Can a self-guided mindfulness mobile app reduce symptoms of anxiety? A randomised controlled trial

Chris Worfolk, Leeds Beckett University

Objective: Mobile-based (mHealth) mindfulness apps are widely used as a form of self-help for anxiety disorders, but little is known about their efficacy. A randomised controlled trial was designed to measure their outcomes.

Methods: 379 participants with generalised anxiety disorder were recruited. Eligible participants were given a mindfulness app (n = 157), relaxing music app (n = 157) or placed on a waiting list (n = 60). Participants with an app were asked to use it daily for four weeks. GAD-7 and PHQ-9 measures were used to track anxiety and mood at baseline, mid (2 weeks) and end (4 weeks) of the study.

Results: 89 participants completed all measures. Both app groups experienced a significant reduction in anxiety at the mid point, but only the mindfulness app group saw a significant reduction between the mid and final points. All groups experienced a significant improvement in mood over the four-week period but only the mindfulness app group experienced a significant improvement at all time points. The mindfulness group experienced a larger clinical effect in both anxiety (d = 0.57) and mood (d = 0.89). Improvement did not correlate with level of engagement.

Conclusions: mHealth mindfulness interventions provide a benefit above and beyond simple relaxation and are an effective way to reduce symptoms of generalised anxiety disorder. Daily practice may not be required to see the benefits from mindfulness.

Keywords: mHealth, mindfulness, anxiety, generalised anxiety disorder

Introduction

Anxiety and depression are a significant burden on the NHS, accounting for more spending than cancer and heart disease (Nuffield Trust, 2014). In many countries, adequate treatment is not available (Alonso et al., 2018). One of the major costs of providing interventions is the therapist (Mukuria et al., 2013).

Further, because of the stigma and long waiting lists, many people do not seek the help they need (Barney, Griffiths, Jorm & Christensen, 2006). This has created a large gap for anonymous, flexible and more cost-effective treatments that would increase utilisation while reducing healthcare spending.

Mindfulness-based therapy is one possible solution. Mindfulness can be defined as

paying attention to the present moment while maintaining a non-judgmental awareness of experiences (Kabat-Zinn, 2009). Mindfulness is typically developed through formal practice sessions known as mindfulness meditations (Hölzel et al., 2011).

Several mechanisms for mindfulness have been proposed, including enhanced self-regulation (Bishop et al., 2004), greater cognitive and emotional flexibility (Gu, Strauss, Bond, & Cavanagh, 2015) and increased ability to control one's attention (Shapiro, Carlson, Astin, & Freedman, 2006).

Mindfulness has been shown to be effective for a wide range of psychological disorders (Khoury et al., 2013), including depression (Hofmann, Sawyer, Witt, & Oh, 2010), generalised anxiety disorder (GAD) (Evans et al., 2008), panic disorder (Kim et al., 2009) and social phobia (Piet, Hougaard, Hecksher, & Rosenberg, 2010). In addition, it has been suggested that mindfulness provides a range of other benefits including increased empathy (Davis & Hayes, 2011), a stronger immune system (Black & Slavich, 2016) and reduced sensitivity to pain (Zeidan & Vago, 2016). It may be more cost-effective than cognitive behavioural therapy (CBT) because it requires less professional training and less time to master (Singh & Gorey, 2017).

Another possible avenue is the field of mobile health (mHealth). 76% of adults in the UK have a smartphone (Ofcom, 2017) so this could represent an accessible and effective way to deliver interventions.

Several mHealth apps have already demonstrated efficacy in supporting Acceptance & Commitment Therapy (ACT) (Levin, Haeger, Pierce, & Cruz, 2017) and alleviating the symptoms of depression (Watts et al., 2013) and post-traumatic stress disorder (PTSD) (Miner et al., 2016). However, these have all required the support of a clinician, which makes delivery complicated and expensive (Kolovos et al., 2018).

A review published by the Anxiety and Depression Association of America (ADAA) identified 52 commercially available mHealth apps targeting anxiety but noted that only two of them were able to provide any evidence of effectiveness in reducing anxiety (Sucala et al. 2017). This is troublesome as ineffective apps may deter patients from seeking further treatment (Price et al. 2014). One of the most popular apps, Headspace, claims to have been downloaded by 16 million people (Griffith, 2017).

The lack of evidence has not stopped individual NHS Trusts and IAPTs (Improving Access to Psychological Therapies) from recommending apps to patients. Bennion, Hardy, Moore, & Millings (2016) used Freedom of Information requests to compile a list of 35 apps currently being recommended by different parts of NHS England despite a lack of clinical evidence. This included four mindfulness-based apps: Headspace, Digipill, Mindfulness Bell and Take A Break.

Evidentially then, there is demand both from the public and from healthcare providers for mobile-based solutions. In order to provide safe, effective and evidence-based care, it is crucial for the scientific literature to begin to examine the effectiveness of this technology.

Slowly, this has started to happen. Pham, Khatib, Stansfeld, Fox, & Green (2016) tested the app "Flowy", a breathing exercise delivered in the form of a game. The user used their finger to indicate when they are breathing in and out and this moves the game along. The game did demonstrate a reduction in panic over the four-week trial, but the reduction in anxiety was not significant.

Enock, Hofmann, & McNally (2014) tested an attention bias modification training app in which users performed a dot-probe exercise where they had to select pictures of faces showing neutral expressions while ignoring pictures of faces showing disgust. They measured social anxiety symptoms and found no significant difference between the experimental and control groups.

Other studies have found a significant reduction in anxiety symptoms but failed to isolate the mobile app as the independent variable or suffered from other methodological flaws. Ivanova et al. (2016) tested using guided vs unguided ACT using an app. They found that unguided therapy did not significantly differ from the therapist-supported group. However, the intervention also used internet-based lessons, exercises, books and CDs.

Proudfoot et al. (2013) tested the app "myCompass", a toolkit app that includes techniques from CBT, Interpersonal Psychotherapy, Problem-solving Therapy and Positive Psychology. They found that the app outperformed both a waiting list and a standard attention treatment. However, the participants were also required to use a computer to access some of the lessons.

Bakker & Rickard (2018) tested the self-monitoring app "MoodPrism" in which users were prompted to complete daily questionnaires about their mood. They found app engagement significantly predicted a reduction in GAD symptoms, but this reduction was only present in their clinical population. Further, the study was based solely on user data from the app, and thus there was no control group to compare against.

Only one study has looked at the efficacy of mindfulness-based apps. Lee & Jung (2018) tested "DeStressify", which failed to significantly reduce anxiety. Further, it tells us little about the ability to treat GAD because the sample was composed of university undergraduates rather than a relevant population.

So far, then, we know that mindfulness is an effective treatment for GAD and that mHealth apps can be an effective treatment for a range of other psychological disorders. However, we do not know if self-guided mHealth apps are an effective way to deliver mindfulness to reduce symptoms of GAD because no study has tested such apps on the relevant population. This is a serious gap because clinicians are already recommending these apps to their patients.

The current study is a randomised controlled trial (RCT) designed to test the efficacy of self-guided mindfulness mHealth apps in reducing the symptoms of anxiety in a sample of people who meet the criteria for GAD. A mindfulness mHealth app was developed and compared to an active control group who received a relaxing music app, and a waiting list control group. Participants' anxiety and mood were

tracked over a four-week period.

It was hypothesised that: (a) participants using the mindfulness app would experience a greater reduction in anxiety symptoms than those in the control groups and that (b) participants using the mindfulness app would experience a greater improvement in mood than those in the control groups.

Relevance to practising clinicians

Mindfulness mHealth apps already play a role in the treatment of psychological conditions, but we do not as of yet know if they work. Answering this question will allow practising clinicians to make informed recommendations to patients presenting symptoms of GAD.

Method

Design

A parallel single-blind independent-groups RCT using a matched-pair design based on participant's GAD-7 inventory scores (see materials sub-section for full details), gender and age (in that order).

Sample

Participants were recruited using volunteer sampling. The study was advertised via Leeds Beckett University and the charity Anxiety Leeds, and by contacting other local and national mental health organisations with which they have a relationship. All organisations were non-NHS. Additionally, adverts were run on social media. All recruitment material advertised the study as looking at the effects of mobile apps on anxiety symptoms.

Eligibility was determined by scoring >= 5 on the GAD-7 inventory, which is the threshold for anxiety (Spitzer, Kroenke, Williams, & Löwe, 2006), being over 18, having an Apple or Android-based smartphone, not currently regularly using mindfulness and not currently in counselling therapy (including CBT). Additionally, anyone who was experiencing, or had previously experienced, suicidal ideation was excluded.

Recruitment took place in two rounds. In the first round, 180 participants were recruited and allocated to all three groups. Due to high levels of attrition in the app groups, a second round of recruitment took place and participants were allocated to one of the two app groups only. Demographic details can be found in the results section.

Materials

Self-report inventories were used to track participants' mood.

Generalized Anxiety Disorder Scale (GAD-7) (Spitzer, Kroenke, Williams, & Löwe, 2006). A seven-item inventory with responses rated on a scale of 0 (not at all) to 3 (nearly every day) with a score of >= 5 representing mild anxiety, >= 10 representing moderate anxiety and >= 15 representing severe anxiety. An example item is "over the past two weeks, how often have you been bothered by feeling nervous, anxious or on edge?"

Internal consistency of the current sample, as measured by Cronbach's alpha, was good (α = 0.84) and comparable to existing literature (Seo & Park, 2015).

Patient Health Questionnaire (PHQ-9)

(Kroenke, Spitzer, & Williams, 2001). A nine-item inventory for tracking mood which uses the same responses as the GAD-7. An example item is "over the past two weeks, how often have you been bothered by feeling down, depressed, or hopeless?"

Internal consistency of the current sample, as measured by Cronbach's alpha, was good (α = 0.83) and comparable to the results of existing literature (Kroenke et al. 2001).

The final question of the PHQ-9 concerns suicidal ideation. As anyone experiencing suicidal ideation was excluded from the study, and to protect participants from potentially distressing questions, this question was removed.

These two measures were selected as they are the standard measures used by IAPTs (Gyani, Shafran, Layard, & Clark, 2013) and are recommended by the International Consortium for Health Outcomes Measurement (ICHOM) (Obbarius et al., 2017) and the National Institute for Health & Care Excellence (NICE) (Kendrick & Pilling, 2012). They are also short, which reduces the risk of abandonment by participants.

The app was built using the JavaScript programming language to allow it to work across both Apple and Android-based devices. This approach had the advantage of allowing any bugs to be fixed quickly, but the disadvantage that participants had to be connected to the internet while using it.

It was built using the React framework (for detailed information see Gackenheimer, 2015) and compiled into a single file by Webpack so that once the app was loaded, participants would not experience any delays while using the app, just like a standard app. The audio files were downloaded on demand, as most apps do.

The app contained two versions of itself and presented a different version based on the login information. One version contained a selection of three mindfulness meditations: one focusing on breath, one focusing on internal bodily sensations and one focusing on external sensory information. Each meditation contained guided audio and was variable in length (5:35-10:58). Information on how to practise mindfulness was also included.

An alternative version of the app contained relaxing music. A selection of seven songs by the musician Chris Zabriskie was chosen due to they're relaxing nature and open source licence. The songs varied in length (5:01-10:56) and information on how to play the songs was also included.

Procedure

Ethical approval was gained from the School of Psychology ethics committee.

Participants were directed to a website where they were put through an initial screening process in which they completed a GAD-7 inventory, demographic questions (age, gender) and about current use of therapy, medication and mindfulness apps.

Eligible participants were randomly assigned to one of three groups: mindfulness app, relaxation app active control and waiting list. Randomisation was done by computer using the Mersenne Twister algorithm (Matsumoto

& Nishimura, 1998).

Those allocated to the mindfulness app and relaxation app active control group were emailed instructions on how to download and use the app. They were not told which app they were getting or that there were different apps. Participants were instructed to use the app once per day over a four-week period. Daily interaction was preferable because practice levels have a significant impact on mindfulness (Jha, Morrison, Parker, & Stanley, 2017). Throughout the period they were sent regular reminder emails.

Participants completed GAD-7 and PHQ-9 measures at baseline, mid-intervention (2 weeks) and immediately post-intervention (4 weeks). This was done in the app. Their use of the app was also monitored to track treatment compliance.

Those allocated to the waiting list group were sent an email informing them that they would receive the app and the results of the study after it had concluded. They were asked to complete the same measures at the same time points as the app groups, which they did via a website.

After the study was complete, participants were given a debrief, and all participants were given access to the mindfulness app. In addition, the preliminary results of the study were shared with them. Participants were offered sources of support including the NHS, Samaritans and Mind.

Data analysis

Statistical analyses were carried out using IBM SPSS Statistics version 24.

Kolmogorov-Smirnov tests and one-way ANOVAs were run on participant's age and GAD-7 scores on registration. Chi-square tests were run on gender, duration of anxiety and use of medication. Fisher's exact correction was used where necessary. All tests suggest there was no significant difference between groups.

Kolmogorov-Smirnov tests were run for each time point and group on both the GAD-7 and PHQ-9 scores. Results indicate all data were normally distributed (all $p > .05$). Mauchly's test was consulted for each ANOVA and Greenhouse-Geisser's correction was used where necessary.

The primary measure of outcome was a reduction of scores on the GAD-7. Data were analysed using a 3x3 mixed factorial ANOVA with between-groups (mindfulness app group, relaxation app active control group and waiting list group) and time as a repeated measure (baseline, 2 weeks, 4 weeks). The same test was also run for mood based on the PHQ-9 scores. Bonferroni post hoc tests were used, and effect sizes were calculated using Cohen's *d* (Cohen, 1988).

Results

Demographics

379 participants were recruited with 89 being included in the final data analysis due to ineligibility ($n = 5$) and attrition ($n = 285$). A participant flowchart is shown in Figure 1. 21.0% of the mindfulness app group, 19.1% of the relaxation app control group and 43.3% of the waiting list group were included in the final analysis.

Demographic information can be seen in Table 1. Participants were predominantly female with the majority of them having suffered from anxiety for over 12 months.

Engagement

Figure 2 shows the number of days participants engaged with the app. Only the 28 days of the study period are included; usage after this, including submitting the final set of measures, was not counted. Figure 3 shows the number of users active on each day of the trial.

A t-test revealed no significant difference between the mindfulness app group (M = 8.52, SD = 5.72) and the relaxation app active control group (M = 8.17, SD = 4.92)'s engagement with the app (t (61) = .26, p = .359).

There was no significant correlation between baseline GAD-7 scores and engagement with the app in either the mindfulness app group (r = .05, p = .777, n = 33, R^2 = .003) or the relaxation app active control group (r = .18, p = .345, n = 30, R^2 = 0.032).

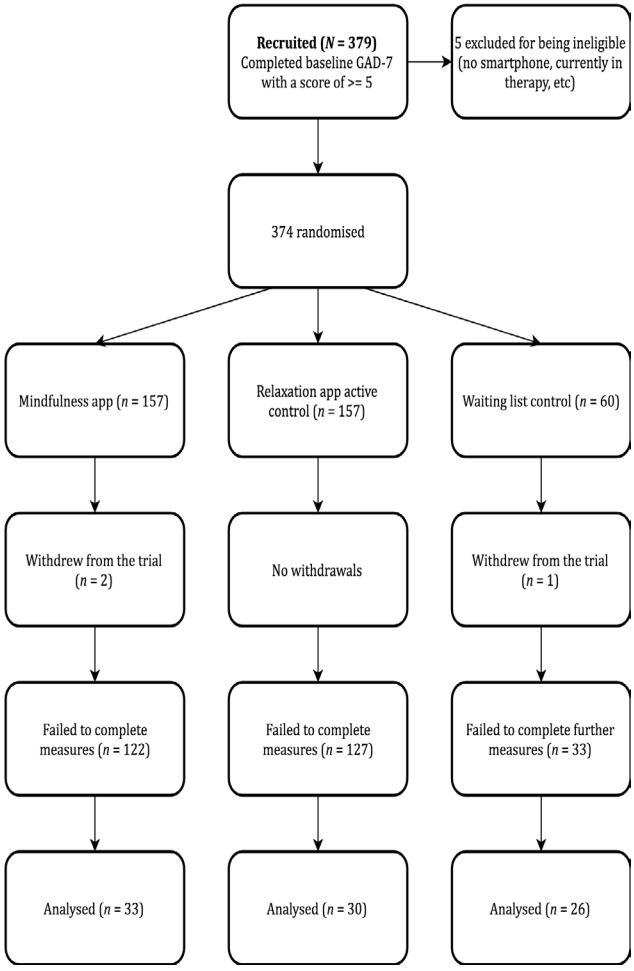

Figure 1. Participant flow, showing the different groups and what stage participants reached in the study

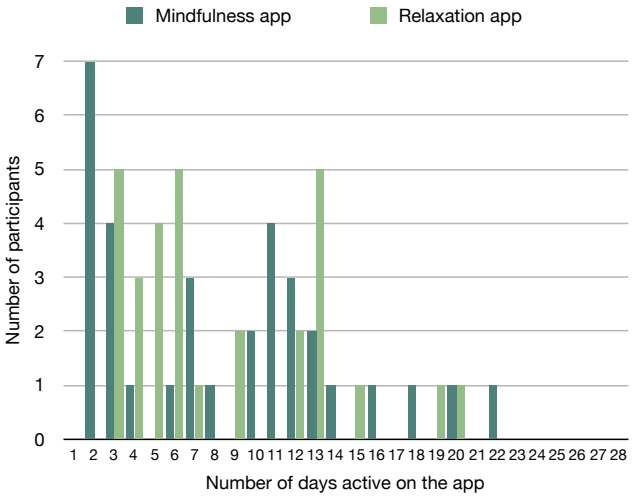

Figure 2. Number of days participants were active on the app

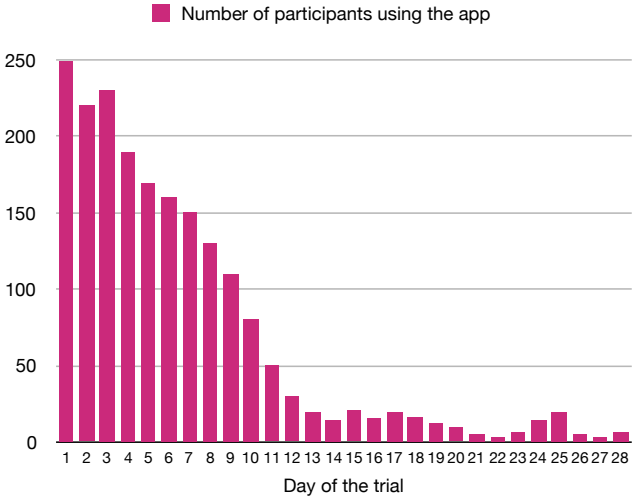

Figure 3. Number of users that were active on the app on each day of the trial

Table 1. Demographic information of participants included in the analysis

		Experimental	*Control*	*Waiting list*	*Total*
Age	*Mean (SD)*	38.03 (9.21)	37.67 (11.43)	39.77 (13.52)	38.42 (11.25)
Gender	*Female*	31 (94%)	28 (93%)	23 (88%)	82 (92%)
	Male	2 (6%)	2 (7%)	3 (12%)	7 (8%)
Duration of anxiety	*Less than 3 months*	1 (3%)	0	0	1 (1%)
	3-6 months	1 (3%)	2 (7%)	1 (4%)	4 (4%)
	6-12 months	1 (3%)	2 (7%)	0	3 (3%)
	More than 12 months	29 (88%)	26 (87%)	24 (92%)	79 (89%)
	Unspecified	1 (3%)	0	1 (4%)	2 (2%)
Using medication	*Yes*	22 (67%)	21 (70%)	16 (62%)	59 (66%)
	No	11 (33%)	9 (30%)	10 (38%)	30 (34%)
GAD-7 score at registration	*Mean (SD)*	15.17 (4.19)	14.61 (3.99)	13.55 (4.46)	14.53 (4.19)

Percentages are rounded and therefore may not total up to 100%. GAD-7 scores were recorded at registration to screen for eligibility and allow pair matching. No PHQ-9 scores were recorded at registration.

Anxiety

A factorial ANOVA revealed a significant main effect of time on GAD-7 scores between the start and end of the trial (F (1.60, 137.66) = 37.35, p < .001). Mean GAD-7 scores at each time point are shown in Figure 4.

Post hoc comparisons revealed a significant difference at all time points (all p < .001) with GAD-7 scores were significantly lower at mid-point compared to baseline, significantly lower at final compared to mid-point, and significantly lower at final compared to baseline.

There was no significant main effect of group (F (2, 86) = .043, p = .958). However, there was a significant interaction effect between time and group (F (3.20, 137.66) = 4.53, p = .004).

A series of one-way ANOVAs revealed there was no significant difference between groups

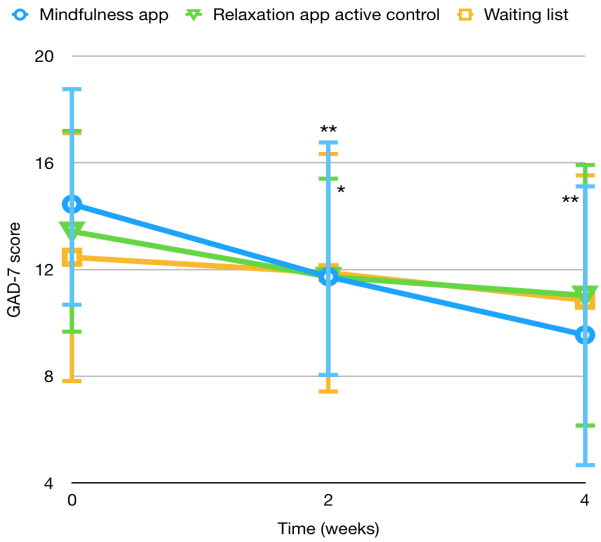

Figure 4. Mean GAD-7 scores over time

** Indicates significance (p < .001)
* Indicates significance (p < .01)

at baseline (F (2, 86) = 1.56, p = .216), mid-point (F (2, 86) = .01, p = .989) or end of the trial (F (2, 86) = .770, p = .466).

A one-way ANOVA was then performed for each group. There was a significant effect of time in the mindfulness app group (F (1.59, 50.94) = 39.51, p < .001). Post hoc comparisons revealed the difference was significant at all time points (all p < .001) with GAD-7 scores being significantly lower at mid-point compared with baseline, significantly lower at final compared with mid-point, and significantly lower at final compared with baseline. Cohen's d indicates a medium effect size (d = 0.57).

There was a significant effect of time in the relaxation app active control group (F (1.43, 41.48) = 9.29, p = .001). Post hoc comparisons revealed GAD-7 scores were significantly lower at mid-point compared to baseline (p = .004), and significantly lower at final compared to baseline (p = .008). However, there was no significant difference at final compared to mid-point (p = .436). Cohen's d indicates a medium effect size (d = 0.54).

There was no significant effect of time in the waiting list group (F (2, 50) = 2.97, p = .061).

Mood

A factorial ANOVA revealed a significant main effect of time on PHQ-9 scores between the start and end of the trial (F (1.73, 143.55) = 26.59, p < .001). Mean PHQ-9 scores at each time point are shown in Figure 5.

Post hoc comparisons revealed a significant difference at all time points with PHQ-9 scores being significantly lower at the mid-point compared to baseline (p = .021), significantly lower at final compared to mid-point (p < .001), and significantly lower at final compared to baseline (p < .001).

There was no significant main effect of group (F (2, 83) = .599, p = .552). However, there was a significant interaction effect between time and group (F (3.46, 143.55) = 4.06, p = .006).

A series of one-way ANOVAs revealed there was no significant difference between groups at baseline (F (2, 85) = .931, p = .398), mid-point (F (2, 84) = 1.12, p = .331) or end of the trial (F (2, 85) = 1.76, p = .179).

A one-way ANOVA was then performed for each group. There was a significant effect of time in the mindfulness app group (F (1.39, 44.31) = 27.50, p < .001). Post hoc comparisons revealed the difference was significant at all time points with PHQ-9 scores being significantly lower at mid-point compared with baseline (p = .015), significantly lower at final compared with mid-point (p < .001), and significantly lower at final

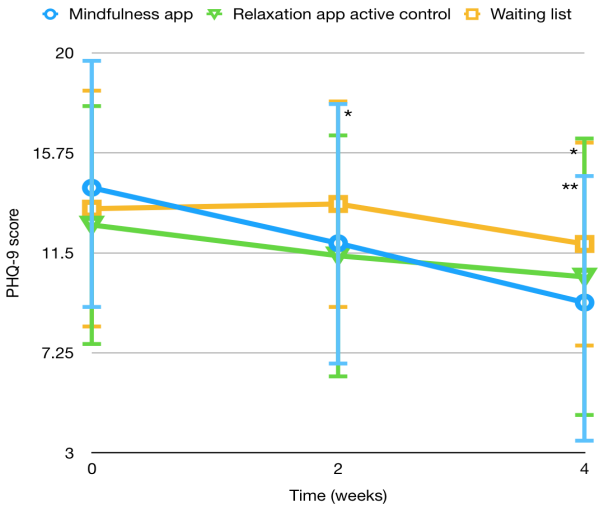

Figure 5. Mean PHQ-9 scores over time

** Indicates significance (p < .001)
* Indicates significance (p < .05)

compared with baseline ($p < .001$). Cohen's d indicates a large effect size ($d = 0.89$).

There was a significant effect of time in the relaxation app active control group (F (2, 56) = 4.75, p = .012). Post hoc comparisons revealed there was no significant difference at mid-point compared to baseline (p = .183) or at final compared to mid-point (p = .494). However, there was a significant difference between baseline and final (p = .042) such that PHQ-9 scores were lower at final compared to baseline. Cohen's d indicates a small effect size ($d = 0.40$).

There was a significant effect of time in the waiting list group (F (2, 46) = 4.45, p = .017). Post hoc comparisons revealed there was no significant difference at mid-point compared to baseline ($p > .999$) or final compared to baseline (p = .096). However, there was a significant difference at final compared to mid-point (p = .025) such that PHQ-9 scores were lower at final compared to mid-point. Cohen's d indicates a small effect size ($d = 0.34$).

Engagement and effect size

The difference between baseline and final GAD-7 and PHQ-9 scores for each participant were calculated, and these were then correlated with the number of days each participant was active on the app.

There was no significant relationship between improvement in scores and days active. Pearson's r (two-tailed) scores are shown in Table 2.

Discussion

This is the first study to examine whether mindfulness-based mHealth apps can reduce symptoms of GAD in the relevant population. It was hypothesised that participants in the mindfulness group would experience a greater reduction in symptoms of GAD and the results support this hypothesis.

To summarise, both the mindfulness app and the relaxation app active control group experienced a significant reduction in anxiety at the mid-point. This reduction continued in the mindfulness app group but did not continue in the relaxation app active control group. The waiting list group saw no significant reduction in anxiety.

Participants in the mindfulness app group experienced a significant improvement in mood

Table 2. Pearson's r scores showing correlation between improvement in scores and number of active days

	r	p	n	R^2
Mindfulness app				
GAD-7	-.10	.587	30	.001
PHQ-9	-.15	.419	30	.021
Relaxation app active control				
GAD-7	.27	.153	30	.072
PHQ-9	.27	.156	30	.071

at all time points. Participants in the relaxation app active control group also experienced this, but at a slower rate, only being significant between the start and end. Participants in the waiting list group experienced no significant improvement in mood between baseline and final but did experience an improvement between mid-point and final.

These findings suggest that mindfulness has a sustained benefit above and beyond relaxation. Both mindfulness and relaxation apps have the potential to decrease the impact of anxiety and affective disorders, but that mindfulness may produce more rapid results and a larger clinical effect.

Given there was no significant difference in levels of engagement across the two app groups, it cannot be that mindfulness was simply a more engaging and acceptable intervention. In both cases, large numbers of users were active at the start of the trial (see Figure 3) but usage then declined. Despite this, the mindfulness group continued to experience the benefit.

One explanation could be that mindfulness produces long-term changes in the brain whereas relaxing music only provides a short-term benefit. Early evidence from neuroscience suggests mindfulness can induce changes in the brain's resting state (Tang & Posner, 2012) even in the early stages of practice (Berkovich-Ohana, Glicksohn, & Goldstein, 2012).

An alternative explanation could be that participants continued to use the mindfulness techniques without accessing the app. Several participants suggested they had done this via email correspondence.

The uplift in mood by the waiting list at the end of the study could be explained by the anticipation of shortly receiving access to the app. Such an effect was demonstrated by Cludius, Schröder, & Moritz (2018) and could also explain why attrition rates in the waiting list group was lower.

These findings broadly agree with existing literature, supporting the idea that mindfulness out-performs relaxation as tested by Jain et al. (2007) using in-person sessions, that mindfulness-based interventions are effective in treating GAD (Khoury et al., 2013), and that mHealth can play a part in reducing symptoms of anxiety (Firth et al., 2017).

The current study found significant results where studies of some other mHealth apps have not. Pham et al. (2016) hypothesised that "Flowy" was too short and lightweight to change clinical symptoms while Enock et al. (2014)'s findings agree with the wider literature that self-directed attention bias modification apps are largely ineffective (Zhang, Ying, Song, Fung, & Smith, 2018). In contrast, mindfulness has a strong evidence base (Keng, Smoski, & Robins, 2011) and is simple to translate into an app.

Although "DeStressify" (Lee & Jung, 2018) did not significantly reduce symptoms of anxiety, it did significantly reduce trait anxiety and improve general health and emotional well-being. As their sample was composed of undergraduate students, one could suggest that the technology does work but that there was simply no GAD to treat.

This would be consistent with the findings of "MoodPrism" (Bakker & Rickard, 2018)

who found that engagement with the app significantly predicted a reduction in anxiety scores for the clinical sample but not the non-clinical sample. The authors offer two possible explanations for this. First, the clinical sample may be more motivated to engage. Second, the mechanism of action was to increase emotional self-awareness, which is a predictor of positive affect (O'Toole, Jensen, Fentz, Zachariae, & Hougaard, 2014). Low emotional self-awareness is associated with both anxiety and depression (Suveg, Hoffman, Zeman, & Thomassin, 2009) and, therefore, increasing emotional self-awareness may lead to a greater magnitude of improvement in clinical samples compared with nonclinical samples.

Increased motivation is unlikely to account for the results of the current study because the severity of GAD symptoms did not correlate with engagement. However, as mindfulness is a closely related concept to emotional self-awareness (Sutton, 2016), this could account for the clinical population of the current study showing improvements in GAD symptoms when the "DeStressify" nonclinical population did not.

There was no correlation between engagement levels and effect size. Historically, research has suggested a strong link between practice time and benefit (Carmody & Baer, 2008; Shapiro et al., 2008; Stanley et al., 2011). However, recent research has suggested that brief mindfulness training sessions can also provide a benefit (Banks, Welhaf, & Srour, 2015; Zeidan, Johnson, Gordon, & Goolkasian, 2010).

There are several possible explanations. One is that our sample size was too small to detect the difference between engagement levels. Only four participants in the mindfulness app group used the app for more than half of the days of the study, with the average usage being 8.52 days. Therefore, we have very little data about the effects of using the app every day.

Second, studies looking at brief mindfulness interventions have typically found improvements in specific areas while finding no effects in others. Zeidan, Johnson, Diamond, David, & Goolkasian (2010) found no improvement in mood but a significant decrease in anxiety. It could be that practice produces diminishing returns in reducing symptoms of GAD, hence the benefit being front-loaded, while the wider benefits of mindfulness, which were not measured in the current study, continue to accumulate with further practice.

Dropout rates were high. Only 20.1% of participants provided sufficient data to be included in the analysis, compared to 43.3% of the waiting list group. This suggests that while the intervention may be efficacious, it may not be effective because the majority of people would not persist with using the app. The current leading treatment for anxiety, cognitive behavioural therapy, has a completion rate of 69.0% (Fernandez, Salem, Swift, & Ramtahal, 2015).

Other studies had higher retention rates including "Flowy", 54.8%, "myCompass", 52.1% and "DeStressify", 82.7%. However, these rates are not out-of-line with apps in the wider community. Chen (2016) noted that the average app in the Google Play Store lost 77% of their daily active users within the first three

days, rising to 90% after 30 days.

The difference between the current study and "DeStressify" is perhaps easiest to explain. The current study used a population with GAD, of which non-adherence is consistently high (Santana & Fontenelle, 2011).

It could be that participants were looking for something novel. Any mention of mindfulness was specifically omitted in describing the study to participants so that they would be blind to the two conditions. Participants expecting something new may have been disappointed when they discovered they were given either mindfulness meditations or relaxing music, both of which are already widely available (Handel, 2011) and so may have been tried before, or been put off by a poor experience (Mani, Kavanagh, Hides, & Stoyanov, 2015).

Another potential explanation for this could be that there is an expectation of mindfulness being relaxing (Dunford & Thompson, 2010), when, in reality, asking people with anxiety to sit with their thoughts can be an unpleasant and uncomfortable experience (Cebolla i Martí, Demarzo, Martins, Soler, & García Campayo, 2017).

In email correspondence, several users indicated they forgot to use the app and would like push notification reminders. "myCompass" offered SMS reminders. The current study offered email reminders, but most went unread (average open rate was 48.2%) and were interrupted at one point when the university's spam filtering system disabled the university email account being used to send the emails.

Limitations

Although the study assessed the effects of mindfulness apps on anxiety, it only did so over a four-week period. Follow-up assessments would aid understanding of whether the benefits are maintained. Similarly, it could be that four weeks is not enough time to learn mindfulness if a participant is naive. Length of study and effect size seems to correlate (Creswell, 2017) so it may be that a longer period of app usage would produce more powerful results.

Participants were recruited using volunteer sampling, which could produce systematic bias (Kekkonen et al., 2015). British participants may have different cultural values, may be more or less comfortable with the idea of mental health and may be more or less familiar with the idea of meditation (Woodall, Morgan, Sloan, & Howard, 2010) compared to other countries. Men were also under-represented in the study, making up 37% of anxiety disorders (McLean, Asnaani, Litz, & Hofmann, 2011) but just 8% of the study participants. The majority of participants were long-term anxiety sufferers (more than 12 months, 89%) so it may be more or less effective for recently diagnosed people or people suffering acute episodes rather than being trait anxious (Roy-Byrne, 2015).

The high attrition rates could have caused systematic bias in the results. Analysis of the dropouts found no significant correlation with demographics or baseline GAD-7 and PHQ-9 scores. However, it could be that participants dropped out because their symptoms became worse.

Directions for future research

Future research should consider how to get people to engage with mHealth mindfulness apps. The current study suggests adherence is a significant problem and existing literature suggests that adherence rates would be far lower outside of an academic study setting (Christensen, Griffiths, & Farrer, 2009). Therefore, understanding how to boost engagement levels is essential to developing effective interventions.

Conclusion

Self-guided mHealth mindfulness apps are an efficacious way of reducing the symptoms of GAD. However, more research is needed to understand how to increase engagement levels.

References

Alonso, J., Liu, Z., Evans-Lacko, S., Sadikova, E., Sampson, N., Chatterji, S., ... & Bruffaerts, R. (2018). Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries. *Depression and anxiety*.

Bakker, D., & Rickard, N. (2018). Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *Journal of affective disorders*, *227*, 432-442.

Banks, J. B., Welhaf, M. S., & Srour, A. (2015). The protective effects of brief mindfulness meditation training. *Consciousness and cognition*, *33*, 277-285.

Barney, L. J., Griffiths, K. M., Jorm, A. F., & Christensen, H. (2006). Stigma about depression and its impact on help-seeking intentions. *Australian & New Zealand Journal of Psychiatry*, *40*(1), 51-54.

Bennion, M. R., Hardy, G., Moore, R. K., & Millings, A. (2017). E-therapies in England for stress, anxiety or depression: what is being used in the NHS? A survey of mental health services. *BMJ open*, *7*(1), e014844.

Berkovich-Ohana, A., Glicksohn, J., & Goldstein, A. (2012). Mindfulness-induced changes in gamma band activity–implications for the default mode network, self-reference and attention. *Clinical Neurophysiology*, *123*(4), 700-710.

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., ... & Devins, G. (2004). Mindfulness: A proposed operational definition. *Clinical psychology: Science and practice*, *11*(3), 230-241.

Black, D. S., & Slavich, G. M. (2016). Mindfulness meditation and the immune system: a systematic review of randomized controlled trials. *Annals of the New York Academy of Sciences*, *1373*(1), 13-24.

Carmody, J., & Baer, R. A. (2008). Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. *Journal of behavioral medicine*, *31*(1), 23-33.

Cebolla i Martí, A. J., Demarzo, M., Martins, P., Soler, J., & García Campayo, J. (2017). Unwanted effects: Is there a negative side of meditation? A multicentre survey. *Plos One, 2017, vol. 12, num. 9*.

Chen, A. (2016, October 22). New data shows losing 80% of mobile users is normal, and why the best apps do better. Retrieved September 4, 2018, from https://andrewchen.co/new-data-shows-why-losing-80-of-your-mobile-users-is-normal-and-that-the-best-apps-do-much-better/

Christensen, H., Griffiths, K. M., & Farrer, L. (2009). Adherence in internet interventions for anxiety and depression: systematic review. *Journal of medical Internet research*, *11*(2).

Cludius, B., Schröder, J., & Moritz, S. (2018). Expectancy Effects in Self-Help Depression Treatment: First Evidence that the Rationale Given for an Online Study Impacts the Outcome. *Behavioural and cognitive psychotherapy*, *46*(2), 195-208.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 2nd.

Collins, P. H., and Bilge, S. (2016). *Intersectionality*. John Wiley & Sons.

Creswell, J. D. (2017). Mindfulness interventions. *Annual review of psychology*, *68*, 491-516.

Davis, D. M., & Hayes, J. A. (2011). What are the benefits of mindfulness? A practice review of psychotherapy-related research. *Psychotherapy*, *48*(2), 198.

Dunford, E., & Thompson, M. (2010). Relaxation and mindfulness in pain: A review. *Reviews in pain*, *4*(1), 18-22.

Enock, P. M., Hofmann, S. G., & McNally, R. J. (2014). Attention bias modification training via smartphone to reduce social anxiety: A randomized, controlled multi-session experiment. *Cognitive therapy and research*, *38*(2), 200-216.

Evans, S., Ferrando, S., Findler, M., Stowell, C., Smart, C., & Haglin, D. (2008). Mindfulness-based cognitive therapy for generalized anxiety disorder. *Journal of anxiety disorders*, *22*(4), 716-721.

Fernandez, E., Salem, D., Swift, J. K., & Ramtahal, N. (2015). Meta-analysis of dropout from cognitive behavioral therapy: Magnitude, timing, and moderators. *Journal of Consulting and Clinical Psychology*, *83*(6), 1108.

Firth, J., Torous, J., Nicholas, J., Carney, R., Rosenbaum, S., & Sarris, J. (2017). Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *Journal of affective disorders*, *218*, 15-22.

Gackenheimer, C. (2015). What is react?. In *Introduction to React* (pp. 1-20). Apress, Berkeley, CA.

Griffith, E. (2017, June 30). Headspace Raises $37 Million to Expand Corporate Partnerships, Create More Content, and Go International. *Fortune*. Retrieved from: http://fortune.com/2017/06/30/headspace-funding/

Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clinical psychology review*, *37*, 1-12.

Gyani, A., Shafran, R., Layard, R., & Clark, D. M. (2013). Enhancing recovery rates: lessons from year one of IAPT. *Behaviour Research and Therapy*, *51*(9), 597-606.

Handel, M. J. (2011). mHealth (mobile health)—using apps for health and wellness. *The Journal of Science and Healing*, *7*(4), 256-261.

Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of consulting and clinical psychology*, *78*(2), 169.

Hölzel, B. K., Lazar, S. W., Gard, T., Schuman-Olivier, Z., Vago, D. R., & Ott, U. (2011). How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. *Perspectives on psychological science*, *6*(6), 537-559.

Ivanova, E., Lindner, P., Ly, K. H., Dahlin, M., Vernmark, K., Andersson, G., & Carlbring, P. (2016). Guided and unguided Acceptance and Commitment Therapy for social anxiety disorder and/or panic disorder provided via the Internet and a smartphone application: a randomized controlled trial. *Journal of anxiety disorders*, *44*, 27-35.

Jain, S., Shapiro, S. L., Swanick, S., Roesch, S. C., Mills, P. J., Bell, I., & Schwartz, G. E. (2007). A randomized controlled trial of mindfulness meditation versus relaxation training: effects on distress, positive states of mind, rumination, and distraction. *Annals of behavioral medicine*, *33*(1), 11-21.

Jha, A. P., Morrison, A. B., Parker, S. C., & Stanley, E. A. (2017). Practice is protective: mindfulness training promotes cognitive resilience in high-stress cohorts. *Mindfulness*, *8*(1), 46-58.

Kabat-Zinn, J. (2009). *Wherever you go, there you are: Mindfulness meditation in everyday life*. Hachette Books.

Kekkonen, V., Kivimäki, P., Valtonen, H., Hintikka, J., Tolmunen, T., Lehto, S. M., & Laukkanen, E. (2015). Sample selection may bias the outcome of an adolescent mental health survey: results from a five-year follow-up of 4171 adolescents. *Public health*, *129*(2), 162-172.

Kendrick, T., & Pilling, S. (2012). Common mental health disorders—identification and pathways to care: NICE clinical guideline. *Br J Gen Pract*, *62*(594), 47-49.

Keng, S. L., Smoski, M. J., & Robins, C. J. (2011). Effects of mindfulness on psychological health: A review of empirical studies. *Clinical psychology review*, *31*(6), 1041-1056.

Khoury, B., Lecomte, T., Fortin, G., Masse, M., Therien, P., Bouchard, V., ... & Hofmann, S. G. (2013). Mindfulness-based therapy: a comprehensive meta-analysis. *Clinical psychology review*, *33*(6), 763-771.

Kim, Y. W., Lee, S. H., Choi, T. K., Suh, S. Y., Kim, B., Kim, C. M., ... & Song, S. K. (2009). Effectiveness of mindfulness-based cognitive therapy as an adjuvant to pharmacotherapy in patients with panic disorder or generalized anxiety disorder. *Depression and anxiety*, *26*(7), 601-606.

Kolovos, S., Dongen, J. M., Riper, H., Buntrock, C., Cuijpers, P., Ebert, D. D., ... & Warmerdam, L. (2018). Cost effectiveness of guided Internet-based interventions for depression

in comparison with control conditions: An individual–participant data meta-analysis. *Depression and anxiety*.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The phq-9. *Journal of general internal medicine*, *16*(9), 606-613.

Lee, R. A., & Jung, M. E. (2018). Evaluation of an mHealth App (DeStressify) on University Students' Mental Health: Pilot Trial. *JMIR mental health*, *5*(1), e2.

Levin, M. E., Haeger, J., Pierce, B., & Cruz, R. A. (2017). Evaluating an adjunctive mobile app to enhance psychological flexibility in acceptance and commitment therapy. *Behavior modification*, *41*(6), 846-867.

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).

Matsumoto, M., & Nishimura, T. (1998). Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, *8*(1), 3-30.

McLean, C. P., Asnaani, A., Litz, B. T., & Hofmann, S. G. (2011). Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. *Journal of psychiatric research*, *45*(8), 1027-1035.

Miner, A., Kuhn, E., Hoffman, J. E., Owen, J. E., Ruzek, J. I., & Taylor, C. B. (2016). Feasibility, acceptability, and potential efficacy of the PTSD Coach app: A pilot randomized controlled trial with community trauma survivors. *Psychological Trauma: Theory, Research, Practice, and Policy*, *8*(3), 384.

Mukuria, C., Brazier, J., Barkham, M., Connell, J., Hardy, G., Hutten, R., ... & Parry, G. (2013). Cost-effectiveness of an improving access to psychological therapies service. *The British Journal of Psychiatry*, *202*(3), 220-227.

Nuffield Trust. Categories of NHS spending per head. 14 November 2014.

Obbarius, A., van Maasakkers, L., Baer, L., Clark, D. M., Crocker, A. G., de Beurs, E., ... & Langford, L. (2017). Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM Depression and Anxiety Working Group. *Quality of Life Research*, *26*(12), 3211-3225.

Ofcom. Communications Market Report – United Kingdom. 3 August 2017.

O'Toole, M. S., Jensen, M. B., Fentz, H. N., Zachariae, R., & Hougaard, E. (2014). Emotion differentiation and emotion regulation in high and low socially anxious individuals: An experience-sampling study. *Cognitive Therapy and Research*, *38*(4), 428-438.

Pham, Q., Khatib, Y., Stansfeld, S., Fox, S., & Green, T. (2016). Feasibility and efficacy of an mHealth game for managing anxiety:"flowy" randomized controlled pilot trial and design evaluation. *Games for health journal*, *5*(1), 50-67.

Piet, J., Hougaard, E., Hecksher, M. S., & Rosenberg, N. K. (2010). A randomized pilot study of mindfulness-based cognitive therapy and group cognitive-behavioral therapy for young adults with social phobia. *Scandinavian*

Journal of Psychology, *51*(5), 403-410.

Price, M., Yuen, E. K., Goetter, E. M., Herbert, J. D., Forman, E. M., Acierno, R., & Ruggiero, K. J. (2014). mHealth: a mechanism to deliver more accessible, more effective mental health care. *Clinical psychology & psychotherapy*, *21*(5), 427-436.

Proudfoot, J., Clarke, J., Birch, M. R., Whitton, A. E., Parker, G., Manicavasagar, V., ... & Hadzi-Pavlovic, D. (2013). Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. *BMC psychiatry*, *13*(1), 312.

Roy-Byrne, P. (2015). Treatment-refractory anxiety; definition, risk factors, and treatment challenges. *Dialogues in clinical neuroscience*, *17*(2), 191.

Santana, L., & Fontenelle, L. F. (2011). A review of studies concerning treatment adherence of patients with anxiety disorders. *Patient preference and adherence*, *5*, 427.

Seo, J. G., & Park, S. P. (2015). Validation of the Generalized Anxiety Disorder-7 (GAD-7) and GAD-2 in patients with migraine. *The journal of headache and pain*, *16*(1), 97.

Shapiro, S. L., Carlson, L. E., Astin, J. A., & Freedman, B. (2006). Mechanisms of mindfulness. *Journal of clinical psychology*, *62*(3), 373-386.

Shapiro, S. L., Oman, D., Thoresen, C. E., Plante, T. G., & Flinders, T. (2008). Cultivating mindfulness: effects on well-being. *Journal of clinical psychology*, *64*(7), 840-862.

Singh, S. K., & Gorey, K. M. (2017). Relative effectiveness of mindfulness and cognitive behavioral interventions for anxiety disorders: meta-analytic review. *Social Work in Mental Health*, 1-14.

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.

Stanley, E. A., Schaldach, J. M., Kiyonaga, A., & Jha, A. P. (2011). Mindfulness-based mind fitness training: A case study of a high-stress predeployment military cohort. *Cognitive and Behavioral Practice*, *18*(4), 566-576.

Sucala, M., Cuijpers, P., Muench, F., Cardoş, R., Soflau, R., Dobrean, A., ... & David, D. (2017). Anxiety: There is an app for that. A systematic review of anxiety apps. *Depression and anxiety*.

Sutton, A. (2016). Measuring the Effects of Self-Awareness: Construction of the Self-Awareness Outcomes Questionnaire. *Europe's journal of psychology*, *12*(4), 645.

Suveg, C., Hoffman, B., Zeman, J. L., & Thomassin, K. (2009). Common and specific emotion-related predictors of anxious and depressive symptoms in youth. *Child Psychiatry and Human Development*, *40*(2), 223.

Tang, Y. Y., & Posner, M. I. (2012). Tools of the trade: theory and method in mindfulness neuroscience. *Social cognitive and affective neuroscience*, *8*(1), 118-120.

Watts, S., Mackenzie, A., Thomas, C., Griskaitis, A., Mewton, L., Williams, A., & Andrews, G. (2013). CBT for depression: a pilot RCT

comparing mobile phone vs. computer. *BMC psychiatry*, *13*(1), 49.

Woodall, A., Morgan, C., Sloan, C., & Howard, L. (2010). Barriers to participation in mental health research: are there specific gender, ethnicity and age related barriers?. *BMC psychiatry*, *10*(1), 103.

Zeidan, F., Johnson, S. K., Diamond, B. J., David, Z., & Goolkasian, P. (2010). Mindfulness meditation improves cognition: Evidence of brief mental training. Consciousness and cognition, 19(2), 597-605.

Zeidan, F., Johnson, S. K., Gordon, N. S., & Goolkasian, P. (2010). Effects of brief and sham mindfulness meditation on mood and cardiovascular variables. *The Journal of Alternative and Complementary Medicine*, *16*(8), 867-873.

Zeidan, F., & Vago, D. R. (2016). Mindfulness meditation–based pain relief: a mechanistic account. *Annals of the New York Academy of Sciences*, *1373*(1), 114-127.

Zhang, M., Ying, J., Song, G., Fung, D. S., & Smith, H. (2018). Attention and Cognitive Bias Modification Apps: Review of the Literature and of Commercially Available Apps. *JMIR mHealth and uHealth*, *6*(5).