Date: 22 March 2022; Swansea University (UK) & Université Grenoble Alpes (France)

Self-administered mindfulness interventions reduce stress in a large, randomized controlled multi-site study

Alessandro Sparacio[1,2], Hans IJzerman[1,3], Ivan Ropovik[4,5], Filippo Giorgini[6], Christoph Spiessens[7], [authors that contributed with the data collection with at least 70 participants], Gabriela M. Jiga-Boy[2]

[1]Université Grenoble Alpes, [2]Swansea University, [3]Institut Universitaire de France, [4]Charles University in Prague, [5]Presov University, [6]University of Milano-Bicocca, [7]Christoph Spiessens Coaching Solutions Ltd.

This project was pre-registered on March 22th, 2022. For any further changes, please contact Alessandro Sparacio

Author Note: Note that we wrote the pre-registration in past tense to avoid errors after finishing the manuscript. Our OSF page can be found here (https://osf.io/6w2zm/)

Abstract

The use of self-administered mindfulness interventions has increased in recent years. The effectiveness of these interventions on regulating stress/emotions, however, is debated. In the present multi-site study (N_{sites} = x, $N_{participants}$ = x), we aimed to investigate the effectiveness of four single, brief stand-alone mindfulness exercises in a population unfamiliar with mindfulness meditation. We tested these four interventions in comparison to non-mindful active control conditions using an adaptive Bayesian design. We found [evidence for the efficacy of x exercises/no evidence for the efficacy of x exercises] with an estimated mean effect size of [xx/xx]. This means that… or We recommend that… [recommendation will be provided].

A large, multisite test of self-administered mindfulness

Engaging in mindfulness meditation, in theory, appears simple: One is asked to sit and focus one's attention on the breath and on the present moment, without needing complex settings or apparatus. Partly due to this apparent simplicity, mindfulness meditation protocols that can be self-administered have increased in accessibility and popularity in recent years (Cavanagh et al., 2018).

Self-administered mindfulness protocols are free or less expensive compared to traditional protocols (Spijkerman, Pots, & Bohlmeijer, 2016) and readily available (Cavanagh et al. 2014; Wahbeh et al. 2014). These mindfulness interventions can be accessed as self-help books, or via computer programmes, smartphone apps, or audio and video recordings. For example, the last few years witnessed a surge in downloads of mindfulness applications such as 'Calm', the top English-language meditation app reportedly with 40 million downloads (Gebel, 2019), and 'Headspace', reportedly with 70 million downloads (Headspace, 2021). Features such as the duration or frequency of these interventions are less defined, because while the applications might offer modular sessions of varying lengths, the onus is on the consumer to decide for how long or how often to engage with it.

Despite their easy accessibility, the evidence behind these mindfulness interventions is debated and some questions remain unanswered, such as: Are these types of intervention truly effective in regulating individuals' stress levels? or Which types of self-administered mindfulness exercises, from the plethora of those available, work to downregulate stress? We attempted to answer these questions first by conducting a survey aimed at mindfulness practitioners in order to know which are the most used or the most popular mindfulness exercises. Based on the results of the survey, we then designed a multi-site, highly powered study to test the effects and the boundary conditions of four types of self-administered mindfulness meditation exercises on stress regulation.

Do self-administered mindfulness interventions reduce levels of stress?

According to Jon Kabat-Zinn (1992, 2003), mindfulness can be defined as "paying attention in a particular way: On purpose, in the present moment, and nonjudgmentally". Mindfulness emphasizes attention to the present moment, with awareness of one's bodily sensations or one's mental content (thoughts, emotions, memories, etc.). Self-administered mindfulness interventions share these features, but compared to traditional established mindfulness protocols (e.g., mindfulness-based stress reduction - MBSR or mindfulness-based cognitive therapy - MBCT), do not require the presence of an instructor or and the commitment to be in a particular place and time regularly, for 8 weeks, allowing practitioners to meditate whenever they choose so (Spijkerman, Pots, & Bohlmeijer, 2016).

Because self-administered mindfulness interventions hold promises for reducing stress levels and are accessible to virtually anyone, it is important to understand whether they bring about the expected results. While some studies showed positive results of self-administered mindfulness interventions on reducing self-reported stress (Cavanagh et al., 2013; Cavanagh et al., 2018), and a recent meta-analysis pointed in the same direction (Taylor et al., 2021), other studies reported less positive outcomes. For instance, Glück and Maercker (2011) did not find evidence that a self-guided mindfulness training effectively decreased levels of perceived stress and distress for a group of meditators compared to a waitlist control group. Finally, a meta-analysis failed to find sufficient evidence regarding the efficacy of self-administered mindfulness interventions on stress reduction after applying publication bias techniques (Sparacio et al., 2022).

Which types of self-administered mindfulness exercises work to downregulate stress?

Another reason why we do not know yet whether self-administered mindfulness interventions are efficacious at regulating stress is that self-administered mindfulness include a range of different exercises (e.g., body scan, mindful breathing, mindful walking, etc.);

current tests evaluate protocols in their entirety, instead of differentiating among these exercises. According to Germer et al. (2016), there are a plethora of mindfulness exercises that can be divided into three different categories: 'awareness', 'present experience', and 'acceptance'. These different categories are not truly separate, but yet, splitting them into three different groups allows us to better understand the potential applied value of certain mindfulness exercises that can be self-administered.

'Awareness' mindfulness exercises are typically characterized by a series of steps: 1) Disengage from an automatic train of thoughts (e.g., pausing a discussion with a friend by taking a long breath), 2) focus the attention on an object that is used as an "anchor" (e.g., bringing the attention to the breath or to parts of the body), and 3) return the attention to the object of focus when one realizes they have been distracted, and then watch where the mind wanders next.

'Present experience' mindfulness exercises instruct participants to pay attention completely to the activity being carried out (e.g., bringing the attention to the sole of the foot while walking). If the mind wanders, the attention is redirected to the present moment and the instructions that are given in this set of exercises have the function of fulfilling this objective.

'Acceptance' mindfulness exercises are characterized by applying a non-judgemental attitude of kindness and curiosity to one's experience. Practitioners that do these types of exercises are invited to cultivate positive feelings towards themselves and to others (e.g., directing loving-kindness to themselves or to someone else).

We currently do not know which category of the exercises described above works best for reducing stress levels. Some studies have tested single brief mindfulness exercises (Feldman et al., 2010; Hutcherson et al., 2008). To our knowledge, however, none investigated the effectiveness of brief stand-alone mindfulness exercises on stress reduction. Thus, the aim of our multi-site project is to investigate which self-administered mindfulness

exercises were more effective in reducing individuals' stress levels, compared to three non-mindful active control conditions randomly sampled.

Study 1: Survey on self-administered mindfulness

To identify the most representative mindfulness exercises to use in our multi-site project, we conducted a survey to ask mindfulness practitioners for their suggestions regarding self-administered mindfulness exercises. that provided a list of the 20 most popular meditation techniques. To compile a final list of exercises, we retained the most popular exercises suggested by the practitioners taking part in the survey that were also included in peer-reviewed literature (Matko et al., 2019).

Procedure

We disseminated a Qualtrics survey to mindfulness practitioners (i.e., instructors and/or researchers) that we retrieved from lists of subscribers of various mindfulness associations and via social media networks (Facebook, Twitter, & LinkedIn) between November 17th and December 31th, 2020. Ninety-five mindfulness practitioners took part in the survey, of the total 40 completed half of it, and 20 respondents arrived at the end.

The survey was divided into two parts: The first part asked respondents to suggest low-intensity self-administered mindfulness interventions designed to decrease levels of stress with a short term efficacy span, or 'state-level' stress, while the second part asked them to suggest high-intensity self-administered mindfulness interventions that can be aimed at changing personality traits that can impact on stress levels (e.g., decreasing 'trait-level' stress). A low-intensity self-administered mindfulness intervention was defined as a "mindfulness intervention that is brief, but long enough to have an impact on stress levels of individuals [and] give participants flexibility (to ensure that they do not abandon the study), while being effective in reducing stress". We did not expect this low-intensity intervention to substantially impact on traits of participants that predisposes them to higher levels of stress

(e.g., trait anxiety, tendency to ruminate). A high-intensity self-administered mindfulness intervention was defined as a "mindfulness intervention that is longer, less flexible, but that might impact not only on stress levels of participants, but also on personality traits of participants that are relevant to the buildup of stress (e.g., trait anxiety, tendency to ruminate)."

We asked the same set of questions for both low- and high-intensity self-administered mindfulness interventions, namely: 1) How long the intervention should last, in order to observe a significant reduction of participants' stress level (*from 1 day to 8 weeks or more*), 2) How many times a week participants should do exercises of self-administered mindfulness? (*1 to 7/week*), 3) Which type of source they thought would be more effective for delivering the intervention (*smartphone app, video file, audio file, self-help book, another type of source - specify which, or 'the source is not relevant'*), and 4) What types of exercises should be included in the intervention (e.g., body scan, focus on breathing, etc.). For each exercise they suggested, we asked them 1) If the order of the exercises they listed in the previous question was important to achieve stress reduction (*yes, maybe, no*), and 2) How long they thought it should be practiced and how often (*length* - or daily duration of the exercise in minutes; *weekly frequency*; *number of weeks of the intervention*).

Results

Results of our survey showed that respondents: 1) estimated that smartphone applications were the most effective types of self-administered mindfulness exercise for reducing stress levels as compared to other cited sources (e.g., books), 2) considered that the order of the mindfulness exercises could potentially be relevant to impact levels of stress[1] and 3) assessed that practicing mindfulness daily had the greatest impact on stress reduction. A full description of the results is beyond the scope of this manuscript, however a R script with

[1] 18 respondents answered "maybe" to this question, 14 answered "yes" and 7 answered "no".

the answers of the respondents and a complete report of the survey in PDF can be found on the OSF page of the project (https://osf.io/aw2vz/).

For planning the multi-site study, we relied on practitioners' answers that included a list of exercises to be considered for a low-intensity self-administered mindfulness intervention[2]. The list compiled practitioners' suggestions and consisted, initially, of 174 mindfulness exercises (see the list of the exercises document at https://osf.io/br86s/). We then removed duplicates and we merged the techniques that we considered to be similar[3] (e.g., 'mindful walking' and 'walk in mindfulness') or that were spelled differently but referred to the same exercise (e.g., 'body scanning' and 'body scan'). We then organized each technique according to its dominant element of mindfulness based on Germer's (2016) system of categorization, which included three categories previously mentioned: 'awareness', 'present experience', and 'acceptance'. Some mindfulness exercises suggested by the practitioners were already categorized according to Germer's system (e.g., 'body scan' was placed as belonging to the category 'awareness'). For the exercises that were not already categorized, the first author made an arbitrary choice based on the similarities between the uncategorized exercise and the other exercises already present in that category (e.g., 'mindful movements' was placed in the category 'present experience').

Finally, for each dominant element, we retained the two exercises that were the most cited: Body scan (39 votes) and mindful breathing (37 votes) for 'awareness', mindful walking (10 votes) and mindful movement (10 votes) for 'present experience', and self-

[2] When we designed the survey, we had not yet outlined the details of the multi-site project, so we asked for general information related to self-administered mindfulness protocols. For the purpose of the multi-site project we used only a portion of all the information provided by the answers of mindfulness practitioners, however we believe that the unused information may help other mindfulness researchers design more rigorous studies on the topic.

[3] This process was done by the first author, however it is possible to infer how the simplification was done by comparing the files with all exercises (i.e., List of the exercises) and the one where similar exercises were merged (i.e., Mindfulness clusters MULTI-SITE; see https://osf.io/br86s/).

compassion break (5 votes) and loving kindness meditation (3 votes) for 'acceptance' (See Mindfulness clusters MULTI-SITE at https://osf.io/br86s/).

We then cross-referenced our set of self-administered mindfulness exercises with the list of 20 popular meditation techniques that Matko et al., (2019) detailed by consulting 637 meditators with an online survey. By retaining the exercises that were present in both approaches we arrived at this final list of self-administered mindfulness exercises: Body scan, mindful breathing, mindful walking and loving-kindness meditation. We decided to test these four mindfulness exercises in the multi-site project described below.

Study 2: The multi-site project

Hypothesis

We hypothesized that participants allocated to any one of the experimental (mindfulness) conditions would experience lower levels of self-reported stress compared to participants allocated to an active control conditions[4].

Additionally we explored effects on the dimensions of pleasure, arousal and dominance of any of the experimental conditions as compared to the active control conditions. Similarly, we also explored the role of neuroticism in reducing levels of stress of individuals for the above mentioned conditions.

Methods

Materials

All materials used in the study, including the ethics (IRB) approval documents of all the sites involved in the project and the meditations scripts are available on our OSF page (https://osf.io/6w2zm/), the analysis code instead can be found on the GitHub repository of

[4] Participants who end up in the control group will have an equal chance of listening to an excerpt of one the three stories that we selected for the present study.

the project (https://github.com/alessandro992/A-large-multi-site-test-of-self-administered-mindfulness).

Participants and sites

Because the guided audio meditations and the active control excerpts recorded for this experiment were in English, we limited the study to English native speakers or participants who self-assessed their level of English proficiency at the C1/C2 levels from the Common European Framework of Reference for Languages (Council of Europe, 2001).

Participants were excluded if they reported having or having had a history of mental illnesses, if they declared having meditated in the previous 6 months, or if they were not fluent in English. Each participant was asked to take part in the survey using a smartphone with headphones or earphones attached, not a computer or laptop, because one experimental condition included mindful walking and, participants being randomly allocated to conditions, we needed to ensure that everyone was equally capable to perform any of the mindfulness activities. Each site was asked to collect data from at least 70 participants and up to a maximum of 130 participants[5] in order to be considered for authorship of the article resulting from this experiment.

The data collection will start on March 23th 2022. Sites will be able to begin recruiting participants up to three weeks after the actual data collection begins. Sites that fail to start data collection up to three weeks from the start of the effective date will be excluded from the project.

Bayesian design analysis

Prior to the data collection, we simulated data based on a Bayes Factor Design Analysis (BFDA; Schönbrodt & Wagenmakers, 2018) to assess the expected efficiency and informativeness of the present design. More specifically, the aim of the design analysis was

[5] We allowed a maximum of three people, per site, to become part of the authorship team.

to establish (1) the expected likelihood of the study to provide compelling relative evidence either in favor of H_0 (BF_{01} = 1/10) or H_1 (BF_{10} = 10); (2) the likelihood of obtaining convincing but misleading evidence, and (3) the likelihood that the study points into the correct direction even if stopped earlier due pragmatic constraints on sample size (Schönbrodt & Wagenmakers, 2018).

Given these aims, we modeled a sequential design with maximum *N* where the data collection continues until either the threshold for compelling evidence is met or the maximum *N* is reached. Initially, forty-one labs indicated interest in the project, we expect to have up to 30 data-collecting labs. Each lab was expected to collect data of at least $N = 70$ participants, with a stopping *N* at 120 (translating to $Min = 420$ and $Max = 720$ participants per condition). We intended to be able to detect an effect size of $d = 0.20$; here we modeled the true value to vary between labs by repeatedly (for each simulation) drawing from a normal distribution, $\delta \sim N(0.20, 0.05)$, with a 95% probability that the effect size falls between $d = 0.10$ and 0.30.

The present study aimed at testing the effectiveness of four stand-alone interventions using a between-subjects adaptive group design, where upon hitting a threshold of compelling evidence in one condition, the plan was to allocate the rest of the participants into other conditions where this has not been the case yet. The design analysis, however, assumed a conservative scenario with equal *N* across all conditions, therefore, simplifying the computations to a single between-subjects *t*-test scenario.

The results (see Figure 1) show that, given the assumed design, the probability of the test to arrive at the boundary of compelling evidence ($BF = 10$ or 1/10) was .79 (.72 at H_1 and .07 erroneously at H_0). The probability of terminating at maximum *N* of 720 per condition was .21; .05 of showing some evidence for H_1 ($BF > 3$), .13 of being inconclusive ($3 > BF > 1/3$), and .03 of showing evidence for H_0 ($BF < 1/3$). For the test of a single condition against controls, the sequential design is expected to be 27% more effective than collecting a fixed

maximum *N* per lab, with the average *N* at stopping point (*BF* boundary and maximum *N*) at 526. Even conservatively assuming a balanced-*N* situation, the informativeness of the design thus appears to be adequate and the use of adaptive design is likely to further enhance informativeness and/or resource efficiency. Figure 1 shows the results of the simulation of the sequential design with 10000 iterations.

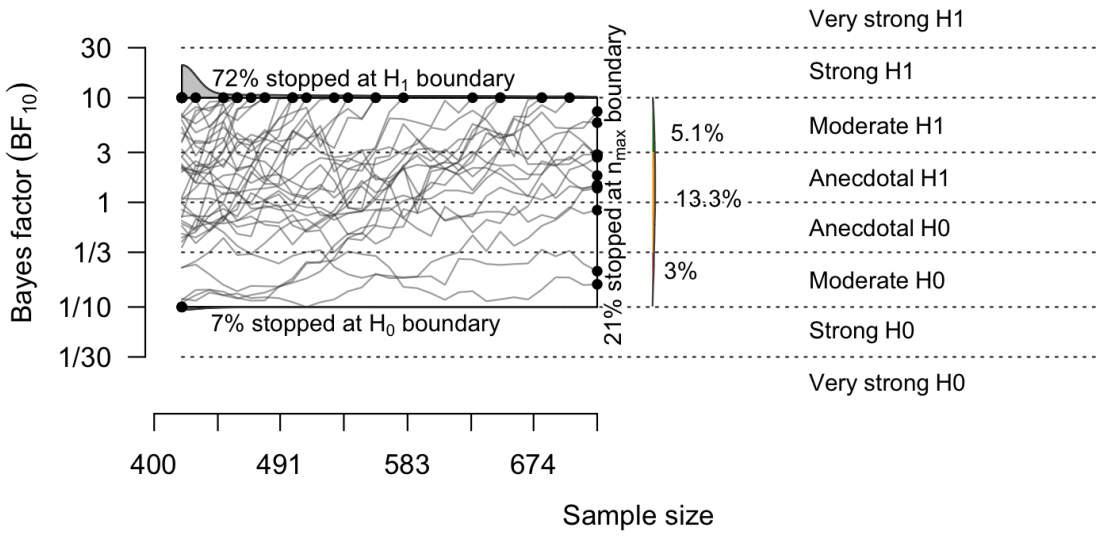

Figure 1. Design analysis for the sequential bayesian design with maximum *N*

Procedure

Participants accessed the experiment via a Qualtrics link. They were provided detailed information about the study (see "Participants Information Sheet" included in the IRB package, https://osf.io/6w2zm/) and were then asked to provide consent prior to participating. Each participant was asked to take part in the survey using a smartphone with headphones or earphones attached, not a computer or laptop, because one experimental condition includes mindful walking and, participants being randomly allocated to conditions, we needed to ensure that everyone would be equally capable to perform the mindfulness activity. To know if this condition was verified we included a question in which we asked participants whether they started the survey from a device other than a smartphone; if they answered negatively,

we asked them to exit the survey and to restart it, this time using a smartphone with headphones or earphones attached.

Participants were also asked to sit in a quiet place such as a room where they could not be disturbed for 20 minutes. After providing informed consent, they were asked to complete the Neuroticism subscale of the IPIP - 5 NEO Domains (IPIP, 2021) to assess the level of neuroticism. Examples of items were "I often feel blue" or "I am filled with doubts about things". Participants could answer a 5-points scale (1 = *Very inaccurate*, 2 = *Moderately Inaccurate*, 3 = *Neither Inaccurate nor Accurate*, 4 = *Moderately Accurate*, 5 = *Very Accurate*). We estimated the scale reliability (i.e., internal consistency) using the omega coefficient. Assuming a unitary-factor model, the reliability estimate in the present sample was ω_u = x.

After the completion of the neuroticism measure, participants were randomly allocated to one of the five conditions[6]. In each condition, participants were presented with a different 15-minutes recording. The four audio files of the mindfulness exercises and the three audio files of the stories of the non-mindful active control condition were recorded by a certified meditator, to keep the voice constant across conditions. Following a brief description of the mindfulness conditions and of the three different stories of the non-mindful active control condition.

Body scan. In body scan, participants were invited to "scan" their parts of the body. Every time the mind wandered, participants were invited to bring back the awareness and the attention to the part of the body they were "scanning".

[6] Participants in the non-mindful active control condition could listen to the excerpt of one of three stories that we selected with an equal chance of getting one story over another.

Mindful breathing. During mindful breathing, participants were invited to "stay with their breath", without changing the way they were breathing. When the mind wandered, participants were told to bring their attention back to the breath with kindness and patience.

Loving kindness-meditation. During this practice, participants were encouraged to direct loving kindness toward themselves and then to extend these feelings of loving kindness towards somebody else.

Mindful walking. The nature of this condition slightly differed from the rest because participants were asked to walk in a quiet place (preferably indoor or in a place as isolated as possible from distractions), while listening to the instructions. During this practice, participants were invited to bring their awareness to the experience of walking and they were invited to "feel" a direct sense of the physical sensations of contact of their feet with the ground.

Non-mindful active control condition. Participants in the control condition had an equal probability of listening to an excerpt coming from one of these three books: "Silverview" (word count: 1838) by John le Carré (Le Carré's, 2021), "the old man and the sea" (word count: 2039) by Ernest Hemingway (Hemingway, 1952) and "Smith of Wootton Major" (word count: 2015) by J. R. R. Tolkien (Tolkien, 1967). These three stories were chosen because they had a similar word count, they were written in an easy to understand English, and did not feature major plot changes so they were unlikely to elicit strong emotions. The full transcript of the seven recordings can be found at our OSF project page (https://osf.io/6w2zm/).

After the 15 minutes of mindfulness exercises or active control listening, participants first answered the State-Trait Anxiety Inventory, Form Y (STAI; Spielberger, Gorshu & Lushene, 1970). Participants were asked to indicate how they felt in that exact moment on a 4-points scale (1 = *Not at all*, 2 = *Somewhat*, 3 = *Moderately so,* 4 = *Very much so*) on 20

items (e.g., “I am tense”; “I feel frightened”). For the present study, we only used the part of the scale that measures state anxiety as we are interested in self-reported stress of individuals with a short-term efficacy span. We calculated the one-factor omega for this scale obtaining a reliability estimate of ωu = x.

Participants then filled in The Self-Assessment Manikin a non-verbal pictorial assessment technique that measures three different dimensions: pleasure, arousal, and dominance in relation to an object or a stimulus (Bradley & Lang, 1994). Participants were invited to observe a series of images and were asked to select the image that best represented their emotional state in that exact moment. This scale demonstrated to have high correlation with scales measuring the same dimensions (i.e., pleasure, arousal, & dominance; Morris, 1995).

Finally, participants answered questions related to their demographic information such as age, gender, country of birth, country of residence, and whether they were students or not. If participants were students, they were asked to indicate in which university they were studying; if participants were not students we inquired about their current job. At the end, participants were thanked and debriefed. Based on the site they belong to, participants were then redirected to another page to receive credits/payments for completing the study.

Analysis Plan

To assess whether any of the chosen mindfulness exercises were more effective than the randomly sampled control conditions to reduce stress in participants, we planned to conduct a sequential design. We sampled participants until a Bayes Factor threshold for compelling evidence was met or until a maximally feasible number of participants was collected (Schönbrodt et al., 2017). Unlike in the frequentist approach, the aim of this procedure was not the control of error rates, but to gauge the relative support provided by the present data in favor of one of two competing hypotheses, H_1 and H_0. In the frequentist

approach, the practice of adjusting the sampling plan after peeking at the data would have increased the Type 1 error rate, but in the Bayesian approach, the data collection could be stopped anytime without having any consequences on the interpretation of the Bayes Factor (Rouder, 2014). We divided our analysis workflow into confirmatory and exploratory. The sequential testing was tied to the confirmatory analysis phase while the exploratory analyses had no role in the decisions related to sampling.

Confirmatory analyses

As the aim of the present study was to assess the effectiveness of four stand-alone mindfulness interventions without any aspirations to draw broad inferences about mindfulness intervention in general, we planned to carry out four independent comparisons with the active controls. Therefore, we carried out four independent-samples Bayesian t-tests to determine whether there was a difference between each mindfulness exercise and the active control condition. We employed a two-tailed test using a non-informative JZS Cauchy prior for the alternative hypothesis with a default *r*-scale of $\sqrt{2}/2$ (Rouder, Speckman, Sun, Morey, & Iverson 2009). To account for the hierarchical nature of the data, we compared the condition means using a Bayesian mixed-effects model that involved a random intercept for lab and for different stories used in the non-mindful active control condition.

We implemented the sequential design as being adaptive, stopping collecting the data for a given test of a condition after obtaining a Bayes Factor of 10 in favor of H_1 or a Bayes factor of 1/10 in favor of H_0., and re-assigning the remaining participants into other conditions, where such a threshold has not been met. We chose a Bayes Factor of 10, because according to the classification of Lee and Wagenmakers (2013), it demarcates the threshold between moderate and strong evidence. We started monitoring the Bayes Factor when there were at least 100 participants in each group. We chose 100 arbitrarily, because with a smaller sample size, the probability of misleading evidence (i.e., a Bayes Factor trajectory that arrives

at the "wrong" boundary, that is, at the H_0 boundary when H_1 is correct and the other way around) would have been higher. Here, using a sample size of 100 and a Bayes Factor of 10 we aimed to substantially decrease the probability of misleading evidence (Stefan et al., 2017).

We monitored the Bayes Factor and recorded the interim analysis every 3 days to check whether the threshold was reached[7]. If the threshold was not reached (BF of 10 or of 1/10) we stopped the data collection after that the maximum number of participants[8] was reached. If for any of the four tests, one of these three conditions was not met (i.e., BF of 10 in favor of H_1; BF of 1/10 in favor of H_0; the maximum number of participants is reached), the data collection was terminated four months after the start of the experiment. When one of these three conditions has been met, we stopped allocating participants to that particular experimental condition and we allocated them to the remaining conditions for which this has not been the case yet. If the data collection ends before all the sites are able to collect at least 70 participants[9] (because we would have reached a Bayes Factor of 10 in favor of H_1 or of 1/10 in favor of H_0 for a given test of a condition), we will continue with the data collection by reopening all groups and equitably allocating each participant to each group until each site would have collected at least 70 participants. In this case, we would use this sample for exploratory analyses.

Exploratory analyses

We also carried out analyses exploring the effect of experimental conditions on pleasure, arousal, and dominance, and for the potential moderating effect of neuroticism. We

[7] Each time we did so, we generated a file in Rmarkdown that we uploaded on the OSF project page. Similarly, we generated a Rmarkdown file, twice a week, to document how many participants were collected for each site and how many participants were assigned to each experimental group.

[8] The maximum number of participants was set to 120*k*, where *k* is the number of sites that took part in the study, and 120 was the maximum number of participants that each site committed to collect.

[9] 70 is the minimum number of participants that we requested to each site to collect to be considered for authorship

thus performed three different Bayesian *t*-tests for each dimension of the Self-Assessment Manikin scale (pleasure, arousal and dominance) comparing our experimental conditions with the control condition. We then looked at The Bayes Factor to establish whether the data favored H_1 or H_0. Similarly as we did for the confirmatory analyses we compared the condition means using a Bayesian mixed-effects model that involved a random intercept for lab and for different stories used in the non-mindful active control condition to account for the hierarchical nature of the data.

To check whether neuroticism moderated the effects of the four experimental conditions on stress, we compared the model with the interaction to the model with only the main effects (using the lmBF function) and we reported the corresponding BF. If the model with the interaction was preferred to the model with only the main effects of a BF of 10 or more, we would have solid evidence of the moderation of neuroticism on stress. We computed the Markov chain Monte Carlo chain (MCMC) to estimate parameters through the posterior function, in this way for each condition we were able to assess whether there were differences in stress levels for people with high, average, or low neuroticism scores.

Dealing with careless or insufficient effort responder

We decided to set a series of rules to deal with careless or insufficient effort responders (C/IE; Currant, 2016), to reduce errors and provide more valid results. We decided to be conservative in our criteria by excluding participants that were providing too unrealistic answers that can hardly be due to casual errors, but at most to a consistent pattern in the response modality (i.e., finishing the survey as soon as possible to get payments/credits). Some possible issues were limited by some settings we fixed in the Qualtrics survey; we forced the answers for the questions connected to our exclusion criteria

(meditation experience, english level, mental illnesses) for the questionnaires related to our dependent variables/moderator instead the answers were requested[10].

To be sure that participants listened to the 15 minutes mindfulness/control condition we blocked the screen with the audio of the meditation/ control condition for 14 minutes (i.e., 840 seconds) so as not to allow participants to proceed to the following survey page until the meditation was actually finished. However, we excluded participants that completed the survey in less than 946 seconds, because as in addition to the 15 minutes of the recording we added a cut score for response time at 2 seconds an item, by following the recommendations of Huang et al., (2012)[11] To exclude participants that provided the same response to every question (i.e., selecting always the answer "strongly agree") we performed a long string analysis. Here we followed the recommendations of Huang et al., (2012) and we excluded from the study individuals with a string of consistent responses equal or greater than 10 (i.e., half of the length of the total scale). Furthermore If a participant responded to less than 50% of the items on a scale, we excluded that participant's score for that scale. Finally we decided not to exclude outliers as we believe that this would have led to the loss of important information and we did not perform any transformation of the data.

Ethics Information, Administrative Organization and Data protection

The study received ethical approval from Swansea University while the sites that participated in the data collection received ethical approval from their local IRBs. To minimize the workload to collaborators, we provided an IRB submission pack that sites could adapt to their institutions' requirements. For most sites, only small adjustments to our IRB pack were necessary for submission. Each site's IRB protocols with ethics details and acceptance of each protocol can be found on the OSF project page at https://osf.io/6w2zm/.

[10]If participants do not provide any answer a message appears asking participants whether they want to continue to the following part of the survey without answering that particular question.

[11] We had 53 questions resulting in 103 seconds to be added to the 840 seconds of the meditation/control condition thus arriving at 946 seconds

Swansea University was the administrative organization for the study. Swansea University was also the data controller for this project. Personal data of participants were processed for the purposes outlined in the information sheet (see the document Information Sheet at https://osf.io/xuznc/). Standard ethical procedures involved participants providing their consent to participate in this study by completing the consent form that was administered at the beginning of the online survey used for the experiment.

We left each site the possibility to recruit participants in their preferred way. Participants could be recruited through the SONA system of the respective institution or via crowdsourcing platforms such as mTurk or Prolific academic. Participants could come from any geographic area if they met our inclusion criteria and could be given either credits or financial compensation in exchange for participation in the study.

Conflict of interests

The certified mindfulness instructor that collaborated on the multi-site project only recorded the meditations, but was not involved in the writing of any parts of the manuscript (e.g., choice of the analyses).

References

Cavanagh, K., Strauss, C., Cicconi, F., Griffiths, N., Wyper, A., & Jones, F. (2013). A randomized controlled trial of a brief online mindfulness-based intervention. *Behaviour Research and Therapy, 51*(9), 573–578

Cavanagh, K., Strauss, C., Forder, L., & Jones, F. (2014). Can mindfulness and acceptance be learnt by self-help?: A systematic review and meta-analysis of mindfulness and acceptance-based self-help interventions. *Clinical Psychology Review*, *34*(2), 118–129

Cavanagh, K., Churchard, A., O'Hanlon, P., Mundy, E., Votolato, P., Jones, F., Gu, J. & Strauss, C. (2018). A randomised controlled trial of a brief online mindfulness-based intervention in a non-clinical population: Replication and extension. *Mindfulness*, *9*(4), 1191–1205

Council of Europe (2001). Common European framework of reference for languages: Learning, teaching, assessment

Feldman, G., Greeson, J., & Senville, J. (2010). Differential effects of mindful breathing, progressive muscle relaxation, and loving-kindness meditation on decentering and negative reactions to repetitive thoughts. *Behaviour Research and Therapy*, *48*(10), 1002–1011

Gebel, M. (2019, February 6). *Calm, the 7-year-old meditation app, says it's now valued at $1 billion*. Business Insider. Retrieved February 4, 2022, from https://www.businessinsider.com/relaxation-app-calm-first-meditation-startup-1-billion-valuation-2019-2?r=US&IR=T.

Germer, C. K., Siegel, R. D., & Fulton, P. R. (2016). *Mindfulness and psychotherapy.* Guilford Press.

Glück, T. M., & Maercker, A. (2011). A randomized controlled pilot study of a brief web-based mindfulness training. *BMC Psychiatry*, *11*(1).

Headspace (2021). *Researching meditation and mindfulness*. https://www.headspace.com/science/meditation-research

Hemingway, E. (1952). *The old man and the sea*. Life.

Hutcherson, C. A., Seppala, E. M., & Gross, J. J. (2008). Loving-kindness meditation increases social connectedness. *Emotion*, *8*(5), 720–724.

International Personality Item Pool. (2001). A scientific collaborator for the development of advanced measures of personality traits and other individual differences. Retrieved from http://ipip.ori.org/.

Le Carré's, J. (2021). *'They told me I was grown up enough to keep a secret': Exclusive extract from Silverview, John le Carré's final novel*. The Guardian. Retrieved February 4, 2022, from https://www.theguardian.com/books/ng-interactive/2021/oct/09/they-told-me-i-was-grown-up-enough-to-keep-a-secret-exclusive-extract-from-silverview-john-le-carres-final-novel.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course.* Cambridge University Press.

Morris, J. D. (1995). Observations: SAM: The self-assessment manikin: An efficient cross-cultural measurement of emotional response. *Journal of Advertising Research, 35*(6), 63–68.

Kabat-Zinn, J., Massion, A. O., Kristeller, J., Peterson, L. G., Fletcher, K. E., Pbert, L., Lenderking, W. R., & Santorelli, S. F. (1992). Effectiveness of a meditation-based stress reduction program in the treatment of anxiety disorders. *The American Journal of Psychiatry, 149*(7), 936–943.

Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: Past, present, and future. *Clinical Psychology: Science and Practice, 10*(2), 144–156.

Matko, K., & Sedlmeier, P. (2019). What Is Meditation? Proposing an empirically derived

classification system. *Frontiers In Psychology*, *10*, 2276.

Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review, 21*(2), 301–308.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237.

Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, *22*(2), 322–339.

Sparacio, A., Ropovik, I., Jiga-Boy, G. M., Forscher, P. S., Paris, B., & IJzerman, H. (2022). Stress regulation via self-administered mindfulness and biofeedback interventions in adults: A pre-registered meta-analysis. *PsyArXiv*, https://psyarxiv.com/zpw28/.

Spielberger, C. D., Gorshu, R. L. and Lushene, R. D. (1970). *Test manual for the state trait anxiety inventory.* Consulting Psychologist Press.

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 randomized controlled trials. *Clinical Psychology Review*, *45*, 102–114.

Stefan, A. M., Gronau, Q. F., Schönbrodt, F. D., & Wagenmakers, E. J. (2019). A tutorial on Bayes Factor Design Analysis using an informed prior. *Behavior Research Methods*, *51*(3), 1042–1058.

Taylor H., Strauss C., Cavanagh K., (2021). Can a little bit of mindfulness do you good? A systematic review and meta-analyses of unguided mindfulness-based self-help interventions. *Clinical Psychology Review*. *89,* 102078.

Tolkien, J., R., R. (1967). *Smith of Wootton Major*. Redbook.

Wahbeh, H., Svalina, M. N., & Oken, B. S. (2014). Group, one-on-one, or Internet?

Preferences for mindfulness meditation delivery format and their predictors. *Open Medicine Journal*, *1*, 66–74.