

# Pre-Analysis Plan: App-Based Mindfulness Meditation

Pierre-Luc Vautrey & Advik Shreekumar

August 6, 2021

## 1 Introduction

Many Americans, and particularly low-income individuals, suffer from mental illnesses such as depression or anxiety without receiving any treatment. Beside the direct consequences of mental illness on well-being, even moderate levels of anxiety may amplify other health problems as well as poverty by impairing decision making. Recent clinical and policy attention has focused on therapies that incorporate mindfulness meditation to produce low-cost and low-stigma interventions.

This study aims to investigate these types of interventions by giving adults access to and incentives to use Headspace, a leading mindfulness meditation app. The study has three purposes:

- (i) Measure the direct effects of mindfulness meditation on mental health, in a large sample of adults interested in the intervention.
- (ii) Measure the additional effects of mindfulness meditation on decision making behavior.
- (iii) Study the formation of individual habits.

We are randomizing participants into a Treatment group and a Waitlist Control, and then further randomizing the Treatment group between one of several incentive schemes. The Control group is further randomized between a cash transfer and a pure control.

## 2 Treatments and Experimental Protocols

### 2.1 Recruitment and Sampling

Participants will apply to enter the study by filling a short survey after clicking Facebook ads, which are distributed on the Facebook platform, Facebook Messenger, and Instagram. The short initial survey will collect consent, contact information and basic demographic information.

We plan to recruit 2,500 participants through the randomization stage.

### 2.2 Main Treatments

All participants in the Treatment group will receive a Headspace premium license valid for at least 3 months at the beginning of the study. In addition, we will randomize each Treatment participant into one of three incentive arms:

- (i) "Pure treatment", receiving only the Headspace license and no incentives to meditate.

- (ii) "Exploration", receiving incentives to meditate almost every day for 4 days.
- (iii) "Routine-forming", receiving incentives to meditate almost every day for 13 days.

All participants in the Waitlist Control group will receive a Headspace premium license after completing the final survey. In addition, we will randomize each Waitlist Control participant into one of two arms:

- (i) "Waitlist Control with Cash Transfer", receiving an unconditional gift card amount approximately equal in value of the Headspace licenses.
- (ii) "Pure Waitlist Control", receiving no unconditional transfers besides the final payment for participation (given to all participants) and incentives obtained in some questions.

### 2.3 Auxiliary Manipulations

During the decision-making endline, we will also randomize two auxiliary interventions:

- (i) Immediate meditation: randomly selected participants within the Treatment group will be incentivized to complete a meditation session right before continuing with the decision-making endline survey. The goal is to test whether there are any immediate effects of meditation on decision-making, on top on longer-term effects from sustained practice.
- (ii) Anxiety primers: within the decision-making endline, participants will be randomly assigned to anxiety or neutral primers. These questions will ask participants to recall and describe an anxious (or neutral) experience that they had, or will elicit answers to questions about how to deal with stressful, unexpected financial constraints. The goals are to (i) give a benchmark, allowing to assess the effects of meditation to the effects of temporary anxiety primers and (ii) test whether meditation makes people more resilient to stressful thoughts by looking at the interaction between treatment and anxiety primers. We will jointly randomize two primers per participant.

### 2.4 Experimental Protocol

Our experiment has the following timeline:

- (i) Recruitment via Facebook ads pointing to a survey that present study, collects consent, demographics and baseline familiarity and practice of meditation.
- (ii) Additional survey to test participants' phone number and collect baseline mental health measures
- (iii) On third survey, participants are randomized equally between five groups: (i) Pure treatment, (ii) Exploration, (iii) Routine-forming, (iv) Waitlist Control with Cash Transfer, (v) Pure Waitlist Control.
- (iv) Short surveys sent by SMS begin, measuring anxiety, beliefs about the effects of meditation, ease of meditating, and willingness to pay for an extension of their Headspace license. 3 surveys.

- (v) Short-run mental health endline. Two weeks after initial randomization, we will survey all participants about their mental health and subjective well-being.
- (vi) Decision making endline. Immediately after the mental health endline, participants will receive an invitation to a decision making endline. We will randomize a subset of the Treatment group to receive incentives to meditate immediately before taking the survey. The survey itself will include cognition and productivity tasks after a randomized anxiety prime, an information avoidance module, and a risk preference module. The anxiety prime will be designed to induce a low or high level of anxiety in participants.
- (vii) Long-run mental health endline. Four weeks after initial randomization, we will survey all participants about their mental health and subjective well-being. Waitlist Control participants will receive their Headspace licenses at the end of the survey.

## 3 Empirical Analysis

### 3.1 Key Outcomes

- Meditation. We will define meditation activity based on session-level usage data from the Headspace app. Headspace offers a variety of sessions, ranging from introductory courses (e.g., “Basics”), to topic-specific courses (e.g., “Pregnancy”, “Mindful Eating”), to ambient noise (e.g., “Temple Rain”, “Lo-Fi Times”). We will not count ambient noise toward’s a user’s meditation total, but will count all other sessions as meditation. Based on this activity data, we will proceed with two key definitions:
  - Minutes meditated per day.
  - Days meditated per week. We will define the start of each day as a period from 4:00am–3:59am.

We will assume that participants in the Waitlist Control are not using the Headspace app at all, imputing  $M = 0$  for them. We note that our measure of meditation will not cover any mindfulness that users practice outside the app (e.g., independently or through a separate service) or with secondary Headspace accounts that we don’t know about.

- Mental health. We will measure participants’ total scores on the Generalized Anxiety Disorder-7 Scale, the Perceived Stress Scale and the Patient Health Questionnaire (PHQ-8, without the question about suicidal thoughts) in the endline survey and in the four-weeks follow-up. We will also study the evolution of anxiety by measuring GAD-7 at baseline and the shortened GAD-2 in the short SMS surveys.
- Risk preferences: we will follow Callen, Isaqzadeh, Long and Sprenger (2014) to measure risk preferences, in particular: the switching point in the first task involving a lottery and a fixed amount, and the certainty premium constructed as in Callen et al. (2014).
- Productivity and Cognitive function:
  - Cognition (Stroop task). We will measure the overall time taken as well as the number of errors participants make.

- Productivity (Proofreading). This task will require participants to highlight simple spelling and grammar errors in a short paragraph. We will measure both the total number of errors participants find, and their productivity (errors found per unit time).

### 3.2 Secondary Outcomes

- Information avoidance (stressful info). We will form a standardized index from our four information avoidance tasks. We'll subtract the control mean and divide by the control standard deviation for each of the questions, and take a simple average of the standardized scores to form an index.
- Effects of cross-randomized anxiety primers on mood, proofreading score and risk preferences
- Propensity to focus on the most relevant information: index aggregating several outcomes:
  - Propensity to demand information about the coin flips in information choice game
  - Propensity to correctly choose to purchase or not purchase the stock when given all information pieces, depending on coin flip value
  - Information avoidance
  - "Risks as feelings": propensity to accept or reject a gamble that loses \$10 with probability 1% and wins \$1 with probability 99%
- Time spent answering incentivized questions and tasks
- Mindfulness scale (FFMQ-15). If FFMQ-15 score is affected, we will conduct mediation analysis on other affected outcomes.
- Beliefs about effects of meditation, perceived difficulty to meditate and willingness to pay for a license extension
  - Measured pre-intervention, 3 times during the first 2 weeks, and at the 4 weeks endline
  - We will study heterogeneous effects in habit formation and takeup along pre-intervention measures of these variables, and will study the evolution of these outcomes over time to understand if people are forming cue-based habits, if their perceived cost of meditation is evolving, or if they are learning about the benefits.

### 3.3 Regression Analysis

We will perform two sets of analyses. The first will be a minimal analysis following conventional practice, based on OLS and two-stage least squares regressions with strata fixed effects. The second will account for baseline covariates via a double machine learning approach.

**Conventional Analyses** We will investigate the effectiveness of incentives at inducing meditation via the following regression:

$$M = \beta_1 \mathbb{1}\{Treatment\} + \beta_2 \mathbb{1}\{ShortIncentives\} + \beta_3 \mathbb{1}\{LongIncentives\} + \alpha_{stratum} + \epsilon$$

The regression will include strata fixed effects, and standard errors will be unclustered. We may specify a different but equivalent set of dummies for the treatment arms to simplify hypothesis tests.

For analysis of mental health and decision making outcomes, we will report intent-to-treat (ITT) effects as well as local average treatment effects from two-stage least squares (2SLS), instrumenting for meditation activity with dummies for each incentive arm.

Our ITT specifications will be of the form:

$$Y_{post} = \beta_1 \mathbb{1}\{Treatment\} + \beta_2 \mathbb{1}\{ShortIncentives\} + \beta_3 \mathbb{1}\{LongIncentives\} + \alpha_{stratum}^{(1)} + \gamma^{(1)} Y_{pre} + \epsilon$$

where  $Y_{pre}$  is a pre-test score that is included where available.

Our 2SLS specification will have three instruments, corresponding to each of the incentive arms. We will conduct inference using conventional asymptotics as well as an Anderson-Rubin procedure that remains valid in the presence of heteroskedasticity and weak instruments. Both the first stage and second stage will include stratum fixed effects, as well as a pre-intervention score if available for an outcome. The reduced form specification will be:

$$M = \pi_1 \mathbb{1}\{Treatment\} + \pi_2 \mathbb{1}\{ShortIncentives\} + \pi_3 \mathbb{1}\{LongIncentives\} + \alpha_{stratum}^{(1)} + \gamma^{(1)} Y_{pre} + \nu$$

The second-stage specification will be:

$$Y_{post} = \beta_1 \hat{M} + \alpha_{stratum}^{(2)} + \gamma^{(2)} Y_{pre} + \epsilon$$

where  $\hat{M}$  is estimated in the first stage.

**Double Machine Learning** We collect a variety of pre-treatment information from pre-randomization surveys. In addition to the conventional specifications detailed above, we will report an additional set of analysis that adjust for these covariates using a double machine learning approach (Chernozhukov et al., 2018).

In particular, because most covariates we observe are categorical, we will report DML results where we perform partialling out using (1) regression tree methods (e.g., random forest or gradient boosted trees), and (2) an ensemble of a penalized linear model and a tree-based method. These exercises will incorporate all available covariates, including strata fixed effects and pre-treatment measures of outcomes where available. For each outcome, our preferred DML specification will be the model that maximizes cross-validated goodness of fit.

### 3.4 Key Contrasts

We will investigate the following questions:

- (i) Do short-term and long-term incentives aid habit formation? We will compare meditation outcomes in the short-term incentives group and in the long-term incentives group to meditation outcomes in the pure incentives group.
- (ii) Does mindfulness meditation reduce anxiety and other mental health issues? We will compare

GAD-7 scores, PSS scores and PHQ-8 scores among participants who meditate more and less, using treatment status as an instrument for meditation activity.

- (iii) Does meditation improve the amount of attention and the ability to allocate attention correctly? We will compare scores on the Stroop task and the Proofreading task among participants who meditate more and less. We will also compare the extent to which being assigned to an anxiety primer alters performance on productivity tasks and decision-making on the Callen et al. 2014 task between participants who meditate more and less.