 min daily guided meditations and these can target specific ailments like stress, anxiety, or depression. Many randomized controlled trials (RCTs) now support the efficacy of Headspace in alleviating, e.g., anxiety, and depression (Bennike et al., 2017; Bostock et al., 2019; Economides et al., 2018; Howells et al., 2016; Noone & Hogan, 2018; Rosen et al., 2018; Yang et al., 2018). Other evidence exists for benefits of, e.g., apps Calm (stress: Huberty et al., 2019; wellbeing: Clarke & Draper, 2020), and Smiling Mind (depressive symptom improvement, resilience; Flett et al., 2019). Apps have been studied in many different populations including physicians (Roy et al., 2020; Wen et al., 2017), academic advisors (Hendricks et al., 2019), and university students (Lyzwinski et al., 2019), with mostly encouraging results. The latter two groups are targets in the present study as well. No meta-analyses have combined effects sizes from mindfulness app studies, though one evaluated online (including website-based) mindfulness-based programs and found salutary main-effects for depression ( $g = .29$ ), anxiety, ( $g = .22$ ), and stress ( $g = .51$ ; Spijkerman et al., 2016).

Mindfulness research has been under criticism for suboptimal methodology of late, and the field has been asked to step up its game (Davidson & Dahl, 2018; Van Dam et al., 2018). Even though several RCTs of mindfulness apps have recently been published, very few studies still use an active control group, the gold standard in MBI research (Goleman & Davidson, 2017). Attrition rates have not been evaluated for mindfulness apps specifically, but a meta-analysis reported a 50% rate for all apps for depressive symptoms collectively (Torous et al., 2020). Sample sizes have also tended to be quite small in studies in the field (most report  $n < 100$ ). An issue in the design of many studies has also been that even though the amount of practice (in minutes) is crucial information, few studies record it. Very few studies have been preregistered to prevent publication bias (Coronado-Montoya et al., 2016). Finally, outcome measures in mindfulness research outside of neuroimaging studies have often tended to be questionnaire-based, which are prone to bias. Biological outcomes would give more accurate information on what is actually happening within practitioners.

The present study is a pre-registered RCT, in which we evaluated a mindfulness app, Welzen, against a specifically tailored active control treatment. The app offered participants guided meditations designed to alleviate symptoms of anxiety, depression, and stress. We recruited a large sample ( $n = 561$ ) from a Finnish university with both students and staff included. A subsample of the treatment arm ( $n = 30$ ) gave salivary cortisol samples pre- and post-study. Our primary hypothesis was that the app would offer small-to-moderate reductions in anxiety and depression, and perhaps in stress, in the treatment arm of the RCT. We also hypothesized that practice minutes would correlate with the amount of benefit participants derive from practice. We expected cortisol levels to decline in saliva samples from pre- to post-measurement. The study was preregistered at Open Science Framework (Hyönä et al., 2018). We collected two waves of data on the RCT design. We also collected exploratory data which are mentioned in the preregistration, but not reported in the study as there was no control group for these data.

## Method

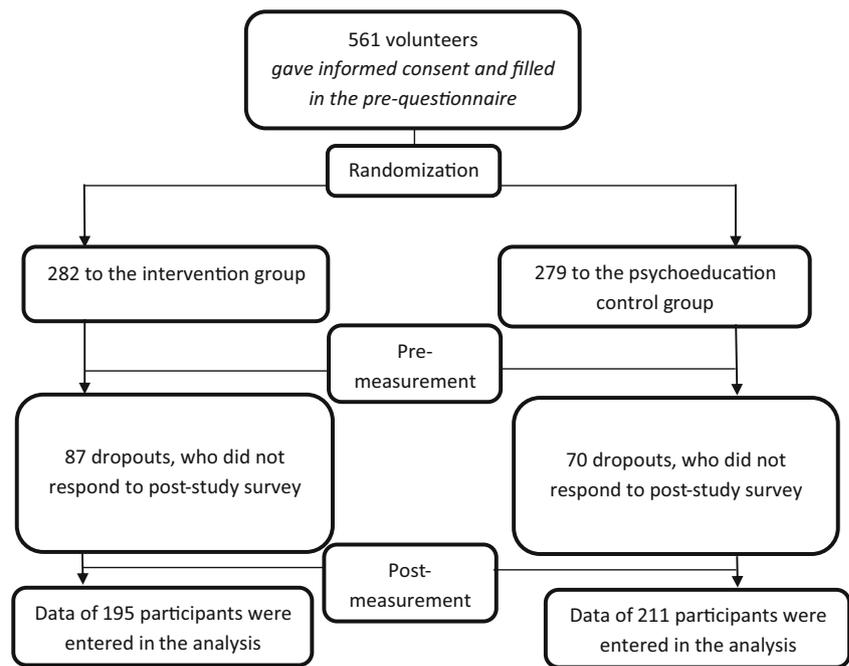
### Participants

Participants were recruited from among the faculty, staff, and students of the University of Turku. The University Wellbeing Services helped in the recruitment. The study was advertised via the university email lists, in the intranet, and with fliers distributed in the campus area. Every volunteer was accepted to participate provided they (1) owned a smartphone, (2) had sufficient English skills (the app was in English), (3) had not practiced mindfulness on a regular basis, and (4) was committed to regular practice for four weeks. Volunteers with a psychiatric diagnosis were instructed to consult their therapist/doctor about the suitability for participation. If permitted by the therapist, they were allowed to take part in the study. Thirty-three participants in the intervention group and 32 in the active control group had a psychiatric diagnosis. A total of 561 people volunteered for the study. They were randomly assigned to the intervention and the psychoeducation control group using the [randomlists.com](https://randomlists.com) website. A graphic depiction of the recruitment process is found in Fig. 1 that also includes the drop-out rate.

### Procedure

The participants volunteered to take part in the study by sending an email to the 2nd author who sent the participant a link to the pre-questionnaires to be filled in. The participants also provided their email address to which a link to the smart phone app could be sent. Moreover, they were asked about their sex, nationality, level and type of education,

**Fig. 1** Process chart from recruitment to post-measurement



current academic status (undergraduate or post-graduate student, staff member, faculty member), previous experience with mindfulness meditation, rated interest in mindfulness meditation, and possible diagnosis of a psychiatric disorder and its treatment.

After participants had filled in the pre-study questionnaires, 9 test tubes together with the instructions for taking the samples and a return envelope were mailed to the cortisol group ( $n = 30$ ). They were provided with another 9 test tubes and a return envelope after the intervention was complete.

The members of the intervention group received an email where they were supplied a link to the daily guided meditations. They also received a one-year premium membership for the app, which meant that they had free access to all its meditation practices. The members of the psychoeducation control group were sent a link to the podcasts appearing in SoundCloud they were instructed to listen to (one recording per day, 6 days a week). They were told that after one month they will be given access to all the contents in the app.

All participants received an email on four consecutive Fridays during the intervention that encouraged them to continue with the program. Immediately after the intervention was completed, all participants received via email a link to the same four questionnaires they had filled in prior to the intervention. The members of the control group received a premium membership that gave them access to all the contents in the app. All participants were encouraged to continue doing practices they found most beneficial. Two months after the completion of the intervention all participants were sent via email a link to the same four questionnaires to be filled in for the third time.

**Intervention** The mindfulness-based intervention was carried out using a smart phone app developed by Welzen (Atlanta, Georgia, USA) specifically for this study. The 7-day mindfulness program was targeted to alleviate symptoms of psychological stress, anxiety and depression as well as increase mindfulness skills. The intervention group repeated the 7-day program four times by carrying out one practice each day. Thus, the intervention lasted for 28 days. Each guided meditation developed and performed by Lena Franklin lasted for approximately 10 min. The guided meditations trained the participants to recognize bodily signs of psychological stress, to focus attention to breathing, to enhance patience, to find balance between external requirements and the inner self, to calm down and relax via breathing, and to strengthen self-compassion via a loving-kindness meditation.

The control group listened to 2–5-min mini-lectures on mindfulness practice. There were 12 lectures the control group listened to twice 6 days a week, repeating the protocol every week for four weeks. Thus, the psychoeducation program also lasted for 28 days. The lectures provided information on mindfulness practice regarding stress (three lectures), anxiety (3 lectures), paying attention (two lectures), and social relations (three lectures).

## Measures

**Perceived Stress** Perceived stress was measured by the Perceived Stress Scale (PSS; Cohen et al., 1983). We translated it into Finnish, after which we asked a professional translator to translate it back to English. The back-translation was then compared to the original questionnaire. As a result, a few

small revisions were made on the Finnish translation. PSS is a 10-item questionnaire that assesses how often a respondent has felt or thought in a certain way during the last month. A 5-point scale is used: never, almost never, sometimes, fairly often, very often. Every item is credited 0–4 points, so the range of the total score is 0–40. Higher scores denote increased perceived stress. The test has good reliability and validity (Cohen et al., 1983). In the present sample, the reliability was also estimated to be good (Cronbach's  $\alpha = .82$ ).

We also employed a single-item stress question developed by the Finnish Institute Occupational Health (Elo et al., 2003) and shown to be a reliable indicator of perceived stress. The English translation of this question reads as follows: "Stress refers to a situation where a person feels tense, restless, nervous or anxious, or she or he has difficulty sleeping, because bothering issues upsets the mind. Do you currently feel this kind of stress?"

**Anxiety** Anxiety was measured by the Generalized Anxiety Disorder (GAD-7) questionnaire (Spitzer et al., 2006) translated into Finnish by the Finnish Institute Occupational Health (<http://www.thl.fi/toimia/tietokanta/mittariversio/109/>). The respondent answers to seven questions concerning her/his feelings of worry, anxiousness, irritability, nervousness, restlessness, and fear during the last two weeks by using a 4-point scale: not at all sure, several days, over half the days, nearly every day. Each item is credited 0–3 points so that the range of the summed score is 0–21, bigger numbers denoting increased anxiety. The validity and reliability of the scale has been estimated to be satisfactory (Spitzer et al., 2006). In the present data set, its reliability was good ( $\alpha = .82$ ).

**Depression** Symptoms of depression were measured by the Beck Depression Inventory (BDI-21; Beck et al., 1988) translated into Finnish by Aalto (2009). The respondent answers 21 items inquiring the person's feelings and life situation during the past week. Each item receives a score from 0 to 3; thus, the summed score ranges from 0 to 63. The validity and reliability of the inventory has been estimated to be satisfactory (Beck et al., 1988). In the present data set, its reliability was assessed to be good ( $\alpha = .85$ ).

**Mindfulness** Mindfulness skills were measured by an abbreviated (24 items) version (FFMQ-SF; Bohlmeijer et al., 2011) of the Five Facet Mindfulness Questionnaire (FFMQ; Baer et al., 2008) translated into Finnish by Rantonen (2014). A 5-point Likert scale is used to answer the questions with the following response alternatives: never or very rarely true, rarely true, sometimes true, often true, very often or always true. The five facets of mindfulness measured by the questionnaire are observing, describing, acting with awareness, non-judging of inner experience, and non-reactivity to inner experience. The sum score varies from 5 to 120. Both versions of the

questionnaire are estimated to be valid and reliable indicators of mindfulness (Baer et al., 2008; Bohlmeijer et al., 2011). In the present study, the reliability of FFMQ-SF was good ( $\alpha = .85$ ).

**Meditation Practice Duration (Minutes)** A meditation practice minute counter was developed by the app firm and embedded within the app. When the participant practiced mindfulness using the app, his/her practice minutes were recorded and stored.

**Cortisol Levels** For a small subset of participants ( $n = 30$ ) in the intervention group, cortisol levels from saliva were collected and analyzed pre- and post-intervention. The sample was taken randomly from the intervention group and limited to 30 due to financial reasons. Three daily samples (after waking up, 30 min after waking up, and right before going to bed) were collected in three consecutive days (two working days and one leisure day). Thus, 18 samples were collected from each participant. They were instructed not to eat or drink 15–30 min prior to taking the sample and wash their mouth with water after eating. Moreover, detailed instructions were provided about how to take the sample and place it in the test tube. A return envelope was also provided for mailing the test tubes to the Finnish Institute of Occupational Health, where the samples were analyzed. One participant failed to return samples and thus, cortisol data was available from 29 participants.

## Statistical Analyses

The data were analyzed with linear mixed models using the *lme4* (1.1–21; Bates et al., 2015) and cumulative link mixed models using the *ordinal* (version 2019.12–10; Christensen, 2019) packages for *R* statistical software (version 3.6.1; R Core team, 2019). All dependent variables were analyzed with linear mixed models, except for the single item stress question, which was analyzed with a cumulative link mixed model. Separate models for each dependent variable were conducted, and they were of the form  $DV \sim \text{Time} * \text{Group} + (1|\text{Participant})$ , in which Time of testing (pre and post) and Group (control vs. treatment) were entered as sum coded fixed effects. Random intercept for participants was included in the random part of the model. Observations exceeding 3 SDs were excluded as outliers from the analyses reported (for the PSS score, 1 participant, and for the GAD and BDI scores, 3 participants were dropped). Analyses were also conducted on all data, and even though this did not change the overall pattern of results, the effects for PSS and BDI scores failed to reach significance when the outliers were included in the analyses. Follow-up comparisons for interactions were conducted with the *emmeans* package (version 1.4.5; Lenth et al., 2020) by computing contrasts between control and treatment groups. Cohen's *d* for the difference between pre- and post-test was

calculated using a pooled estimate of the pre- and post-test variance.

Cortisol measurements were collected from a subsample of the treatment group ( $N=29$ ), and changes between the pre- and post-test in cortisol were examined with a model  $DV \sim \text{Time} + (1|\text{Participant})$ , in which Time of testing (pre and post) was entered as a sum coded fixed effect. Note that there were 9 cortisol measurements at both pre and post-test phases.

As specified in the preregistration, exploratory analyses were conducted to examine whether treatment effects depended on the level of mindfulness after treatment. The post-test mindfulness score was centered and entered into the models as a fixed effect, including its interaction effects with Group and Time. We also explored differences in treatment effects between students, faculty and staff by entering participant group as a fixed effect in the models (bachelor’s degree students as the baseline group), including its interaction effects with Group and Time. In addition to these preregistered exploratory analyses, we also examined whether treatment effects observed in the treatment group were associated with practice time (i.e., minutes of app use). In these analyses, practice time was centered and entered in the models as a fixed effect, including its interaction terms with Time. As the control group had access to the app after the post-test, we also examined practice effects during the whole study period (from the pretest to the follow-up) for all participants.

## Results

### Perceived Stress

In the analysis of the PSS scores (see Table 1), there was a main effect of time, indicating that the scores decreased during the intervention period. Importantly, a Time\*Group interaction indicates that the scores decreased more in the treatment (Cohen’s  $d = -.50$ ) than in the control group ( $d = -.34$ ).

**Table 1** Model estimates for the PSS scores

Predictor	PSS		<i>t</i>
	<i>b</i>	<i>CI</i>	
Intercept	16.51	16.10–16.91	80.01
Time	1.09	0.86–1.31	9.40
Group	0.19	–0.22 – 0.59	0.91
Time: Group	–0.23	–0.45 – –0.00	–1.96
Random Effects			
$\sigma^2$	11.46		
$\tau_{00 \text{ user}}$	16.37		
$N_{\text{user}}$	560		
Observations	965		

However, in the follow-up comparisons the difference between control and treatment groups at post-test failed to reach statistical significance, estimate = .83,  $SE = .50$ ,  $t = 1.66$ ,  $p = .097$ .

The analysis of the single item stress question (see Table 2) showed that the score decreased during the intervention period ( $d = -.43$ ), and that overall, the treatment group had lower stress scores. However, there was no indication of an interaction.

### Anxiety

The analysis of the GAD scores (see Table 3) showed a main effect of time, indicating that the scores decreased during the intervention period ( $d = -.40$ ). However, there was no indication of an interaction involving Group.

### Depression

In the analysis of the BDI scores (see Table 4) there was a main effect of time, indicating that the scores decreased during the intervention period. Moreover, a Time \* Group interaction suggests that the scores decreased more in the treatment ( $d = -.45$ ) than in the control group ( $d = -.29$ ). In the follow-up comparisons, the difference between control and treatment groups at post-test failed to reach significance, estimate = .88,  $SE = .56$ ,  $t = 1.59$ ,  $p = .11$ .

### Mindfulness

The analysis of the FFMQ scores (see Table 5) showed an overall increase in the scores during the intervention period ( $d = .24$ ), but no indication of an interaction.

### Cortisol Levels

Cortisol data was collected from 29 participants. Of these, four did not provide samples in the post-test phase. There was no

**Table 2** Model estimates for the single item stress question

Predictor	Single item stress question		
	<i>OR</i>	<i>CI</i>	<i>Z</i>
Time	1.81	1.56–2.09	8.02
Group	1.24	1.00–1.52	2.00
Time: Group	0.92	0.80–1.05	–1.28
Random Effects			
$\sigma^2$	3.29		
$\tau_{00 \text{ user}}$	3.61		
$N_{\text{user}}$	561		
Observations	967		

**Table 3** Model estimates for the anxiety scores

Predictor	GAD		
	<i>b</i>	CI	<i>t</i>
Intercept	12.79	12.53–13.06	93.16
Time	0.66	0.51–0.82	8.38
Group	0.08	–0.19 – 0.34	0.55
Time: Group	–0.12	–0.27 – 0.04	–1.47
Random Effects			
$\sigma^2$	5.36		
$\tau_{00 \text{ user}}$	7.03		
$N_{\text{user}}$	558		
Observations	958		

indication of changes in the cortisol levels in the treatment group ( $d = .06$ ; see Table 6).

### Exploratory Analyses of the Relationship between Treatment Effects and Mindfulness Skills

The analyses with the post-test FFMQ score as a covariate in the models did not provide evidence for three-way interactions between time, treatment group and mindfulness after treatment. There were two-way interactions between time and FFMQ post-test score for PSS ( $b = .39$ ,  $SE = .18$ ,  $t = 3.28$ ), single item stress question ( $b = .17$ ,  $SE = .07$ ,  $z = 2.30$ ), and BDI ( $b = .41$ ,  $SE = .13$ ,  $t = 3.22$ ), indicating that the decrease in these scores were greater for participants who reported higher mindfulness after the treatment period.

### Exploratory Analyses of the Treatment Effects in Different Participant Groups

There were no indications of three-way interactions between time, treatment group and participant group.

**Table 4** Model estimates for depression scores

Predictor	BDI		
	Estimate	CI	<i>t</i>
Intercept	8.09	7.63–8.55	34.67
Time	0.96	0.72–1.21	7.77
Group	0.20	–0.26 – 0.65	0.84
Time: Group	–0.25	–0.49 – –0.00	–1.99
Random Effects			
$\sigma^2$	12.94		
$\tau_{00 \text{ user}}$	21.85		
$N_{\text{user}}$	558		
Observations	956		

**Table 5** Model estimates for the mindfulness scores

Predictor	FFMQ		
	Estimate	CI	<i>t</i>
Intercept	83.12	82.21–84.03	179.87
Time	–1.22	–1.63 – –0.80	–5.78
Group	–0.29	–1.20 – 0.61	–0.63
Time: Group	0.36	–0.05 – 0.77	1.71
Random Effects			
$\sigma^2$	37.21		
$\tau_{00 \text{ user}}$	95.00		
$N_{\text{user}}$	561		
Observations	966		

### Exploratory Analyses of the Relationship between Treatment Effects and App Use (Dose-Response Relationship)

The total app use in the intervention group ranged from 8.6 min to 678.7 min ( $M = 195.6$  min). The participants assigned to the intervention group used the app from 1 to 71 times, with an average of 21.4 times, during the intervention. In the analyses with total time that the app was used during the intervention as one of the fixed effects, there was a Time \* Practice time interaction in the FFMQ scores,  $b = -.74$ ,  $SE = .36$ ,  $t = -2.06$ , indicating that the more the app was used, the more the FFMQ scores increased from pre- to post-test in the intervention group. No indication of interaction effects was observed in the other measures.

The control group got access to the app after the intervention period, so we also analyzed whether the app use in both the control and the intervention groups was related to changes in the independent measures between the pre-test and follow-up. The total app use during the whole study period ranged from 2.5 min to 1985.7 min ( $M = 235.9$  min). The participants used the app from 1 to 165 times, with an average of 24.8 times, during the whole study period. There were no

**Table 6** Model estimates for cortisol (nmol/l)

Predictor	Cortisol		
	Estimate	CI	<i>t</i>
Intercept	7.52	6.65–8.40	16.79
Time	–.16	–.74 – -.43	–.53
Random Effects			
$\sigma^2$	40.11		
$\tau_{00 \text{ user}}$	3.26		
$N_{\text{user}}$	29		
Observations	470		

indications of interaction effects involving practice time in any of the measures.

## Discussion

We set out to investigate whether a 4-week course on a mindfulness app would result in wellbeing benefits over and above those from an online mindfulness course without the active ingredient of actual meditation practice. We also assessed whether more time spent practicing on the one hand, and increases in mindfulness on the other hand, would result in additional benefit to participants. Our results indicate that practicing mindfulness with the help of a mindfulness app resulted in alleviated stress (though not when measured through the single-item stress question) and depression. The results for anxiety were also close to statistical significance. The effects are evident in comparison with an active control group, which is a gold standard in intervention research. Thus, the present study provides some evidence of the efficacy of using a mindfulness app to alleviate mental suffering. Contrary to our expectation, more time spent practicing did not predict further alleviation in stress, depression, or anxiety, despite the fact that the degree of perceived mindfulness increased as a function of practice time.

Our results regarding intervention effects are in line with the previous literature, which has indicated lessened anxiety and depression and improved wellbeing from practicing mindfulness using apps (e.g., Bostock et al., 2019; Clarke & Draper, 2020). Furthermore, our study provides further evidence for a beneficial intervention effect on stress, as reported for online mindfulness-based interventions in the meta-analysis of Spijkerman et al. (2016), which was not conclusively demonstrated for face-to-face MBIs (see the meta-analysis of Goyal et al., 2014). Finally, cortisol levels were not found to be decreased by the intervention. Yet, the small sample size likely played a role in burying any possible effect. As biological measures may become more readily available in the future, studies would do well to include them more as sensitive and objective outcome measures.

As the need for large-scale, cost-effective, low-risk interventions in mental health is large and growing, our study provides crucial evidence their usability at least in a university student and staff population like the one examined in the present study. The 2020 pandemic situation has created further anxiety and stress in the society and has also made, for the time being, face-to-face mindfulness-based interventions nearly impossible to administer with restrictions on people gathering in the same space. As evidence of their benefits accumulate, mindfulness apps are likely to increasingly meet the need of people searching for peace and calm amid stormy life circumstances.

Time spent practicing did not modify the intervention effects. This was likely due to there being relatively little variance in the practice times. The app offered a 10-min daily meditation practice for participants and there was no added incentive to practice more. As the mean daily practice duration was between 8 and 9 min, participants were quite diligent in their practice routines. It is encouraging that positive intervention effects were obtained with relatively small daily time investments. Yet, effects from additional practice would have needed more extensive individual variability in practice time. Finally, it may be noted that individual variability in the perceived mindfulness did not modify the observed intervention effects. This may not be very surprising given the fact that mindfulness measures are somewhat unreliable indicators for actual increases in mindfulness; thus, salutary effects may have been lost in noise (Goleman & Davidson, 2017).

## Limitations, Strengths, and Future Directions

Even though the study was uncommonly robust for the field (a large sample size combined with an active control group), some cautionary notes are in order. First, even though attrition rate was lower than has been reported for non-mindfulness mHealth interventions (31%; Melville et al., 2010), the study's rate of 28.0% could still have been lower, perhaps with the help of gamification in app design, by adding incentives to remain in the study via a reward system, or by providing more encouraging feedback. Second, the sex balance in the sample was lopsided with more than five times more females than males. This is a common finding in volunteer-based mindfulness studies, as females appear to volunteer more often than males to mindfulness studies (e.g., Lahtinen & Salmivalli, 2020). Using a sample of volunteers is also a limitation, as there may be a reason to self-select, which leads to bias. The sample may then include an excessive number of people eager to meditate, or, on the other hand, participants with pronounced suffering in their lives. Third, the saliva sample subgroup in our study was too small to draw firm conclusions related to effects on stress-hormone levels. Moreover, due to financial reasons the cortisol group had no control group. Thus, these results should be treated with appropriate caution.

Fourth, the active control, though an unusually robust control condition for the field, could still have been designed to more closely resemble the treatment condition – it could have been an app otherwise identical to the main app, just without the meditation practices. The best studies in this regard have offered a “placebo” treatment as an active control condition via the same modality as the treatment condition, e.g. the Health Enhancement Program used in many face-to-face MBI studies (Eisendrath et al., 2014; Goleman & Davidson, 2017; MacCoon et al., 2012). Fifth, the use of self-report measures for main outcomes may be considered a limitation,

though this is how depression and anxiety are commonly measured.

Notable strengths in the study are its large sample size, its robust active control condition, its implementation of strict estimates of actual meditation practice minutes and its preregistration. The methodology in much of mindfulness research has been notoriously shoddy (Coronado-Montoya et al., 2016; Goleman & Davidson, 2017; Van Dam et al., 2018). Thus, high-quality meta-analyses typically can only include <5% of all available studies due to lapses in methodological rigor (Goyal et al., 2014; Kuyken et al., 2015). As the most popular mindfulness apps now have users in the millions, the future of mindfulness practice is likely to increasingly involve the use of apps. High-quality studies evaluating their efficacy are thus likely to be in high demand and the field should take up the gauntlet.

**Data Availability Statement** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Author Contributions** OL: Designed part of the study, partly supervised JA's work, and wrote the abstract, introduction, and discussion parts of the manuscript. He also acts as the corresponding author and is responsible for responding to the peer reviews.

JA: Designed and conducted most of the study and wrote a Master's thesis based on the results.

JK: Conducted statistical analyses for the study and wrote the results section.

LF: Designed the apps evaluated in the study. This author did not have any role in designing the study or evaluating the apps.

JH: Designed part of the study, wrote the method section of the manuscript, commented on the draft, and supervised A2's work.

**Funding** Open access funding provided by University of Turku (UTU) including Turku University Central Hospital.

## Declarations

**Ethical Approval** Ethical review of the study plan was approved by the University of Turku Ethics Committee (decision number: 34/2018).

**Research Involving Human and Animal Rights** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

**Conflict of Interest** A4 developed the studied apps. Other authors declare no other potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were

made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Baer, R. A., Smith, G. T., Lykins, E., Button, D., Krietemeyer, J., Sauer, S., Walsh, E., Duggan, D., & Williams, J. M. G. (2008). Construct validity of the five facet mindfulness questionnaire in meditating and nonmeditating samples. *Assessment*, 15(3), 329–342. <https://doi.org/10.1177/1073191107313003>.
- Bates, D., Maechler, M., Bolker, B., Walker, S., & Haubo Bojesen Christensen, R. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.
- Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck depression inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77–100. [https://doi.org/10.1016/0272-7358\(88\)90050-5](https://doi.org/10.1016/0272-7358(88)90050-5).
- Bennike, I. H., Wieghorst, A., & Kirk, U. (2017). Online-based mindfulness training reduces behavioral markers of mind wandering. *Journal of Cognitive Enhancement*, 1(2), 172–181. <https://doi.org/10.1007/s41465-017-0020-9>.
- Bohlmeijer, E., Ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R. (2011). Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. *Assessment*, 18(3), 308–320. <https://doi.org/10.1177/1073191111408231>.
- Bostock, S., Crosswell, A. D., Prather, A. A., & Steptoe, A. (2019). Mindfulness on-the-go: Effects of a mindfulness meditation app on work stress and well-being. *Journal of Occupational Health Psychology*, 24(1), 127–138. <https://doi.org/10.1037/ocp0000118>.
- CDC (2020). Household Pulse Survey. <https://www.cdc.gov/nchs/covid19/pulse/mental-health.htm>
- Christensen, R. H. B. (2019). Ordinal - regression models for ordinal data. R package version 2019.12–10. <https://CRAN.R-project.org/package=ordinal>.
- Clarke, J., & Draper, S. (2020). Intermittent mindfulness practice can be beneficial, and daily practice can be harmful. An in depth, mixed methods study of the “calm” app's (mostly positive) effects. *Internet Interventions*, 19, 100293. <https://doi.org/10.1016/j.invent.2019.100293>.
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 385–396. <https://doi.org/10.2307/2136404>.
- Coronado-Montoya, S., Levis, A. W., Kwakkenbos, L., Steele, R. J., Turner, E. H., & Thombs, B. D. (2016). Reporting of positive results in randomized controlled trials of mindfulness-based mental health interventions. *PloS one*, 11(4), e0153220. <https://doi.org/10.1371/journal.pone.0153220>.
- Davidson, R. J., & Dahl, C. J. (2018). Outstanding challenges in scientific research on mindfulness and meditation. *Perspectives on Psychological Science*, 13(1), 62–65. <https://doi.org/10.1177/1745691617718358>.
- de Vibe, M. F., Bjørndal, A., Fattah, S., Dyrdal, G. M., Halland, E., & Tanner-Smith, E. E. (2017). Mindfulness-based stress reduction (MBSR) for improving health, quality of life and social functioning in adults: A systematic review and meta-analysis. *Campbell Systematic Reviews*, 13(11), 1–264. <https://doi.org/10.4073/csr.2017.11>.

- Economides, M., Martman, J., Bell, M. J., & Sanderson, B. (2018). Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone app: A randomized controlled trial. *Mindfulness*, *9*(5), 1584–1593. <https://doi.org/10.1007/s12671-018-0905-4>.
- Eisendrath, S. J., Gillung, E. P., Delucchi, K. L., Chartier, M., Mathalon, D. H., Sullivan, J. C., Segal, Z. V., & Feldman, M. D. (2014). Mindfulness-based cognitive therapy (MBCT) versus the health-enhancement program (HEP) for adults with treatment-resistant depression: A randomized control trial study protocol. *BMC Complementary and Alternative Medicine*, *14*(1), 95. <https://doi.org/10.1186/1472-6882-14-95>.
- Elo, A. L., Leppänen, A., & Jahkola, A. (2003). Validity of a single-item measure of stress symptoms. *Scandinavian Journal of Work, Environment & Health*, *444*–451. <https://www.jstor.org/stable/40967322>
- Flett, J. A., Hayne, H., Riordan, B. C., Thompson, L. M., & Conner, T. S. (2019). Mobile mindfulness meditation: A randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness*, *10*(5), 863–876. <https://doi.org/10.1007/s12671-018-1050-9>.
- Goleman, D., & Davidson, R. J. (2017). *Altered traits: Science reveals how meditation changes your mind, brain, and body*. Penguin.
- Goyal, M., Singh, S., Sibinga, E. M., Gould, N. F., Rowland-Seymour, A., Sharma, R., et al. (2014). Meditation programs for psychological stress and well-being: A systematic review and meta-analysis. *JAMA Internal Medicine*, *174*(3), 357–368. <https://doi.org/10.1001/jamainternmed.2013.13018>.
- Hendricks, D. S., Lavery, M. R., Bouillon, L. E., & Gotfried, R. S. (2019). Evaluating the effectiveness of a mindfulness app among academic advisors. *Journal of Technology in Behavioral Science*, *5*, 1–7. <https://doi.org/10.1007/s41347-019-00122-w>.
- Howells, A., Ivztan, I., & Eiroa-Orosa, F. J. (2016). Putting the ‘app’ in happiness: A randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. *Journal of Happiness Studies*, *17*(1), 163–185. <https://doi.org/10.1007/s10902-014-9589-1>.
- Huberty, J., Green, J., Glissmann, C., Larkey, L., Puzia, M., & Lee, C. (2019). Efficacy of the mindfulness meditation mobile app “calm” to reduce stress among college students: Randomized controlled trial. *JMIR mHealth and uHealth*, *7*(6), e14273. <https://doi.org/10.2196/14273>.
- Hyönä, J., Aaltonen, J., Lahtinen, O., Franklin L., & Kaakinen J. (2018). Mindfulness meditation supported by a smartphone app. Retrieved from <https://osf.io/mtfb4/#!>. Accessed 30 Apr 2021.
- Kabat-Zinn, J. (1990). *Full catastrophe living: The program of the stress reduction clinic at the University of Massachusetts Medical Center*.
- Kunttu, K., Pesonen, T., & Saari, J. (2016). Korkeakouluopiskelijoiden terveystutkimus 2016. Ylioppilaiden terveydenhoitosäätiön tutkimuksia 48. Arkimedia Oy, Helsinki.
- Kuyken, W., Hayes, R., Barrett, B., Byng, R., Dalgleish, T., Kessler, D., et al. (2015). Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (PREVENT): A randomised controlled trial. *The Lancet*, *386*(9988), 63–73. [https://doi.org/10.1016/S0140-6736\(14\)62222-4](https://doi.org/10.1016/S0140-6736(14)62222-4).
- Lahtinen, O., & Salmivalli, C. (2020). An effectiveness study of a digital mindfulness-based program for upper secondary education students. *Mindfulness*, *11*(11), 2494–2505. <https://doi.org/10.1007/s12671-020-01462-y>.
- Lenth, R., Singmann, H., Love, J., Buerkner, P., & Herve, M. (2020). Emmeans: Estimated marginal means, aka least-squares means (1.4.5)[computer software].
- Lyzwinski, L. N., Caffery, L., Bambling, M., & Edirippulige, S. (2019). The mindfulness app trial for weight, weight-related behaviors, and stress in university students: Randomized controlled trial. *JMIR mHealth and uHealth*, *7*(4), e12210. <https://doi.org/10.2196/12210>.
- MacCoon, D. G., Imel, Z. E., Rosenkranz, M. A., Sheffel, J. G., Weng, H. Y., Sullivan, J. C., Bonus, K. A., Stoney, C. M., Salomons, T. V., Davidson, R. J., & Lutz, A. (2012). The validation of an active control intervention for mindfulness based stress reduction (MBSR). *Behaviour Research and Therapy*, *50*(1), 3–12. <https://doi.org/10.1016/j.brat.2011.10.011>.
- Mani, M., Kavanagh, D. J., Hides, L., & Stoyanov, S. R. (2015). Review and evaluation of mindfulness-based iPhone apps. *JMIR mHealth and uHealth*, *3*(3), e82. <https://doi.org/10.2196/mhealth.4328>.
- Melville, K. M., Casey, L. M., & Kavanagh, D. J. (2010). Dropout from internet-based treatment for psychological disorders. *British Journal of Clinical Psychology*, *49*(4), 455–471. <https://doi.org/10.1348/014466509X472138>.
- Noone, C., & Hogan, M. J. (2018). A randomised active-controlled trial to examine the effects of an online mindfulness intervention on executive control, critical thinking and key thinking dispositions in a university student sample. *BMC psychology*, *6*(1), 13. <https://doi.org/10.1186/s40359-018-0226-3>.
- Parsons, C. E., Crane, C., Parsons, L. J., Fjorback, L. O., & Kuyken, W. (2017). Home practice in mindfulness-based cognitive therapy and mindfulness-based stress reduction: A systematic review and meta-analysis of participants’ mindfulness practice and its association with outcomes. *Behaviour Research and Therapy*, *95*, 29–41. <https://doi.org/10.1016/j.brat.2017.05.004>.
- Pierceall, E. A., & Keim, M. C. (2007). Stress and coping strategies among community college students. *Community College Journal of Research and Practice*, *31*(9), 703–712. <https://doi.org/10.1080/10668920600866579>.
- Rantonen, O. (2014). Tietoisuustaitojen hierarkkinen malli. Master’s thesis. [Turku].
- Rosen, K. D., Paniagua, S. M., Kazanis, W., Jones, S., & Potter, J. S. (2018). Quality of life among women diagnosed with breast Cancer: A randomized waitlist controlled trial of commercially available mobile app-delivered mindfulness training. *Psycho-oncology*, *27*(8), 2023–2030. <https://doi.org/10.1002/pon.4764>.
- Roy, A., Druker, S., Hoge, E. A., & Brewer, J. A. (2020). Physician anxiety and burnout: Symptom correlates and a prospective pilot study of app-delivered mindfulness training. *JMIR mHealth and uHealth*, *8*(4), e15608. <https://doi.org/10.2196/15608>.
- Spijkerman, M. P. J., Pots, W. T. M., & Bohlmeijer, E. T. (2016). Effectiveness of online mindfulness-based interventions in improving mental health: A review and meta-analysis of randomised controlled trials. *Clinical Psychology Review*, *45*, 102–114. <https://doi.org/10.1016/j.cpr.2016.03.009>.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>.
- Teasdale, J. D., Segal, Z. V., Williams, J. M. G., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology*, *68*(4), 615–623. <https://doi.org/10.1037/0022-006X.68.4.615>.
- Torous, J., Lipschitz, J., Ng, M., & Firth, J. (2020). Dropout rates in clinical trials of smartphone apps for depressive symptoms: A systematic review and meta-analysis. *Journal of Affective Disorders*, *263*, 413–419. <https://doi.org/10.1016/j.jad.2019.11.167>.
- Van Dam, N. T., van Vugt, M. K., Vago, D. R., Schmalzl, L., Saron, C. D., Olendzki, A., et al. (2018). Mind the hype: A critical evaluation and prescriptive agenda for research on mindfulness and meditation. *Perspectives on Psychological Science*, *13*(1), 36–61. <https://doi.org/10.1177/1745691617709589>.
- Wang, K., Varma, D. S., & Prosperi, M. (2018). A systematic review of the effectiveness of mobile apps for monitoring and management of

- mental health symptoms or disorders. *Journal of Psychiatric Research*, 107, 73–78. <https://doi.org/10.1016/j.jpsychires.2018.10.006>.
- Wen, L., Sweeney, T. E., Welton, L., Trockel, M., & Katznelson, L. (2017). Encouraging mindfulness in medical house staff via smartphone app: A pilot study. *Academic Psychiatry*, 41(5), 646–650. <https://doi.org/10.1007/s40596-017-0768-3>.
- Yang, E., Schamber, E., Meyer, R. M., & Gold, J. I. (2018). Happier healers: Randomized controlled trial of mobile mindfulness for stress management. *The Journal of Alternative and Complementary Medicine*, 24(5), 505–513. <https://doi.org/10.1089/acm.2015.0301>.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.