Analysis Plan for Mindfulness Project

- **Research questions and hypotheses**
    - **Primary research question:** Can an eight-week mindfulness-based program improve the mental health of college students with mild to severe symptoms of depression and anxiety in China?
        - H1 (alternative hypothesis): Students in the MBP will show significantly greater improvement in depression and anxiety symptoms from pre-test to post-test compared to the control group.
        - H0 (null hypothesis): There will be no significant difference in the change in depression and anxiety symptoms from pre-test to post-test compared to the control group.
- **Primary variables of interest**
    - **Depression:** PHQ-9 (Patient Health Questionnaire-9). In line with scoring guidelines, we will sum responses to the nine questions, each rated 0 (not at all) to 3 (nearly every day), for a total score of 0-27. We will use this continuous outcome in the primary analyses, with higher scores indicating greater depression severity (e.g., 5-9 mild, 10-14 moderate, 15-19 moderately severe, 20-27 severe). These data were collected at baseline, three times during the intervention, and immediately post-intervention (as primary outcomes). We also plan to measure this outcome at the 3-month follow-up (key secondary), the 6-month follow-up (secondary), and post-graduation (secondary).
    - **Anxiety:** GAD-7 (Generalized Anxiety Disorder-7). In line with scoring guidelines, we will sum responses to the seven questions, each rated 0 (not at all)

to 3 (nearly every day), for a total score of 0-21. We will use this continuous outcome in the primary analyses, with higher scores indicating greater anxiety severity (e.g., 5-9 mild, 10-14 moderate, 15-21 severe). These data were collected at baseline, immediately post-intervention, and biweekly during the intervention. We also plan to measure this outcome at 3-month follow-up, 6-month follow-up, and post-graduation.

Practice time: In weekly surveys, we asked students in the treatment group which days they had practiced and on average how much they practiced on those days. For each participant, we then multiply the number of days practiced by the average practice time to obtain their weekly practice time outside of class for that week. In addition, we also ask whether students in the treatment group had attended the 90-minute class during the previous week, and we also include this as a part of their total weekly practice time. We focus on the treatment arm for within-person dose–response. We will also check whether students in the control group report any mindfulness-like practice in weekly surveys.

- **Statistical analysis plan:**

Means, standard deviations, and ranges will be calculated for all continuous variables. Frequencies and percentages will be calculated for all categorical variables. Balance tests will be performed between the treatment and control group, as well as between attrited and non-attrited students. Since the randomization is at the individual level, we use robust standard errors.

To obtain estimates of the effect of participation in the mindfulness program on the entire sample, we calculate the intention to treat (ITT) effect of the program on those assigned to the treatment, regardless of their take-up or engagement in the program. To do this, we regress the

outcome on treatment assignment for the whole sample using analysis of covariance (ANCOVA). The model is specified as follows:

$$Y_{i,\text{post}} = \beta_0 + \beta_1 T_i + \beta_2 Y_{i,\text{pre}} + \mathrm{X}_i'\gamma + \epsilon_i$$

where $Y_{i,\text{post}}$ is the outcome variable for individual *i* measured at post intervention (immediate post intervention, 3-month, 6-month, and post-graduation graduation follow-ups); T_i is the treatment assignment indicator (1=treatment, 0=control); $Y_{i,\text{pre}}$ is the baseline measure of the continuous mental health outcome (depression or anxiety) for individual *i*; X_i' is a vector of other baseline covariates for individual *i* (including age, gender, academic year, academic major, urban/rural childhood residence, and only child status); γ is a vector of coefficients corresponding to the covariates in X_i'; and β_1 is the main coefficient of interest, estimating the average difference in the post-treatment outcome between the treatment and control groups, while holding the baseline outcome and other covariates constant.

In this paper, we also seek to understand the effect of the intervention on those with higher levels of compliance to the intervention protocol (i.e., the effect of engaging in the mindfulness program through individual outside of practice as opposed to simple random assignment to it). To do this, we estimate the Local Average Treatment Effect (LATE). To address potential endogeneity, we employ a two-stage least squares (2SLS) instrumental variable (IV) strategy, using the initial random assignment to the treatment group (T_i) as a valid instrument for the treatment receipt (D_i). Here, we define treatment receipt as attending at least 50% of the eight weekly sessions (four out of eight) and engaging in formal mindfulness practice

for at least 25% of the total suggested time (180 minutes per week * 0.25 = 45 minutes per week). The model is specified as follows:

$$D_i = \alpha_0 + \alpha_1 T_i + \alpha_2 Y_{i,\text{pre}} + \mathrm{X}_i'\delta + u_i$$

where D_i is an indicator variable equal to 1 if individual *i* received the treatment and 0 otherwise; T_i is the instrument, equal to 1 if individual *i* was assigned to the treatment group and 0 otherwise; $Y_{i,\text{pre}}$ is the baseline measure of the outcome; X_i' is a vector of other baseline covariates; and α_1 is the key first-stage coefficient, representing the effect of assignment on receipt.

The second stage equation, which regresses the final outcome on the predicted values of treatment receipt, is as follows:

$$Y_{i,\text{post}} = \beta_0 + \beta_1^{\text{LATE}} \widehat{D}_i + \beta_2 Y_{i,\text{pre}} + \mathrm{X}_i'\gamma + \epsilon_i$$

where $Y_{i,\text{post}}$ is the outcome variable for individual *i* measured after the intervention, $\widehat{D}_i$ is the *predicted* value of treatment receipt for individual *i*, and β_1^{LATE} is the Local Average Treatment Effect (LATE) (the primary coefficient of interest).

- **Attrition and missing data**

We define attrition as non-participation at the time of the post-intervention survey. For survey attrition, we do not impute missing data. However, we will check whether there is a correlation with treatment status. If so, we will construct robust bounds for the treatment estimates (e.g., Lee bounds).

For item missing data in PHQ-9/GAD-7, we remove the score if more than 20% of the data on that scale are missing. As with attrition, we also check whether there is a correlation

between the missing items and treatment status. If 20% or less of the data are missing, we impute the mean score of the other responses for that individual.

- **Multiple outcomes/hypotheses**

To address the testing of multiple outcomes (anxiety and depression), we apply the Bonferroni correction. To do this, we divide the original alpha by the number of comparisons (i.e., 0.05 / 2 tests = 0.025).

- **Heterogeneity**

We conduct moderation analyses of the treatment effect on depression and anxiety symptoms. We will test for heterogeneous treatment effects by estimating an Ordinary Least Squares (OLS) regression model that includes an interaction term between the treatment indicator and the subgroup variable. The model that we use for this is the following:

$$Y = \beta_0 + \beta_1 * \text{Treatment} + \beta_2 * \text{Subgroup} + \beta_3 * (\text{Treatment} * \text{Subgroup}) + \text{Covariates} + \varepsilon$$

The potential moderators that we investigate in this analysis include the following:

- Baseline anxiety symptom severity: We include this as a continuous mean score on the GAD-7.
- Baseline depression symptom severity: We include this as a continuous mean score on the PHQ-9.
- Level of childhood adversity: We include this as a continuous mean score on the Childhood Trauma Questionnaire – Short Form (CTQ-SF-25).
- Gender: We include this as a binary variable (male/female)