The Selah Trial: A Preference-based Partially Randomized Waitlist Control Study of Three Stress Management Interventions

Rae Jean Proeschold-Bell, Ph.D.

Duke Global Health Institute, Duke University Center for Health Policy and Inequalities Research, Durham, NC

David E. Eagle, Ph.D.

Duke Department of Sociology, Duke Global Health Institute, Duke University Center for Health Policy and Inequalities Research, Durham, NC

Logan C. Tice, MA

Duke Global Health Institute, Duke University Center for Health Policy and Inequalities Research, Durham, NC

Alyssa Platt, MS

Duke Global Health Institute, Department of Biostatistics and Bioinformatics, Duke University, Durham, NC

Jia Yao, MA

Duke Global Health Institute, Duke University Center for Health Policy and Inequalities Research, Durham, NC

Jessie S. Larkins, MEd, MDiv LCMHCA, NCC

Duke Global Health Institute, Duke University Center for Health Policy and Inequalities Research, Durham, NC

Eunsoo Timothy Kim, Ph.D.

Duke Global Health Institute, Duke University Center for Health Policy and Inequalities Research, Durham, NC

Joshua A. Rash, Ph.D.

Department of Psychology, Memorial University of Newfoundland, St. John's, NL, Canada

Please direct comments to

Rae Jean Proeschold-Bell

at rae.jean@duke.edu.

Abstract

Chronic stress undermines psychological and physiological health. We tested three remotely delivered stress management interventions among clergy, accounting for intervention preferences. United Methodist clergy in North Carolina enrolled in a partially randomized, preference-based waitlist control trial. The interventions were: mindfulness-based stress reduction (MBSR), Daily Examen prayer practice, and Stress Proofing (stress inoculation plus breathing skills). Co-primary outcomes were symptoms of stress (Calgary Symptoms of Stress Inventory) and 48-hour ambulatory heart rate variability (HRV) at 12 weeks compared to waitlist control. Survey data were collected at 0, 12, and 24 weeks and 48-hour ambulatory HRV at 0 and 12 weeks. The 255 participants were 91% White and 48% female. Forty-nine participants (22%) without a preference were randomly assigned between the three interventions (n = 40) and waitlist control (n = 9). Two hundred six participants (78%) with a preference were randomly assigned to waitlist control (n = 62) or their preferred intervention (n = 144). Compared to waitlist control, MBSR [mean difference (MD) = −0.30, 95% CI: −0.41, −0.20; P < .001] and Stress Proofing (MD = −0.27, 95% CI: −0.40, −0.14; P < .001) participants had lower stress symptoms at 12 weeks; Daily Examen participants did not until 24 weeks (MD = −0.24, 95% CI: −0.41, −0.08). MBSR participants demonstrated improvement in HRV at 12 weeks (MD = +3.32 ms; 95% CI: 0.21, 6.44; P = .036). MBSR demonstrated robust improvement in self-reported and objective physical correlates of stress; Stress Proofing and Daily Examen resulted in improvements in self-reported correlates of stress. These brief practices were sustainable and beneficial for United Methodist clergy during the heightened stressors of the COVID pandemic.

Keyword: mindfulness, prayer, heart rate variability, vagal tone, stress management, occupational stress

This is an author-produced PDF of an article accepted for publication in the *Translational Behavioral Medicine*. The full article citation is: Proeschold-Bell, R.J., Eagle, D., Tice, L., Platt, A., Yao, J., Larkins, J., Kim, E., Rash. (2024). The Selah trial: A preference-based partially randomized waitlist control study of three stress management interventions, *Translational Behavioral Medicine* https://doi.org/10.1093/tbm/ibae017

Lay summary

A common source of stress, which can harm physical and mental health, is work. Clergy engage in a profession that requires toggling between varied and interpersonally complex tasks, providing emotional labor, and experiencing stressors such as public criticism. Practical, brief practices are needed to manage occupational stress. We invited all United Methodist clergy in North Carolina to enroll in a stress management study. Participants chose their preferred of three interventions: mindfulness-based stress reduction (MBSR), Daily Examen prayer practice, or Stress Proofing (a combination of stress inoculation plus breathing skills). Clergy without a preference were randomly assigned to one of the three interventions and a waiting group. Clergy with a preference were randomly assigned to either begin the intervention or wait at least 6 months and provide data while waiting. Participants practiced each of the three interventions at high levels across 24 weeks. Compared to clergy who waited for an intervention, MBSR participants evidenced robust improvement in self-reported (stress and anxiety symptoms) and physiological (heart rate variability measured across 48 hours) outcomes, whereas Stress Proofing and the Daily Examen only resulted in improvements in self-reported outcomes. The three brief practices were sustainable and beneficial for United Methodist clergy during the heightened stressors of the COVID pandemic.

Practice: Mindfulness-based stress reduction can be used to improve physiological and psychological correlates of stress, whereas the Stress Proofing program of stress inoculation therapy plus breathing exercises and the Daily Examen prayer practice can be used to improve psychological correlates of stress, among United Methodist Church clergy.

Policy: United Methodist denominational officials who want to improve stress management and associated health outcomes among clergy should encourage use of synchronous, web-delivered mindfulness-based stress reduction (MBSR) programs and can consider offering the Daily Examen prayer practice and Stress Proofing for clergy who do not respond favorably to MBSR.

Research: Future research should: (i) examine whether these three stress management programs operate through their purported mechanisms of action; (ii) test ambulatory heart rate variability (HRV) at longer-term follow-up for Stress Proofing and the Daily Examen prayer practice and determine whether improvements in HRV observed for MBSR persist at longer-term follow-up; and (iii) test other structured prayer practices for stress management for clergy and possibly other religious populations.

Introduction

Stress is a complex phenomenon occurring when the demands of a situation exceed the resources that one has to cope effectively. Stress involves a physiological component (i.e. bodily changes) and a psychological component (e.g. perception of circumstances in life). When not

managed well, stress can contribute to physiological problems (e.g. coronary artery disease, stroke) and psychological concerns (e.g. major depression).

Clergy are often frontline professionals responsible for care during times of family and community crisis, and are often the first person sought out by individuals struggling with serious mental illness. Identifying evidence-based ways for clergy to manage and reduce their stress symptoms is important for clergy themselves and those they help. Physical and mental health concerns among Mainline Protestant clergy, including average-to-elevated prevalence of chronic diseases and symptoms of depression and anxiety, may relate to exposure to chronic stressors. The clergy occupation is emotionally demanding with leadership responsibilities, public criticism, and few breaks. The translation of efficacious stress management practices for professionals engaged in emotional work is needed; previous translational research indicates the value of disseminating psychosocial interventions in the workplace.

Among the interventions developed for stress management, mindfulness-based stress reduction (MBSR) is well-established with demonstrated efficacy at improving symptoms of stress and anxiety among diverse populations. Unfortunately, poor engagement in mindfulness-based interventions can be a barrier to obtaining beneficial effects; dropout rates can exceed 25% and weak associations have been observed between self-reported home practice and intervention outcomes among 28 studies ($r = 0.26$, 95% CI: 19–0.34). Another intervention, stress inoculation treatment, has demonstrated efficacy for anxiety and depression, although research on adherence to treatment over time is limited. Aspects of stress inoculation therapy, such as engaging in stressful activities to gain confidence in one's stress-endurance abilities, draw on cognitive behavioral therapy, which reliably results in improvement of stress symptoms. We sought to translate and test efficacious programs of MBSR and stress inoculation treatment with clergy, attending to the busy lifestyles and tendency of clergy to overextend themselves due to the emotional and sacred nature of their work. We further sought to test a prayer practice without an evidence base for stress management, but which includes pathways (e.g. attention focused on the moment, nonjudgmental thinking) that lead to stress symptom improvement for MBSR. Prayer practices are personalized to the occupational population of clergy, which we hoped would promote regular practice. All interventions tested were delivered remotely due to the COVID-19 pandemic.

To assess stress management outcomes, stress can be thought of as a latent construct that varies temporally and requires assessment of multilevel indicators to adequately capture. Stress can be measured as an exposure or a response, and can be approximated through symptoms that are behavioral, cognitive (e.g. appraisals), affective (e.g. mood), or physiological (e.g. operation of the autonomic nervous system). In this study, we sought to capture stress response through the use of measures that capture physical and psychological symptoms associated with chronic stress, and a biological marker associated with stress [i.e. heart rate variability (HRV)].

HRV reflects variations in heart rate that index the capacity of the parasympathetic nervous system to alter heart rate in order to effectively meet the demands of a stressful event. We chose HRV as a biological indicator of stress for several reasons, including (i) neurobiological evidence suggesting that HRV is impacted by stress, and supporting its use for the objective assessment of stress; (ii) pragmatic considerations, including ease of collecting ambulatory heart rate data; and (iii) observations from our pilot trial suggesting that HRV would be responsive to

the interventions offered and supporting the feasibility of measuring HRV as an objective marker of stress within our study population. Furthermore, HRV reliably covaries with stress during stress-inducing procedures, and lower levels of HRV are reliably associated with depressed mood, anxiety disorders, and reports of heightened occupational distress. Additionally, HRV is a strong indicator of morbidity and risk of mortality in longitudinal studies.

In a meta-analysis specific to mindfulness-based interventions, only three studies evaluated long-term (i.e. 24-hour) recordings of HRV, which may represent a better indicator of chronic stress. Results of these studies were equivocal and difficult to interpret due to limitations that included relatively small sample sizes (n = 19–42), or recruitment of a health population characterized by dysregulation of the autonomic nervous system (i.e. fibromyalgia).

The current study aimed to test the effectiveness of three stress management interventions (i.e. MBSR; the Daily Examen prayer practice; and a set of stress inoculation skills entitled Stress Proofing) shown to be acceptable and feasible in a pilot study among United Methodist clergy. Outcomes of interest included symptoms of stress and HRV (co-primary outcomes), anxiety symptoms (secondary outcome), and depressive symptoms (exploratory outcome). We avoided a one-size-fts-all approach and tested three interventions to allow for treatment personalization; an additional aim was to determine the effect of participant intervention preference on outcomes. We hypothesized that participants in each of the three active intervention conditions would experience improvements in each outcome compared to the waitlist control condition, and that participants with specific intervention preferences would experience larger improvements.

Methods

Study Design

We conducted a partially randomized preference trial with a waitlist control. Preference-based trials, a type of pragmatic clinical trial design, recognize that individuals have treatment preferences that are likely to affect outcomes due to expectancy effects or degree of engagement, which are particularly important in behavioral interventions. Frequently in the partially randomized design, study participants without specific preferences are randomized to intervention, whereas those with a specific preference are allowed to choose an intervention. In our design (also used by Carlson et al.), participants without specific preferences were randomized to the waitlist control group or one of three interventions, and participants with a preference were randomly assigned to begin their preferred intervention or to the waitlist control group. Henceforth, this is referred to as our “trial-phase” cohort. Following enrollment in the partially randomized waitlist-controlled preference trial, after the start of the COVID-19 pandemic, enrollment was re-opened to a fully “observational” cohort whose participants chose an intervention without a randomization structure for sensitivity analyses pertinent to generalizability (see Supplementary Fig. A1). The study protocol was registered under ClinicalTrials.gov identifier NCT04625777 and published online. All procedures were approved by

the Duke University Campus Institutional Review Board and all participants gave informed consent.

Participants

The United Methodist Church (UMC) in the USA includes 5.7 million members and constitutes 9% of all religious congregations. The UMC is categorized as a moderate-to-liberal Protestant Christian denomination, known for its focus on social and economic justice, respect for diverse religious beliefs, and openness to modernity. The two organizing bodies of the UMC in North Carolina (NC) in 2021 had a combined membership of nearly 1600 clergy, 2300 churches, and 500 000 congregants.

Our target population was ~1600 active UMC clergy in North Carolina, USA, in 2019–21; this sample of UMC clergy has been shown to be demographically similar to all Mainline clergy in the USA [36]. UMC clergies are ordained (or on the path to ordination) within a specific organizing body called an Annual Conference. Study inclusion criteria were UMC clergy ages 18 or older with a current appointment in either the North Carolina or Western North Carolina Annual Conference. To enhance ecological validity, there were no stressor health-related study exclusion criteria. Once in the study, participants with certain health conditions were excluded from HRV data collection (detailed below) but included in survey data collection.

Study Procedures

Recruitment and enrollment: Trial cohort

We invited eligible clergy from November 2019 to January 2020 via mail and email addresses provided by the UMC conferences and announcements at gatherings. We directed interested participants to a website to enroll prior to February 2020 by (i) completing the Treatment Acceptability and Preferences Scale for each intervention; (ii) expressing their preferences, if any, among the three active interventions; and (iii) providing study consent. Randomized assignments for participants with and without preferences were performed in February 2020, as specified in Supplementary Table A1. Participants assigned to immediate intervention chose among a list of dates for their desired intervention workshop. The Selah trial-phase cohort is participants who enrolled prior to 1 March 2020 and provided data while participating in an immediate intervention or the waitlist condition.

Recruitment and enrollment

Observational cohort We re-opened enrollment in March 2020, anticipating that interest in stress management may increase with the start of the COVID-19 pandemic. Participants who enrolled after 28 February 2020 selected any intervention workshop and dates from a list; they were not randomized and represent a fully observational cohort that received their chosen

intervention. The Selah observational cohort is participants enrolled after 1 March 2020, plus post-waitlist data from participants who enrolled before 1 March 2020.

Randomization

For the trial-phase cohort, we asked participants whether they preferred any of the three interventions during the enrollment process. We randomized participants who stated no preference to be able to immediately receive one of the three active interventions or to the waitlist using an allocation ratio of 1:1:1:1. We randomized participants who preferred one intervention to immediately receive their preferred active intervention or to a waitlist with a 3:1 intervention vs waitlist ratio for MBSR and Stress Proofing, and a 5:4 intervention vs waitlist ratio for the Daily Examen (for all preference scenarios, see Supplementary Table A1). The analysis statistician wrote code to generate the random allocation sequence in Stata version 16. Two staff members were responsible for assessing randomization results and informing participants of intervention allocation.

Masking

One staff member who was not the analysis statistician executed the randomization codes so that the analysis statistician could remain masked to intervention allocation until data were collected and analysis decisions finalized. Staff cleaning HRV data were masked to intervention assignment. All others, including participants and intervention workshop instructors, were aware of group assignments. For details on masking, see Supplementary Appendix C.

Intervention and data collection implementation

Implementation of the interventions and data collection for the trial-phase and observational cohorts occurred during April 2020 to October 2021. The three active interventions were designed as multi-session workshops and are described in detail in the protocol manuscript. As a pandemic related modification to the trial, workshops were conducted synchronously online in groups of 5–25 participants rather than in-person. Adaptations to study design due to COVID19 have been described in detail previously. Participants provided data from their homes or personal settings.

Interventions

MBSR

MBSR teaches a set of meditation activities, along with attention to attitudes. We contracted with certified instructors from Duke Integrative Medicine who used a criterion-standard mindfulness-based intervention based on Jon Kabat-Zinn's model. An instructor taught eight weekly 90-minute, synchronous, web-based videoconference sessions on awareness of breath, body scans, walking meditation, "choiceless" open awareness, Loving Kindness Meditation, yoga, and bringing awareness to the present moment. The course provided meditation instruction, periods of guided practice, and group discussion. Following their first session, participants were asked to practice MBSR for 45 minutes per day for 6 months. After the eight sessions, participants were offered a 4-hour online "Day of Mindfulness" which included participants and community members not enrolled in the study.

The Daily Examen prayer practice

The Daily Examen is a Jesuit reflective prayer practice developed by Ignatius of Loyola and widely practiced by Christians from many traditions. Two certified spiritual directors who are experts in the Daily Examen developed and co-taught each training. They used a five-step, modern adaptation of the Daily Examen [42]: (i) Become aware of God's presence; (ii) review the events of the past 24 hours, recalling things for which you are grateful; (iii) review the events of the past 24 hours, noticing where you experienced God's presence; review what stands out and pay attention to what emotions arise; (iv) Consider what went well or wrong, and if needed, ask God for forgiveness; and (v) look ahead to the next day; consider one thing you should do and where you need God's assistance.

With some similarities to mindfulness-based practices, these steps help participants attend to the present by reflecting on positive emotions, moving past negative emotions, and aligning their actions with their perception of God's wishes, with decreased judgment of their thoughts and feelings. The trainings consisted of three 90-minute, synchronous, web-based sessions and involved didactic content, practice, and small group discussion. Participants were asked to commit to practicing the Daily Examen daily for 10–15 minutes over 6 months following their first workshop session. Two and 6 weeks following their workshop, participants had the option to meet with their instructors in an online small group to discuss their practice.

Stress Proofing inoculation combination

Stress Proofing is a set of stress reduction skills with aspects of stress inoculation training selected and packaged by the NC Systema organization founder into four weekly 90-minute, synchronous, web-based sessions. The founder was the instructor for all study training sessions. The four seasons series began with education on the stress response and awareness of one's own stress response. The training diverged from traditional stress inoculation training and focused on physical activities to undo the stress response, including walking while diaphragmatic breathing, triangle and square breathing, tension control, stretching, and massage. The instructor discussed stress inoculation training and encouraged participants to embrace physical discomfort to learn

to tolerate discomfort. The session content recommended a variety of beneficial lifestyle practices, including prioritizing nutrition and sleep and disengaging from technological devices an hour before sleep. We asked participants to practice diaphragmatic breathing and any other Stress Proofing activities daily for 6 months. Workshops were taught by the founder.

Waitlist condition

Waitlist participants waited at least 6 months to participate in interventions. During this time, they completed surveys at 0, 12, and 24 weeks. We invited participants without disqualifying health conditions (as noted below in Measures: HRV) to additionally provide a 48-hour continuous ambulatory heart rate recording coinciding with their 0- and 12-week surveys. After completing the waiting period, participants could update their intervention preference and receive an intervention while providing survey and HRV data.

Measures

Co-primary outcomes: Stress symptoms and HRV

Stress symptoms were measured using the subscales of anger, muscle tension, cardiopulmonary arousal, neurological/ gastroenterological, and cognitive disorganization (total 41 items) of the Calgary Symptoms of Stress Inventory (C-SOSI), a reliable and valid measure guided by mindfulness-based theory in its development. Participants were asked to indicate on a scale of 0 (never) to 4 (frequently) of how often they experienced each symptom when presented with a stressor. We used continuous mean scores of all the items (range 0–4), with higher mean scores indicating worse symptoms. The study Cronbach's alpha was 0.96.

Ambulatory HRV was measured across a 48-hour period and indexed using the time-domain metric root mean square of successive RR differences (RMSSD) because it is less affected by breathing and a more suitable outcome measure in ambulatory studies than frequency-domain measures. Participants were mailed a Bittium eMotion Faros 180 heart rate recording device with electrodes two weeks prior to the intervention and taught to connect the device to two pregelled (Ag/AgCl) disposable Ambu BlueSensor wet-gel ECG electrodes placed beneath the right clavicle and left ribcage. Participants were instructed to wear this ambulatory heart rate monitoring device for a 48-hour period at week 0 and week 12, during which time they proceeded with their usual work, exercise, bathing, and sleep routines. Heart rate was measured using continuous electrocardiographic (ECG) recording sampled at a rate of 1000 Hz and used to calculate HRV. Study staff imported the 48-hour ECG recording to Kubios HRV Premium V3.4.1 software, partitioned it into 5-minute segments, visually inspected it to allow for manual correction of ectopic beats, detrended it, and subjected it to Kubios' automatic artifact correction algorithm. Two individual-level cosine function parameters were estimated across 5-minute segments by Ordinary Least Squares regression to quantify the circadian variability parameters:

(i) Midline Estimating Statistic of Rhythm (MESOR), defined as the rhythm-adjusted 24-hour mean, and (ii) amplitude, defined as the distance between MESOR and the maximum of the cosine curve (i.e. half the extent of rhythmic change in a cycle). Participants were excluded from HRV data collection if they had underlying medical conditions, including a diagnosis of tachycardia; being pregnant or becoming pregnant during the course of data collection; being diagnosed with COVID-19; having a pacemaker; or documentation of other cardiovascular-related chronic or acute morbidities that could impact the integrity of HRV data (Supplementary Table A2). See Supplementary Methods for detailed HRV procedures.

Secondary and exploratory outcomes

The secondary outcome, symptoms of anxiety, was measured using the seven-item Generalized Anxiety Disorder-7 (GAD-7) scale (sum scores ranged from 0 to 21, with scores of ≥8 screening positive for elevated anxiety symptoms) [46]. The study Cronbach's alpha was 0.90. The exploratory outcome of depressive symptoms was measured using the eight-item Patient Health Questionnaire-8 (PHQ-8; sum scores ranged from 0 to 24, with scores of ≥10 screening positive for elevated depressive symptoms) [47]. The study Cronbach's alpha was 0.87.

Demographic, intervention preference, and clinical measures

Sociodemographic and clinical measures were obtained by self-administered surveys during the baseline, 12-, and 24-week time periods. The clinically relevant constructs included physical activity, body mass index (BMI), and caffeine and alcohol intake. Preference measures were collected by self-administered surveys; the Treatment Acceptability and Preferences Scale [38] was administered during the enrollment period and preference for online vs in-person intervention was administered at baseline (see Supplementary Methods for details).

Engagement measure: Daily practice reports via text

We sent participants a daily text message for 24 weeks during the active intervention period. MBSR participants reported the number of minutes practiced the prior day. Daily Examen participants reported whether or not they had practiced the prior day. Stress Proofing participants reported whether they had conducted 0, 1, or 2 "resets" (i.e. Stress Proofing practices) the prior day.

Data collection procedure

Intervention participants

For both trial-phase and observational cohort participants, we collected survey data, solicited by email and administered in REDCap database software 12.4.28, at intervention start (0 weeks), 12 weeks, and 24 weeks. We collected HRV data at 0 and 12 weeks. We mailed participants Bittium eMotion Faros 180 ambulatory heart rate recording devices. After collecting

their 48-hour sample, participants returned devices by mail. Study staff cleaned and processed the data using Kubios Premium software 3.4.1. Baseline data collection occurred after randomization among trial participants, thus it was possible for dropout to occur prior to baseline data collection.

Waitlist participants

We initiated data collection from waitlist participants in groups of 20 starting in June 2020, July 2020, September 2020, and February 2021 to span the range of data collection from immediate intervention participants. The spacing of data collection during the waiting period matched that of intervention participants: surveys at 0, 12, and 24 weeks, and HRV data at 0 and 12 weeks. We asked waitlist participants who proceeded to start an intervention following the waiting period to provide data again during their intervention period following this same schedule. We included the post-waitlist intervention data in observational arm (i.e. sensitivity) analyses.

Incentives

We compensated participants $25 for each occasion of 48-hour ambulatory HRV data submitted, $20 each for 0- and 12-week surveys, and $25 for 24-week surveys.

Data availability

The datasets generated during the current study are not publicly available but we will make de-identified data available for reasonable requests compliant with ethical approvals from the sending and receiving hosts' institutional ethics review boards.

Statistical analysis

The Selah study team administers a biennial panel survey of all UMC clergy in North Carolina, the Clergy Health Initiative Longitudinal Survey. The 2019 wave of the panel survey obtained a 73% response rate. We compared descriptive statistics between the Selah trial phase and the 2019 panel survey to determine characteristics associated with self-selection of clergy into the Selah study and to inform representativeness of the study participants and generalizability of results.

Based on results from our pilot study, we expected an average baseline C-SOSI score of 0.92 (SD = 0.46) across all interventions, with 12-week follow-up scores of 0.7 (SD = 0.58) for MBSR, 0.55 (SD = 0.36) for Stress Proofing, and 0.51 (SD = 0.38) for the Daily Examen. A per-arm

sample size of 40 for Daily Examen, 47 for Stress Proofing, and 195 for MBSR would have 80% power to detect similar differences, using a two-sample t-test with unequal variance given 20% loss to follow-up and a design effect of 1.3 (with an average cluster size of 12) due to clustering caused by group delivery of the intervention. For HRV, a per-arm sample size of 140 would have 80% power to detect a medium effect size (standardized mean difference of 0.5) for a two-sample t-test, assuming a similar follow-up rate and design effect as with C-SOSI. All sample size calculations assumed an alpha of 0.0167 based on a Bonferroni correction on three hypotheses (for three interventions) and were conducted using PASS 2021 software. We aimed for a combined sample of 400 intervention and control participants (see Supplementary Appendix for details).

Use of the partially randomized preference design with a waitlist control during the trial phase meant that the analytic data would be a mix of data from participants who were: (i) randomized to one of three active interventions to start immediately, or to a waitlist control group (for trial participants who had no preference) and (ii) assigned to one's preferred intervention and then randomized to start immediately or to the waitlist control group (for trial participants who had a preference and were allowed to select their intervention). This approach increased the likelihood that intervention arms would be imbalanced on baseline characteristics in ways that may be associated with study outcomes in an unadjusted analysis, even with the inclusion of a shared waitlist control. In addition, randomization was performed prior to baseline data collection, with substantial dropout before data collection commenced. A propensity score covariate adjustment method was selected using covariate balancing propensity scores; details are provided in Supplementary Methods. All analyses use an as-treated approach to calculate treatment effects.

Main outcome models used linear mixed-effects models with random intercepts at the level of the individual to account for repeated measurements within individuals. Random slopes on a binary treatment indicator for the group assignment were used to produce a random intercept for each workshop and a separate intercept for the un-clustered control arm in order to account for partial clustering due to group-administered treatment. We calculated an interclass correlation coefficient (ICC) for each outcome model to quantify the level of clustering due to group treatment. We included a cubic functional form for calendar time to protect against time confounding (details included in Supplementary Methods).

Because propensity score adjustment was used to balance baseline levels of the outcomes (in addition to other prognostic characteristics), we chose to model treatment effects longitudinally using a constrained longitudinal data analysis modeling (cLDA) approach, which models baseline as an outcome and assumes baseline levels of the outcome are equal across arms [53]. Due to known variation in timing of the 12- and 24-week surveys, time was modeled continuously in weeks from baseline using linear splines with knots at 12 and 24 weeks for C-SOSI, GAD-7, and PHQ-8 outcomes and one knot at 12 weeks for HRV outcomes. Treatment effects were between-arm differences in outcomes at 12 weeks (primary) and 24 weeks

(secondary). We used robust sandwich standard errors to account for the fact that propensity scores were estimated.

We extracted and compiled the text data to calculate the proportion of participants engaging in their assigned practice on each day throughout the 24 weeks. Analysis of practice data was purely descriptive with no hypothesis testing.

Subgroup analysis was performed to ascertain whether treatment effects were different for participants who received an intervention that they uniquely preferred versus those that had tied or no preference or who received an intervention other than the one for which they expressed a unique preference. Treatment effect estimates by preference status were calculated using binary interaction terms with treatment and time terms with calculated linear combinations for treatment estimate by preference status. These subgroup analyses are strictly exploratory, thus analytic focus should be on effect estimates, confidence intervals, and magnitude of interaction terms rather than statistical significance and any patterns observed would need to be confirmed in a future randomized study for definitive conclusions to be made.

Data collected from trial-phase participants who provided post-waitlist data while receiving an intervention or from participants who enrolled in the study after 1 March 2020 and therefore were not randomly assigned to immediate intervention vs control were considered fully observational and separate from the trial-phase data collection and thus were excluded from the main analysis. A sensitivity analysis was performed pooling these data with the trial-phase data to ascertain whether results remained when all available data were used.

Missing data were present both in baseline covariates used to generate propensity scores and 12- and 24-week outcome data due to study dropout. Therefore, sensitivity analyses were performed using multiple imputation with chained equations (MICE) (see Supplementary Materials) to assess the extent to which missing data and study dropout may have affected the magnitude and direction of treatment effect estimates. Propensity scores were calculated separately for each of the imputation datasets and estimates combined using Rubin's rules.

All statistical tests used an alpha of 0.05. Because we were interested in examining the effectiveness of each intervention separately and there were two correlated primary outcomes of interest (C-SOSI and HRV MESOR), P-values were adjusted for two tests separately for each intervention based on trial-phase data using the Benjamini–Hochberg procedure. Original and corrected P-values are presented only for primary outcomes. Statistical analyses were conducted using Stata Statistical Software: Release 17, except that the propensity scores were calculated using R Statistical Software: Version 4.1.1 and the CBPS package.

Results

Study Flow

As shown in Supplementary Fig. B1, 1642 eligible UMC clergy were invited to participate; prior to 1 March 2020, 390 consented. Assignment to interventions was preference-based with

310 (79.5%) indicating a preference for one single intervention or ambivalence between two interventions, and 80 (20.5%) indicating no preference among the three interventions. Of the 310 clergy with a preference, 207 were assigned to their uniquely preferred intervention (or, if ambivalent between two interventions, to one of the two) to occur immediately (with 144 going forward to participate in interventions), while 103 were randomly assigned to the waitlist (with 62 providing baseline data). The 80 clergy without any preference were randomly assigned between the three interventions and the waitlist control group. Of the 80 without any preference, 60 (20 per intervention) were randomly assigned to interventions, of whom 40 went on to participate and provide baseline data, and 20 of the 60 were randomly assigned to the waitlist, (with 9 providing baseline data). Thus, the waitlist analysis sample consisted of 71 participants. Of the 390 who consented before 1 March 2020, 135 ultimately declined to participate in the trial. 255 provided survey data and 157 also provided HRV data; of the 255, 8% (n = 20) were excluded from HRV data collection and 31% (n = 78) had missing HRV data. See Supplementary Table A2 for exclusion reasons and counts.

Among the 255 participants, 71 were assigned to the waitlist control group and 184 stated a unique intervention preference at study registration. Of the 184, 14 changed their preference between study registration and intervention launch. Reasons for not receiving one's initial preferred intervention were participant-driven, possibly due to new circumstances from the pandemic, our switch from in-person to online-only delivery, or logistical reasons such as specific intervention dates and times. Of the 184 trial participants who received an immediate intervention, 40 did not have any preference at study registration and were randomly assigned between 3 interventions, and 144 received an intervention that they indicated they preferred either at study registration or at intervention launch. Thus, 22% of immediate intervention participants were fully randomized both to intervention type and to immediate vs waitlist, and 78% were randomized only to immediate vs waitlist.

As shown in Supplementary Fig. B1, 47 participants randomized to waitlist participated in post-waitlist interventions and were included in sensitivity analyses. An additional 63 individuals consented to participate in the interventions after 1 March 2020 (see Supplementary Fig. B2), of whom 50 participated and were included in sensitivity analyses.

Sample characteristics

Table 1 reports comparisons of baseline characteristics between participants assigned to the immediate interventions vs waitlist. Participants were evenly split between females (47.5%) and males (52.5%), with a mean age of 53.9 (SD = 11.2) years, predominantly white (90.6%) or Black (5.9%), married or cohabitating (89.4%), and serving a church (82.4%). Supplementary Tables B2 and B3 report baseline characteristics of the subsample of participants that provided HRV data and pooled trial and observational participants. Supplementary Table B1 depicts comparisons of characteristics of Selah trial phase participants with those of the eligible population (i.e. those who participated in the 2019 panel survey). Females were more likely to

participate in the Selah study than males (47.5% in Selah vs 33.7% in the panel study, P < .001). Those who were bi-vocational (P < .001), had BMI <30 (P = .043) and self-reported diabetes (P = .002) were less likely to participate in the Selah study; those with elevated depressive symptoms were more likely to participate (P = .005).

Engagement

Participation in intervention sessions was high across all three active interventions (Supplementary Table B4). The median size of intervention groups was eight (Interquartile Range, IQR: 4,10) across a total of 21 groups. For Stress Proofing, 87.5% attended three out of four main sessions and 62.5% of participants attended four out of four main sessions, with more than half attending the optional follow-up session. For the Daily Examen, 95.8% of participants attended all three sessions, with more than half attending at least one optional follow-up session. For MBSR, the median participant attended seven of eight sessions.

The text message response rate for all interventions peaked at three weeks: Stress Proofing, 85.9%; Daily Examen, 90.2%; MBSR, 86.5%; see Supplementary Fig. B3. At 24 weeks, the combined text response rate was 70.0%. Reports of any practice the day before at 24 weeks were high: Stress Proofing, 68%; Daily Examen, 72%; MBSR, 78%. For those reporting any practice, the average reported minutes of MBSR practice per day across the 24 weeks was 28.4 (SD=16.8) minutes.

Propensity scores

Supplementary Table A3 shows the variable specifications for propensity score models. Distributions of propensity scores indicated good overlap, with overdispersion at low propensity scores across all intervention types (Supplementary Fig. B4). Comparisons of unweighted descriptive statistics between treatment conditions indicated relatively good a priori balance between participants in most sociodemographic characteristics (Supplementary Table B5). Those in Stress Proofing and MBSR exhibited higher baseline levels of stress, anxiety, and depressive symptoms than their counterparts in Daily Examen and the waitlist. After propensity score adjustment, differences between treatment arms were systematically reduced.

Primary outcome analyses

Stress symptoms (C-SOSI) The baseline adjusted mean C-SOSI score across all arms was 1.01 (95% CI: 0.95, 1.06) (Table 2). Mean differences in C-SOSI scores between an active treatment arm and waitlist at 12 weeks were most pronounced for MBSR [Mean Difference (MD) = −0.30, 95% CI: −0.41, −0.20; P < .001] and Stress Proofing participants (MD = −0.27, 95% CI: −0.40, −0.14; P < .001), with less evidence of substantial differences for Daily Examen participants

(MD = −0.08, 95% CI: −0.21, 0.05). By 24 weeks post-baseline, differences between active treatment arms and the waitlist control grew more substantial, with stronger evidence of differences between Daily Examen and waitlist participants (MD = −0.24, 95% CI: −0.41, −0.08), in addition to larger differences between Stress Proofing and MBSR vs waitlist participants.

HRV

Baseline adjusted mean MESOR across all arms was 24.6 ms (95% CI: 22.2, 27.1) and amplitude was 7.88 ms (95% CI: 6.54, 9.22, Table 2). At 12 weeks, participants in MBSR had a modest 3.32 ms higher mean MESOR (95% CI: 0.21, 6.44; P = .036) and a 1.94 ms higher amplitude (95% CI: 0.17, 3.72) than similar participants in waitlist control. There was no evidence of a significant difference in MESOR or amplitude for the other intervention arms vs the waitlist.

Secondary and exploratory outcome analyses

Anxiety symptoms (GAD-7)

The baseline adjusted mean GAD-7 score across all arms was 4.92 (95% CI: 4.50, 5.34, Table 2). Similar to the stress symptoms results, participants in Stress Proofing (MD = −1.29 points, 95% CI: −2.27, −0.26 at 12 weeks; MD = −1.45, 95% CI: −2.68, −0.23 at 24 weeks) and MBSR (MD = −1.85 points, 95% CI: −2.66, −1.04 at 12 weeks; MD = −2.40, 95% CI: −3.43, −1.36 at 24 weeks) had fewer symptoms of anxiety at 12 and 24 weeks than comparable participants in the waitlist control, with the greatest differences for MBSR. Differences between Daily Examen and waitlist participants were modest at 12 weeks (MD = −0.51, 95% CI: −1.42, 0.40) and more pronounced at 24 weeks (MD = −1.36, 95% CI: −2.49, −0.24).

Depressive symptoms (PHQ-8)

The baseline adjusted mean PHQ-8 score across all arms was 5.51 (95% CI: 5.11, 5.91, Table 2). Similar to the results for anxiety symptoms, at 12 weeks, Stress Proofing and MBSR participants had fewer symptoms of depression than comparable participants in the waitlist control (Stress Proofing: MD = −1.72, 95% CI: −2.82, −0.63; MBSR: MD = −2.07, 95% CI: −3.15, −1.00), with sustained differences between participants in each intervention vs waitlist control at 24 weeks (Stress Proofing: MD = −1.60, 95% CI: −3.00, −0.20; MBSR: MD = −2.46, 95% CI: −3.69, −1.24). Differences between Daily Examen and waitlist participants were observed at both 12 weeks (MD = −1.21, 95% CI: −2.37, −0.05) and 24 weeks (MD = −1.48, 95% CI: −2.88, −0.08).

Subgroup analyses

Participants who received a uniquely preferred intervention (n = 174) largely resembled those who did not (n = 81) in terms of baseline sociodemographic and clinical characteristics. Being white (P = .01), having a lower BMI (P = .001), and having higher levels of physical activity

(P = .015) were correlated with engaging in an initially preferred intervention (Supplementary Table B11).

When estimates were stratified by receipt of an initially uniquely preferred intervention, there was little evidence of heterogeneity in treatment effects for stress, anxiety, and depressive symptom outcomes at 12 weeks (Figs. 1-4, Supplementary Fig. B5). For Stress Proofing outcomes at 24 weeks, there was some evidence that those who did not have a unique preference (or, in rarer cases, whose initial preference did not match their intervention group) had a larger treatment effect (lower stress and depressive symptoms compared to waitlist) than those who had and received their initial uniquely preferred intervention (stress symptoms interaction effect = 0.20, 95% CI: 0.08, 0.32; depressive symptoms interaction effect = 1.84, 95% CI: 0.44, 3.24).

Sensitivity analysis

When missing outcome and covariate values were imputed using MICE methods, results for stress, anxiety, and depressive symptoms at 12 weeks did not differ substantially from complete case estimates (Supplementary Table B9). However, magnitudes of treatment effects at 24 weeks were attenuated when missing values were imputed, suggesting that participants with lower stress, anxiety, and depressive symptom scores may have been more likely to drop out of the study between 12 and 24 weeks. In contrast, multiply imputed estimates for HRV outcomes moved in the direction of better HRV outcomes across all interventions. Results of the trial data were largely similar to pooled results with observational and post-waitlist data (Supplementary Table B10).

Table 1 Baseline characteristics of active intervention and waitlist study arms for trial period participants

	Waitlist	**Stress**	**Daily Examen**	**Mindfulness-Based stress reduction**	**Total**
	(N = 71)	(N = 48)	(N = 71)	(N = 65)	(N = 255)
Age (in years)					
Mean (SD)	54.8 (10.1)	53.4 (10.6)	54.5 (11.8)	52.6 (12.1)	53.9 (11.2)
Sex					
Female	32 (45.1%)	30 (62.5%)	28 (39.4%)	31 (47.7%)	121 (47.5%)

Male	39 (54.9%)	18 (37.5%)	43 (60.6%)	34 (52.3%)	134 (52.5%)
Race and Ethnicity					
White and not Latinx	65 (91.5%)	45 (93.8%)	60 (84.5%)	61 (93.8%)	231 (90.6%)
African American and not Latinx	5 (7.0%)	3 (6.3%)	5 (7.0%)	2 (3.1%)	15 (5.9%)
Asian/Pacific American, Native American, Latinx, multiracial, and other	1 (1.4%)	0 (0.0%)	6 (8.5%)	2 (3.1%)	9 (3.5%)
Marital and habitation status					
Not married, or married but separated/divorcing	8 (11.3%)	8 (16.7%)	6 (8.5%)	5 (7.7%)	27 (10.6%)
Married or cohabitating	63 (88.7%)	40 (83.3%)	65 (91.5%)	65 (91.5%)	228 (89.4%)
Any children live at home					
No	38 (53.5%)	28 (59.6%)	39 (54.9%)	29 (46.0%)	134 (53.2%)
Yes	33 (46.5%)	19 (40.4%)	32 (45.1%)	34 (54.0%)	118 (46.8%)
Clergy appointment					
Pastoral charge	58 (81.7%)	39 (81.3%)	61 (85.9%)	52 (80.0%)	210 (82.4%)
Extension or other	13 (18.3%)	9 (18.8%)	10 (14.1%)	13 (20.0%)	45 (17.6%)
Bi-vocational					

No	68 (95.8%)	45 (93.8%)	70 (98.6%)	63 (96.9%)	246 (96.5%)
Yes	3 (4.2%)	3 (6.3%)	1 (1.4%)	2 (3.1%)	9 (3.5%)
Hours per week worked as UMC full-time clergy					
Mean (SD)	49.5 (9.9)	50.2 (9.6)	49.4 (11.1)	49.0 (11.4)	49.5 (10.5)
Stress from congregation(s)/work from Nov 2019 to registration, [0–3]					
Mean (SD)	1.8 (0.7)	2.1 (0.8)	1.9 (0.7)	1.9 (0.8)	1.9 (0.7)
Financial stress					
Not at all or slightly stressful	48 (68.6%)	32 (68.1%)	51 (71.8%)	41 (66.1%)	172 (68.8%)
Moderately, very, or extremely	22 (31.4%)	15 (31.9%)	20 (28.2%)	21 (33.9%)	78 (31.2%)
Alcoholic drink intake					
None	23 (32.9%)	16 (34.0%)	27 (38.6%)	20 (32.8%)	86 (34.7%)
Occasional drink (not every week)	17 (24.3%)	13 (27.7%)	17 (24.3%)	13 (21.3%)	60 (24.2%)
1–2 drinks	11 (15.7%)	7 (14.9%)	13 (18.6%)	12 (19.7%)	43 (17.3%)
3–6 drinks	11 (15.7%)	8 (17.0%)	6 (8.6%)	11 (18.0%)	36 (14.5%)
about a drink a day	6 (8.6%)	2 (4.3%)	4 (5.7%)	4 (6.6%)	16 (6.5%)
more than a drink a day	2 (2.9%)	1 (2.1%)	3 (4.3%)	1 (1.6%)	7 (2.8%)
Self-reported current heavy alcohol use					

No	68 (100.0%)	46 (97.9%)	68 (97.1%)	59 (98.3%)	241 (98.4%)
Yes	0 (0.0%)	1 (2.1%)	2 (2.9%)	1 (1.7%)	4 (1.6%)
Caffeinated beverage intake per day					
None	9 (12.9%)	5 (10.6%)	10 (14.3%)	6 (9.8%)	30 (12.1%)
1 cup	18 (25.7%)	11 (23.4%)	13 (18.6%)	17 (27.9%)	59 (23.8%)
2–3 cups	31 (44.3%)	23 (48.9%)	37 (52.9%)	27 (44.3%)	118 (47.6%)
4–5 cups	10 (14.3%)	7 (14.9%)	8 (11.4%)	8 (13.1%)	33 (13.3%)
6 or more cups	2 (2.9%)	1 (2.1%)	2 (2.9%)	3 (4.9%)	8 (3.2%)
Metabolic equivalents (METs) per week					
Mean (SD)	71.0 (89.4)	62.0 (70.2)	74.7 (93.2)	56.7 (76.3)	66.7 (83.9)
BMI					
Mean (SD)	30.3 (6.9)	31.0 (6.9)	30.4 (6.6)	31.5 (7.8)	30.8 (7.0)
Obesity					
Not obese (BMI <30)	38 (53.5%)	23 (47.9%)	38 (53.5%)	32 (49.2%)	131 (51.4%)
Obese (BMI 30+)	33 (46.5%)	25 (52.1%)	33 (46.5%)	33 (50.8%)	124 (48.6%)
High blood pressure					
No (including missing)	43 (60.6%)	32 (66.7%)	47 (66.2%)	45 (69.2%)	167 (65.5%)
Yes, current or history	28 (39.4%)	16 (33.3%)	24 (33.8%)	20 (30.8%)	88 (34.5%)
Diabetes					

No (including missing)	58 (81.7%)	45 (93.8%)	61 (85.9%)	59 (90.8%)	223 (87.5%)
Yes, current or history	13 (18.3%)	3 (6.3%)	10 (14.1%)	6 (9.2%)	32 (12.5%)
Depression screens					
Negative (PHQ-8 <10)	63 (90.0%)	39 (81.3%)	62 (87.3%)	46 (71.9%)	210 (83.0%)
Positive (PHQ-8 10+)	7 (10.0%)	9 (18.8%)	9 (12.7%)	18 (28.1%)	43 (17.0%)

[a] Among all the Selah Trial participants, there were small numbers of missing values for: whether they had children at home (n = 3); financial stress (n = 5); alcohol intake (n = 7); whether they were experiencing heavy alcohol use (n = 10); caffeine intake (n = 7); and PHQ-8 depression screens (n = 2).

[b] In the Selah Trial sample, there are 217 full-time clergy: 62 in the waitlist arm, 40 in the Stress Proofing arm, 58 in the Daily Examen arm (3 participants missed values for number of hours worked per week), and 57 in the MBSR arm. Among these full-time clergy, 3 were missing values for the number of work hours per week.

[c] Participants engaged in a mean of 15 minutes of activity during the week and burned a total of 66.7 metabolic equivalents (METs) which woudl equate to consuming ~3.502L of additional oxygen given that 1 MET is ~3.5 ml of oxygen consumed per kilogram bodyweight per minute (ml/kg/min). Among the Selah Trial participants, 1 participant value for metabolic equivalents was missing.

Discussion

We performed a partially randomized, participant-preference, waitlist control study to evaluate the effectiveness of three potentially stress-reducing interventions on self-reported symptoms of stress and one biological marker of parasympathetic nervous system activity among an occupational group of United Methodist clergy with challenging work. We measured engagement and analyzed trial data from 255 participants who underwent randomization.

Separate profiles of evidence emerged when each intervention was independently compared to the waitlist control. Participants allocated to MBSR evidenced improvement in self-reported symptoms of stress, anxiety, and depressed mood, and in HRV MESOR and amplitude from pre- to postintervention at 12 weeks, with improvements in symptoms of stress, anxiety, and depressed mood maintained at 24 weeks. Participants allocated to Stress Proofing evidenced improvements in symptoms of stress, anxiety, and depressed mood from pre- to post-intervention at 12 weeks and maintained at 24 weeks but did not evidence change in HRV parameters. Improvement in symptoms of stress, anxiety, and depressed mood among participants who completed the Daily Examen were not evidenced until 24 weeks with no change evidenced in HRV parameters. Stated alternately, participation in MBSR resulted in stable and enduring improvement in self-reported and physiological correlates of stress, while participation in Stress Proofing resulted in enduring improvement in self-reported correlates of stress, and participation in the Daily Examen resulted in delayed improvements in self-reported correlates of stress.

We included MBSR as a gold standard stress management intervention. The enduring improvements in symptoms observed among participants allocated to MBSR are consistent with a systematic review of the effects of MBSR interventions among nonclinical samples that reported significant improvements in symptoms of stress, anxiety, and depression when compared to nonactive control conditions. In the current study, these findings may be attributed to regular engagement across 24 weeks; study participants practiced an average of 28 minutes per day, which is consistent with a systematic review of 43 studies of MBSR and mindfulness-based cognitive therapy which found an average home practice duration of about 30 minutes per day, six days per week. The beneficial outcomes are noteworthy with an average practice of 28 minutes on days of any practice.

We hypothesized that the Daily Examen may influence symptoms of stress through mechanisms similar to those for MBSR. Both practices develop the ability to observe and describe thoughts, feelings, and behaviors, which may help bring attention to the present (or past 24 hours) as opposed to worry about the future. Furthermore, both practices encourage nonreactivity toward thoughts and feelings, which may promote calmness. Statistically significant improvements in self-reported correlates of stress were not evidenced among participants allocated to the Daily Examen until 24 weeks. Among participants who provided engagement data, Daily Examen practice was high throughout the 24 weeks. A prayer practice may be particularly acceptable for populations such as clergy and other people of faith. The Daily Examen, at just 15 minutes/day, may be more feasible than MBSR to sustain past 24 weeks. Only one empirical study on the Daily Examen other than our pilot study has been published focusing on positive emotions; participants randomly assigned to practice the Examen increased in self-transcendent positive emotions but not eudemonic motivation after 2 weeks. The current study is the first to empirically investigate the effects of the Daily Examen on stress outcomes.

Consistent with studies on stress inoculation therapy, Stress Proofing—a set of stress inoculation, breathing, and walking exercises plus lifestyle changes—led to post-intervention and enduring improvements in self-reported correlates of stress. Stress Proofing exercises differed from mindfulness-based exercises in that participants were not explicitly taught to direct their thoughts to the present. For example, in the MBSR awareness of breath exercise, participants were taught to notice their breath without changing it as a way to focus on the present, whereas in Stress Proofing, participants were taught to change their breathing (e.g. triangle, square, and deep breathing) without intentional present focus. The goal of the Stress Proofing breathing exercises was to lower heart rate and impact the autonomic nervous system; other studies have found that diaphragmatic breathing decreases diastolic and systolic blood pressure, salivary cortisol, respiratory rate, and anxiety symptoms, although researchers have called for more high-quality studies to determine clinical utility. Stress Proofing did not evidence improvement in HRV at 12 weeks despite its physical practices. Like the Daily Examen, suggested daily practice for Stress Proofing was 15 minutes.

Only MBSR participants evidenced a statistically significant improvement in HRV, a noninvasive biological marker of the strength of the parasympathetic nervous system as measured at the sinoatrial node [65]. Following completion of MBSR, participants evidenced improvement in two long term HRV parameters: the MESOR which reflects trait-like activity of the parasympathetic nervous system, and amplitude which reflects higher day-to-day variability of parasympathetically mediated HRV. Previous research evaluating the effect of MBSR on HRV is equivocal, with one systematic review that identified 19 randomized controlled trials (RCTs) evaluating the effect of mindfulness-based interventions on HRV, reporting no statistically significant difference pre- to post-intervention (Hedges' g = 0.38, 95% CI = −0.014 to 0.77) [28]. It is important to note that there were only two RCTs included in this review that evaluated MBSR which included long-term HRV (i.e. 24 hours) as an outcome, one within 168 people who lived with fibromyalgia and another within 19 people who experienced benign heart palpitations. As such, the current trial provides the most robust evidence to date for the effect of MBSR on long-term HRV (capturing the experience of everyday stressors) when delivered with fidelity among a community sample.

Confidence in the effect of MBSR on HRV observed in the present study is strengthened for five reasons: (i) propensity score adjustment was performed to ensure that results could not be explained by variation in baseline characteristics associated with HRV, such as age and gender, BMI, and physical activity; (ii) HRV was quantified using long-term (v) practice data indicated high engagement. We can conclude that MBSR resulted in improvement in parasympathetic cardiac control with good confidence, although we cannot make similar conclusions about activation of the sympathetic nervous system as no measure was collected.

The Selah study is one of few to conduct a behavioral trial of a specific prayer practice. The Daily Examen was preferred by more participants than the other two interventions, although it had attenuated effects even at 24 weeks. However, given the degree of acceptability shown across interventions, it may be possible to increase the amount of prayer time or combine the Daily Examen with elements of MBSR to lead to physiological benefit, although this would require further testing. While beyond the scope of this manuscript, the current dataset could be used to evaluate potential mechanisms (i.e. mediators and moderators) of treatment effects, including degree of engagement.

In this trial, there was clear engagement for all three practices during the initial 3-month intervention period and persisting for an additional 3 months. Such engagement could have been driven by a number of factors. First, interventions were delivered virtually which can improve accessibility, particularly among those with competing demands (e.g. caring for dependents). On the other hand, remote delivery lacked the incentives of social connection and bonding that can occur with in-person connection with colleagues. Of note, a review of RCTs reporting on the effects of online preventative mindfulness interventions for nonclinical populations reported high attrition in over half of the eight included trials [69]. Second, the need for stress symptom reduction during the COVID-19 pandemic may have contributed to increased engagement. Finally, the feasibility and acceptability of the interventions evaluated may have contributed to

increased engagement. In support of this, we observed similarly high levels of engagement in our pre-pandemic pilot study in which practices were taught in-person in retreat settings [23]. Contrary to our expectations, in exploratory sub analyses, having a unique intervention preference vs being indifferent between two or more interventions (predominant case) or receiving an intervention that was not one's initial preference (rarer case) was largely not related to study outcomes. Outcomes of interest were only observed to vary by preference type among two comparisons. Stress Proofing participants without a unique preference for Stress Proofing observed greater reduction in symptoms of stress at 24 weeks, suggesting greater durability of effects when a unique preference was not present. The observation that both at baseline and each endpoint, outcomes of interest were largely similar between those that had and received their unique preferences and those that were indifferent between two or more interventions or did not receive their unique preference suggests that simply having an a priori preference and being able to choose an intervention corresponding to that preference did not correlate with the intervention's eventual effectiveness. This may be the case if, e.g., lack of having a preference and yet still enrolling in the study was associated with being more open to instruction and practice, and/or expecting great benefit from any stress management intervention offsetting any "advantage" that having and receiving a unique preference might convey. Future study is required to confirm these findings.

This study had several limitations. We did not conduct intervention fidelity checks, although the interventions were delivered by certified MBSR instructors, and all instructors followed the same materials throughout the trial. HRV data was collected at 12 but not 24 weeks, which limits our understanding of the durability of treatment effects on HRV parameters. HRV data collected at 24 weeks would have been particularly interesting for the Daily Examen intervention which saw improvements in self-reported correlates of stress at 24 weeks and not at 12 weeks. Observed results are limited to individuals in a single profession (clergy from a single, predominantly white denomination) in a single geographic area. We do not know the exact mechanisms explaining observed differences, as we did not measure exposures to stressors or cognitive appraisals associated with stressors, and because our outcomes were temporally removed from acute stressors.

The treatment personalization of partially randomized preference trials may enhance translation and external validity of a study, although estimates of treatment effects may suffer from similar biases as may be seen in an observational study due to confounding between characteristics that give rise to preference and the outcome under study. Some previous studies suggest that analyzing outcomes of a study on the subsample of participants who were indifferent to their intervention allocation can provide an unbiased treatment effect estimate [69]. Our sample size did not allow for this analysis. Another approach—used in the current study—is to control for confounding characteristics between preference and the outcome, which may lead to estimates with more precision and a relatively unbiased estimate of the treatment effect [70]. The propensity score adjustment approach can only correct confounding bias if the

propensity model is correctly specified, which cannot be definitively confirmed. Treatment assignment and partial randomization occurred prior to baseline data collection which limited our ability to characterize participants who withdrew from the study and leaves the possibility that randomization may have affected the baseline level of outcomes. Finally, the partially randomized structure of intervention assignment means there were limited cases in which a participant received an intervention that they specifically did not initially prefer, thus we could not explicitly measure the effect of preference on study outcomes, only a selection effect.

This study also had several strengths. We evaluated both emerging and well-validated interventions and did so using a study design that accounted for participants' preferences, which in behavioral trials have the potential to affect engagement and outcomes. Instructors were well-trained and consistent across cohorts within each intervention, increasing the chances of consistent within-treatment delivery (i.e. a single instructor delivered the Stress Proofing content, the same two instructors co-taught each Daily Examen class, and although there were four MBSR instructors, all met well-established MBSR certification standards). The trial methods were adapted to COVID-19 in a way that approximated real-world conditions, and practice was measured for 24 weeks with high daily response rates. We collected self-report and physiological measures for a relatively large sample. HRV is not subject to expectancy effects and provides confidence in the improvements seen in the MBSR participants. Simultaneously, the self-reported symptom outcomes are useful in indicating that participants across interventions felt noticeably better, even if it was a placebo effect. The fact that Daily Examen participants did not have significantly improved scores across multiple self-reported outcomes until 24 months also provides some confidence in the timing of feeling better for each intervention for the average participant. Further, we were able to use existing survey data from the study population to describe participants who selected into the trial, informing generalizability and possible bias.

Table 2 Between arm, mixed effects regression estimated differences in outcomes between immediate intervention and waitlist by follow-up time point for trial period participants

	Time point	Stress Proofing		Daily Examen		Mindfulness-based stress reduction	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Survey outcomes							
C-SOSI (scale range 0–4): n = 233; N = 603	Baseline Mean[a]	1.00 [0.93, 1.07]	1.01 [0.95, 1.06]	1.00 [0.93, 1.07]	1.01 [0.95, 1.06]	1.00 [0.93, 1.07]	1.01 [0.95, 1.06]
	12 weeks	−0.28 [−0.41, −0.16]	−0.27 [−0.40, −0.14][b]	−0.07 [−0.21, 0.07]	−0.08 [−0.21, 0.05][c]	−0.27 [−0.39, −0.16]	−0.30 [−0.41, −0.20][d]
	24 weeks	−0.31 [−0.49, −0.13]	−0.31 [−0.48, −0.14]	−0.22 [−0.40, −0.05]	−0.24 [−0.41, −0.08]	−0.37 [−0.50, −0.23]	−0.42 [−0.55, −0.28]
GAD-7 (scale range 0–21): n = 233; N = 603	Baseline Mean[a]	4.89 [4.37, 5.41]	4.92 [4.50, 5.34]	4.89 [4.37, 5.41]	4.92 [4.50, 5.34]	4.89 [4.37, 5.41]	4.92 [4.50, 5.34]
	12 weeks	−1.22 [−2.18, −0.26]	−1.29 [−2.27, −0.31]	−0.47 [−1.37, 0.44]	−0.51 [−1.42, 0.40]	−1.62 [−2.43, −0.80]	−1.85 [−2.66, −1.04]
	24 weeks	−1.33 [−2.55, −0.10]	−1.45 [−2.68, −0.23]	−1.33 [−2.48, −0.17]	−1.36 [−2.49, −0.24]	−2.13 [−3.21, −1.05]	−2.40 [−3.43, −1.36]
PHQ-8 (scale range 0–24): n = 232 participants; N = 600 observations	Baseline Mean[a]	5.43 [4.86, 6.01]	5.51 [5.11, 5.91]	5.43 [4.86, 6.01]	5.51 [5.11, 5.91]	5.43 [4.86, 6.01]	5.51 [5.11, 5.91]
	12 weeks	−1.82 [−2.95, −0.69]	−1.72 [−2.82, −0.63]	−1.22 [−2.39, −0.05]	−1.21 [−2.37, −0.05]	−2.06 [−3.13, −0.98]	−2.07 [−3.15, −1.00]
	24 weeks	−1.85 [−2.96, −0.74]	−1.60 [−3.00, −0.20]	−1.52 [−2.87, −0.18]	−1.48 [−2.88, −0.08]	−2.46 [−3.62, −1.31]	−2.46 [−3.69, −1.24]
HRV outcomes							
MESOR (unit = 1 MS): n = 142; N = 253	Baseline Mean[a]	24.6 [22.2, 27.1]	24.6 [22.2, 27.1]	24.6 [22.2, 27.1]	24.6 [22.2, 27.1]	24.6 [22.2, 27.1]	24.6 [22.2, 27.1]
	12 weeks	−1.99 [−5.33, 1.35]	−0.64 [−3.92, 2.65][e]	1.64 [−2.81, 6.08]	1.38 [−2.75, 5.52][f]	3.31 [0.01, 6.61]	3.32 [0.21, 6.44][g]
Amplitude (unit = 1 MS): n = 142; N = 253	Baseline Mean[a]	7.74 [6.34, 9.14]	7.88 [6.54, 9.22]	7.74 [6.34, 9.14]	7.88 [6.54, 9.22]	7.74 [6.34, 9.14]	7.88 [6.54, 9.22]
	12 weeks	0.84 [−1.13, 2.81]	1.33 [−0.68, 3.33]	1.21 [−0.80, 3.21]	1.08 [−0.68, 2.84]	1.86 [0.07, 3.65]	1.94 [0.17, 3.72]

Number of participants (n) and number of observations (N) for the regression model of each outcome are reported.

Depending on the survey outcome, 46 Stress Proofing participants, 63 Daily Examen participants, and 57–58 MBSR participants were compared to 66 waitlist control participants. For each of the HRV outcomes, 30 Stress Proofing participants, 36 Daily Examen participants, and 37 MBSR participants were compared to 39 waitlist control participants.

[a] Regression estimated mean rather than difference between arms. Means are constrained to be the same across all arms at baseline.

[b] $P < .001$ before and after Benjamini–Hochberg correction.

[c] $P = .222$ before correction; $P = .445$ after correction.

[d] $P < .001$ before and after correction.

[e] $P = .704$ before and after correction.

[f] $P = .511$ before and after correction.

[g] $P = .036$ before and after correction.

Intraclass correlation coefficients are reported in Results Supplementary Table B12.

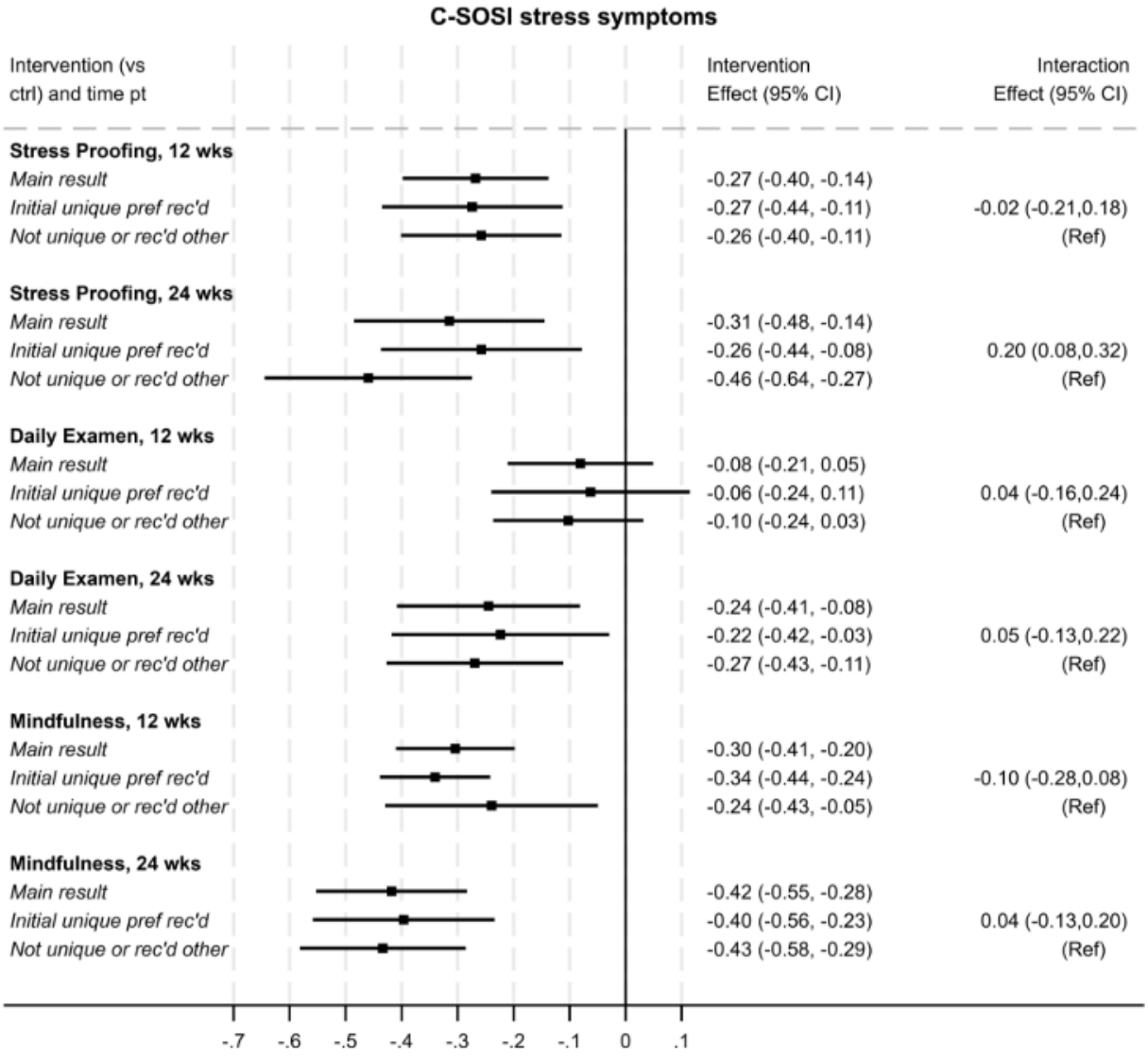

Figure 1 Subgroup analysis of heterogeneity of treatment effects on stress symptoms by intervention preference type

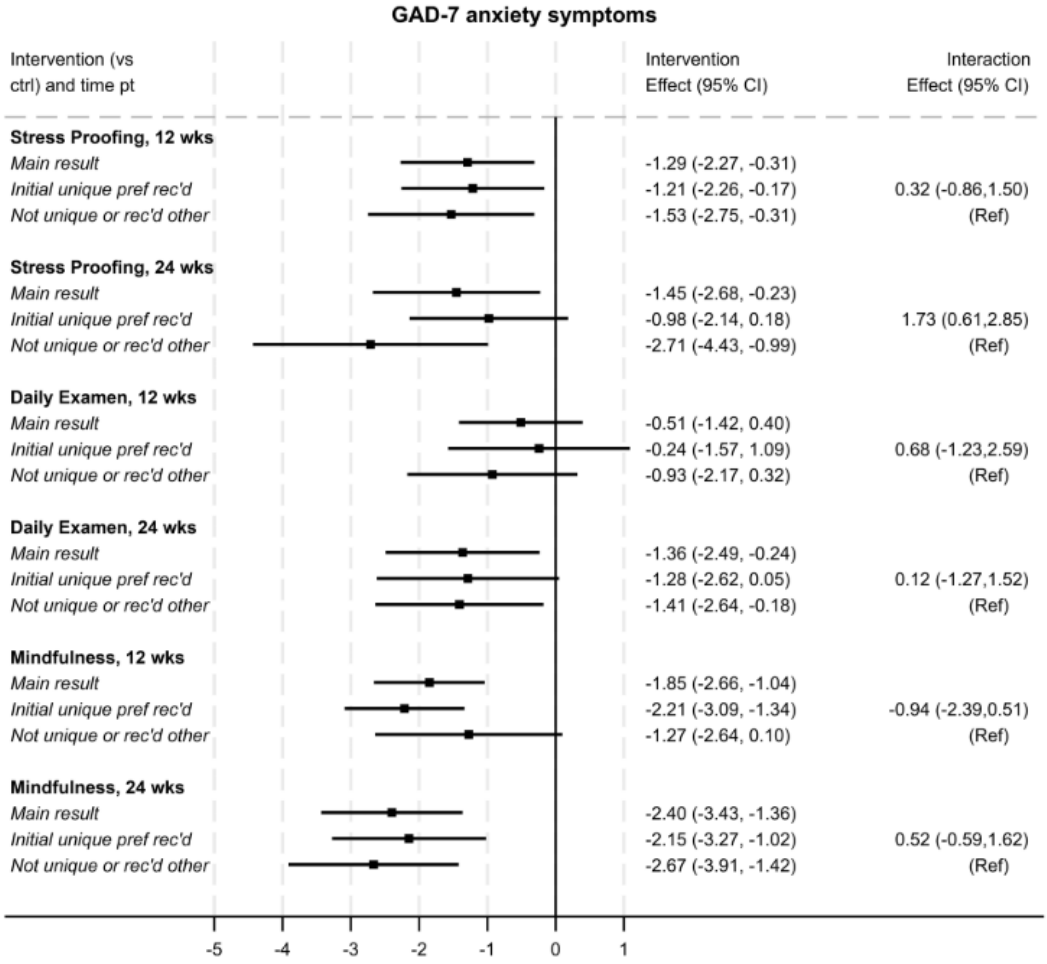

Figure 2 Subgroup analysis of heterogeneity of treatment effects on anxiety symptoms by intervention preference type

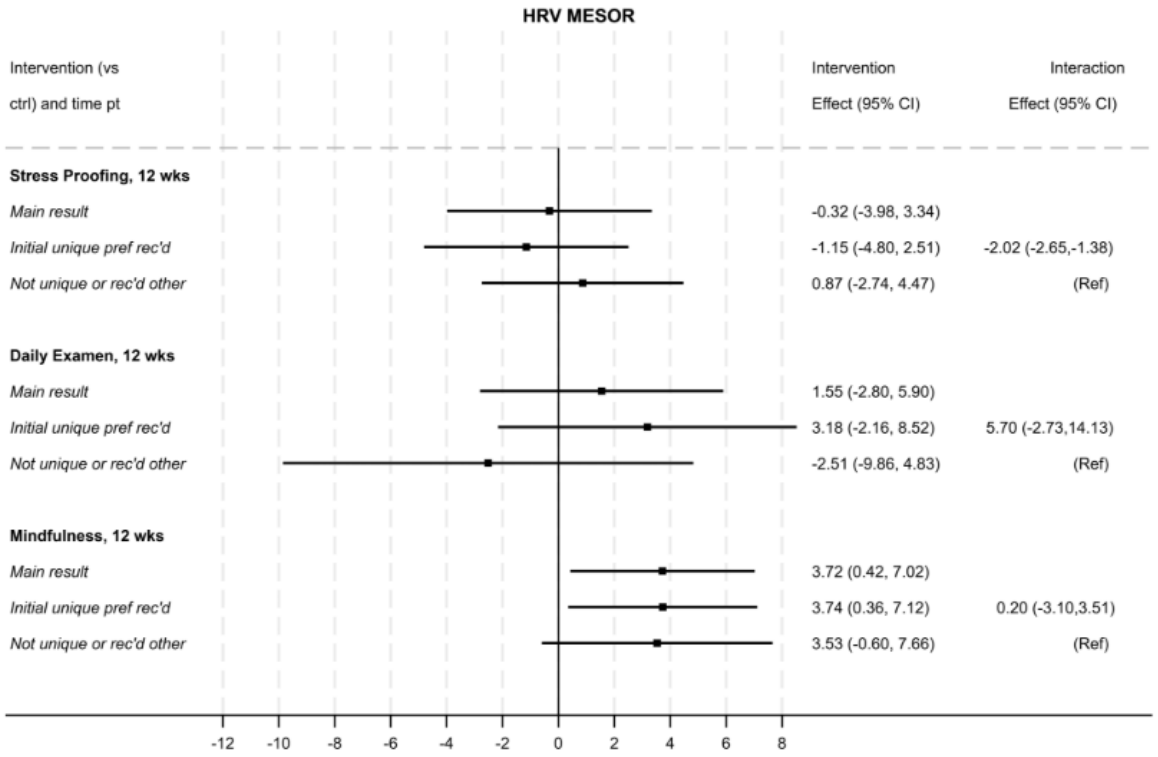

Figure 3 Subgroup analysis of heterogeneity of treatment effects on heart rate variability MESOR by intervention preference type

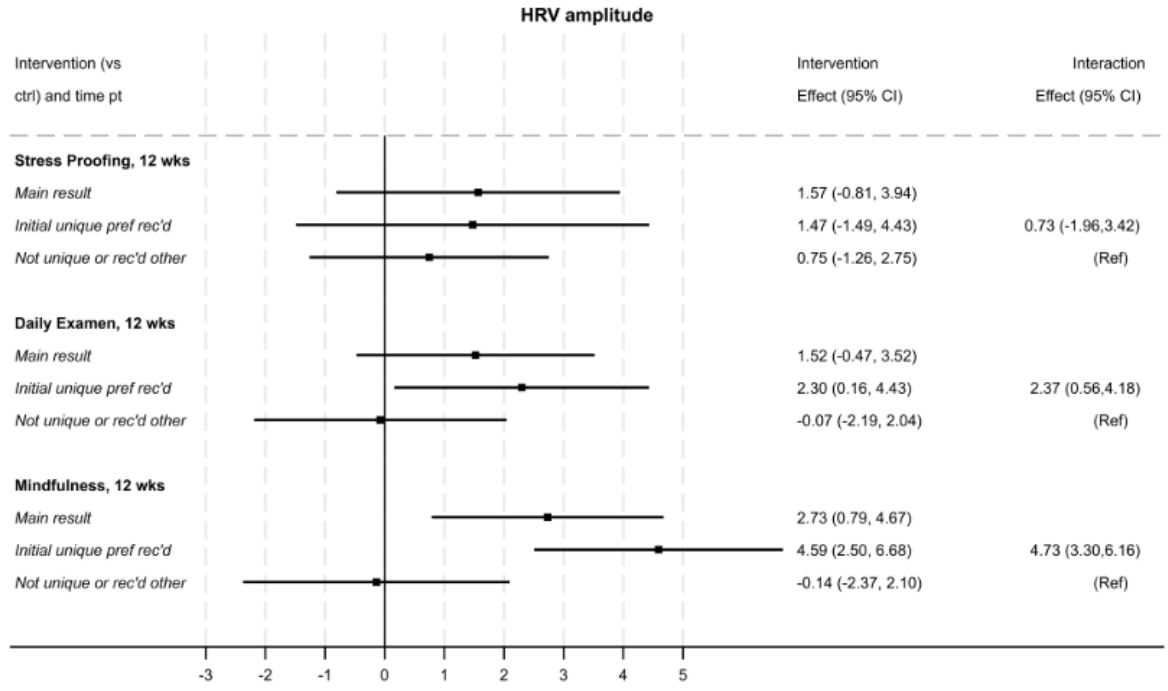

Figure 4 Subgroup analysis of heterogeneity of treatment effects on heart rate variability amplitude by intervention preference type

Conclusion

The Selah Stress Management Trial tested three separate behavioral interventions compared to a control group. Participants who provided text message data engaged in each intervention with great frequency and with enduring practice through 24 weeks, indicating that each intervention was acceptable to clergy who are busy and engage in challenging emotional and administrative activities. Each intervention group experienced improvements in self-reported correlates of stress at 24 weeks, during a particularly stressful time of the COVID-19 pandemic, and with two of the interventions requiring only 15 minutes of practice per day. Only MBSR, which, when practiced and reported, was practiced on average 28 minutes per day, resulted in statistically significant improvement in HRV from pre- to post-intervention. Despite a robust literature of MBSR's effects on self-reported correlates of stress, this is the first study to show a significant improvement on long-term HRV parameters. These findings show the strongest evidence of improvement for MBSR, although Stress Proofing and the Daily Examen may be considered if individuals do not prefer MBSR. There is a clear need for stress symptom reduction among many occupational groups; these findings provide evidence of effectiveness of three manageable and scalable interventions for United Methodist clergy.

Supplementary data

Supplementary data is available at *Translational Behavioral Medicine* online.

Acknowledgments The authors greatly appreciate the numerous instructors who designed and implemented the interventions: Glenn Murphy, Mark Shaw, Karen Keen, Riitta Whaley, Maya McNeilly, Phyllis Hicks, and Julie Kosey. We are grateful to Claire Cusick for recruitment, Beth Stringfield, Blen Biru, Jessica Choi, and Sofa Labrecque and 19 students for interview, text message, and heart rate variability data collection, and Elizabeth Wallack for processing and quantification of heart rate variability data files. We appreciate consultations with Liz Turner and Joe Egger of the Duke Global Health Institute Research Design and Analysis Core and Carl Weisner of Duke Divinity School, literature review from Celia Hybels and Anna Holleman, and manuscript comments from the Clergy Health Initiative Writing Group. This study was funded by the Rural Church Area of The Duke Endowment.

Funding Sources This study was funded by the Rural Church Area of The Duke Endowment (grant number ORC-2096-SP). Conflict of interest statement. None declared.

Human Rights All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee

and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the Duke University IRB, protocol 2019-0238.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Welfare of Animals This article does not contain any studies with animals performed by any of the authors.

Transparency Statements Study Registration: This study was registered after the study began. The study is registered at ClinicalTrials.gov, https:// clinicaltrials.gov/ct2/show/NCT04625777. Analytic Plan Preregistration: The analysis plan was not formally preregistered. Data Availability: De-identified data from this study are not available in an a public archive. De-identified data from this study will be made available (as allowable according to institutional IRB standards) by emailing the corresponding author. Analytic Code Availability: Analytic code used to conduct the analyses presented in this study are not available in a public archive. They may be available by emailing the corresponding author. Materials Availability: Some of the materials used to conduct the study are presented in a public archive: https://spiritedlife.org/program-info/the-practices/. Other materials may be available by emailing the corresponding author.

References

Cohen S, Gianaros PJ, Manuck SB. A stage model of stress and disease. Perspect Psychol Sci 2016; 11:456–63. https://doi.org/10.1177/1745691616646305

Lovallo WR. Stress and Health. 3rd edn. SAGE Publications, Los Angeles, CA, 2016. https://doi.org/10.4135/9781071801390

Richardson S, Shaffer JA, Falzon L et al. Meta-analysis of perceived stress and its association with incident coronary heart disease. Am J Cardiol 2012; 110:1711–6. https://doi.org/10.1016/j.amjcard.2012.08.004

Cohen S, Janicki-Deverts D, Miller GE. Psychological stress and disease. JAMA 2007; 298:1685–7. https://doi.org/10.1001/jama.298.14.1685

Proeschold-Bell RJ, LeGrand SH. High rates of obesity and chronic disease among United Methodist clergy. Obesity 2010; 18:1867– 70. https://doi.org/10.1038/oby.2010.102

Webb BL, Chase K. Occupational distress and health among a sample of Christian clergy. Pastoral Psychol 2019; 68:331–43. https://doi.org/10.1007/s11089-018-0844-y

Lindholm G, Johnston J, Dong F et al. Clergy wellness: an assessment of perceived barriers to achieving healthier lifestyles. J Relig Health 2016; 55:97–109. https://doi.org/10.1007/s10943-014-9976-2

Proeschold-Bell R, Miles A, Toth M et al. Using effort-reward imbalance theory to understand high rates of depression and anxiety among clergy. J Prim Prev 2013; 34:439–53. https://doi.org/10.1007/s10935-013-0321-4

Proeschold-Bell R, Steinberg DM, Yao J et al. Using a holistic health approach to achieve weight-loss maintenance: results from the Spirited Life intervention. Transl Behav Med 2020; 10:223–33. https://doi.org/10.1093/tbm/iby117

Clemow LP, Pickering TG, Davidson KW et al. Stress management in the workplace for employees with hypertension: a randomized controlled trial. Transl Behav Med 2018; 8:761–70. https://doi.org/10.1093/tbm/iby018

Janssen MHMG, Heerkens YF, Kuijer W et al. Effects of mindfulness-based stress reduction on employees' mental health: a systematic review. PLoS One 2018;13: e0191332. https://doi.org/10.1371/journal.pone.0191332

Kriakous SA, Elliott KA, Lamers C et al. The effectiveness of mindfulness-based stress reduction on the psychological functioning of healthcare professionals: a systematic review. Mindfulness 2021; 12:1–28. https://doi.org/10.1007/s12671-020-01500-9

Abbott RA, Whear R, Rodgers LR et al. Effectiveness of mindfulness-based stress reduction and mindfulness-based cognitive therapy in vascular disease: a systematic review and meta-analysis of randomised controlled trials. J Psychosom Res 2014; 76:341–51. https://doi.org/10.1016/j.jpsychores.2014.02.012

Parsons CE, Crane C, Parsons LJ et al. Home practice in mindfulness-based cognitive therapy and mindfulness-based stress reduction: a systematic review and meta-analysis of participants' mindfulness practice and its association with outcomes. Behav Res Ther 2017; 95:29–41. https://doi.org/10.1016/j.brat.2017.05.004

Saunders T, Driskell JE, Johnston JH et al. The effect of stress inoculation training on anxiety and performance. J Occup Health Psychol 1996; 1:170–86. https://doi.org/10.1037//10768998.1.2.170

Kashani F, Kashani P, Moghimian M et al. Effect of stress inoculation training on the levels of stress, anxiety, and depression in cancer patients. Iran J Nurs Midwifery Res 2015; 20:359–64. https://www.ncbi.nlm.nih.gov/pubmed/26120337

Hofmann SG, Asnaani A, Vonk IJJ et al. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012; 36:427–40. https://doi.org/10.1007/s10608-012-9476-1

Meichenbaum D. Stress inoculation training: a preventative and treatment approach. In: Lehrer PM, Woolfolk RL, Sime WE (eds.), Principles and Practice of Stress Management. The Guilford Press: New York, NY, 2005, 497–516. 19.

Lindsay EK, Creswell JD. Mechanisms of mindfulness training: monitor and acceptance theory (MAT). Clin Psychol Rev 2017; 51:48–59. https://doi.org/10.1016/j.cpr.2016.10.01120

Crosswell AD, Lockwood KG. Best practices for stress measurement: how to measure psychological stress in health research. Health Psychol Open 2020; 7:205510292093307–12. https://doi.org/10.1177/205510292093307221

Thayer JF, Ahs F, Fredrikson M et al. A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health. Neurosci Biobehav Rev 2012;36:747–56. https://doi.org/10.1016/j.neubiorev.2011.11.009

Kim H, Cheon E, Bai D et al. Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry Investigation 2018; 15:235–45. https://doi.org/10.30773/pi.2017.08.1723

Proeschold-Bell RJ, Eagle DE, Tice LC et al. The Selah pilot study of spiritual, mindfulness, and stress inoculation practices on stress-related outcomes among United Methodist clergy in the United States. J Relig Health 2023; 62(4):2686–710. https://doi.org/10.1007/s10943023-01848-x

Kemp AH, Quintana DS, Felmingham KL et al. Depression, comorbid anxiety disorders, and heart rate variability in physically healthy, unmedicated patients: implications for cardiovascular risk. PLoS One 2012;7: e30777. https://doi.org/10.1371/journal.pone.0030777 .

Kemp AH, Brunoni AR, Santos IS et al. Effects of depression, anxiety, comorbidity, and antidepressants on resting-state heart rate and its variability: an ELSA-Brasil cohort baseline study. Am J Psychiatry 2014; 171:1328–34. https://doi.org/10.1176/appi.ajp.2014.13121605

Järvelin-Pasanen S, Sinikallio S, Tarvainen MP. Heart rate variability and occupational stress-systematic review. Ind Health 2018; 56:500–11. https://doi.org/10.2486/indhealth.2017-0190.

Tsuji H, Larson MG, Venditti FJ et al. Impact of reduced heart rate variability on risk for cardiac events: the Framingham Heart Study. Circulation 1996; 94:2850–5. https://doi.org/10.1161/01.cir.94.11.2850

Brown L, Rando AA, Eichel K et al. The effects of mindfulness and meditation on vagally mediated heart rate variability: a meta-analysis. Psychosom Med 2021; 83:631–40. https://doi.org/10.1097/PSY.0000000000000900

Dunne PJ, Lynch J, Prihodova L et al. Burnout in the emergency department: randomized controlled trial of an attention-based training program. J Integr Med 2019; 17:173–80. https://doi.org/10.1016/j.joim.2019.03.009

Owens JE, Schorling J, Plews-Ogan M et al. A randomized controlled trial evaluating Mindfulness-Based Stress Reduction (MBSR) for the treatment of palpitations: a pilot study. Int J Cardiol 2016; 223:25–7. https://doi.org/10.1016/j.ijcard.2016.08.183

Grossman P, Deuring G, Walach H et al. Mindfulness-based intervention does not influence cardiac autonomic control or the pattern of physical activity in fibromyalgia during daily life. Clin J Pain 2017; 33:385–94. https://doi.org/10.1097/AJP.0000000000000420

Ma J, Lewis MA, Smyth JM. Translational behavioral medicine for population and individual health: gaps, opportunities, and vision for practice-based translational behavior change research. Transl Behav Med 2018; 8:753–60. https://doi.org/10.1093/tbm/ibx032

Walter SD, Turner R, Macaskill P et al. Beyond the treatment effect: evaluating the effects of patient preferences in randomised trials. Stat Methods Med Res 2017; 26:489–507. https://doi.org/10.1177/0962280214550516

Carlson LE, Zelinski EL, Speca M et al. Protocol for the MATCH study (Mindfulness and Tai Chi for cancer health): a preference-based multi-site randomized comparative effectiveness trial (CET) of Mindfulness-Based Cancer Recovery (MBCR) vs. Tai Chi/Qigong (TCQ) for cancer survivors. Contemp Clin Trials 2017; 59:64–76. https://doi.org/10.1016/j.cct.2017.05.015

The United Methodist Church Online Directory & Statistics 2021. 2021 Annual Conference Membership and Attendance in the US. http://www.umdata.org/UMFactsHome.aspx (19 January 2024, date last accessed).

Chaves M, Hawkins M, Holleman A et al. Introducing the fourth wave of the National Congregations Study. J Sci Study Relig 2020; 59:646–50. https://doi.org/10.1111/jssr.12684

Steensland B, Park JZ, Regnerus M et al. The measure of American Religion: toward improving the state of the art. Social Forces 2000; 79(1):291–318. https://doi.org/10.1093/sf/79.1.291

Sidani S, Epstein DR, Bootzin RR et al. Assessment of preferences for treatment: validation of a measure. Res Nurs Health 2009; 32:419–31. https://doi.org/10.1002/nur.20329

StataCorp. Stata Statistical Software. Stata Press; Release 16. 2019.

Tice LC, Eagle DE, Rash JA et al. Rationale and preferences trial design to test three approaches to reduce stress symptoms among clergy. Trials 2021; 22:892. https://doi.org/10.1186/s13063-021-05845-x

Kabat-Zinn J. Meditation is Not What You Think. New York: Hachette Books, 2018. 42. Thibodeaux MS. Reimagining the Ignatian Examen: Fresh Ways to Pray from Your Day. Chicago, IL: Loyola Press, 2015.

Carlson LE, Thomas BC. Development of the Calgary symptoms of stress inventory (C-SOSI). Int J Behav Med 2007; 14:249–56. https://doi.org/10.1007/BF03003000

Penwell L. Validation of the Calgary symptoms of stress inventory (C-SOSI) for predicting adherence to a stress reduction technique. Doctoral Dissertation, West Virginia University. ProQuest Dissertation, 2012. https://search.proquest.com/docview/1112056408

Penttilä J, Helminen A, Jartti T et al. Time domain, geometrical and frequency domain analysis of cardiac vagal outflow: effects of various respiratory patterns. *Clin Physiol* 2001; 21:365–76. https://doi.org/10.1046/j.1365-2281.2001.00337.x

Johnson SU, Ulvenes PG, Øktedalen T et al. Psychometric properties of the general anxiety disorder 7-Item (GAD-7) scale in a heterogeneous psychiatric sample. *Front Psychol* 2019; 10:1713. https://doi.org/10.3389/fpsyg.2019.01713

Kroenke K, Strine TW, Spitzer RL et al. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord* 2009; 114:163–73. https://doi.org/10.1016/j.jad.2008.06.026

Harris PA, Taylor R, Minor BL et al.; REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019; 95:103208. https://doi.org/10.1016/j.jbi.2019.103208

Tarvainen MP, Niskanen J, Lipponen JA et al. Kubios HRV—heart rate variability analysis software. Comput Methods Programs Biomed 2014; 113:210–20. https://doi.org/10.1016/j.cmpb.2013.07.024

Vansteelandt S, Daniel RM. On regression adjustment for the propensity score. Stat Med 2014; 33:4053–72. https://doi.org/10.1002/sim.6207

Imai K, Ratkovic M. Covariate balancing propensity score. J R Stat Soc Series B: Stat Methodol 2014; 76:243–63. https://doi.org/10.1111/rssb.12027

Flight L, Allison A, Dimairo M et al. Recommendations for the analysis of individually randomised controlled trials with clustering in one arm—a case of continuous outcomes. BMC Med Res Methodol 2016; 16:165. https://doi.org/10.1186/s12874-016-0249-5

Coffman CJ, Edelman D, Woolson RF. To condition or not condition? Analysing "change" in longitudinal randomised controlled trials. BMJ Open 2016;6: e013096. https://doi.org/10.1136/bmjopen-2016-013096

White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. Stat Med 2011; 30:377–99. https://doi.org/10.1002/sim.4067

Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley & Sons, 2004. https://doi.org/10.1002/9780470316696

Rubin M. When to adjust alpha during multiple testing: a consideration of disjunction, conjunction, and individual testing. *Synthese* 2021; 199:10969–11000. https://doi.org/10.1007/s11229-021-032 76-4

Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. Ann Stat 2001; 29:1165–88. https://doi.org/10.1214/aos/1013699998

StataCorp. Stata Statistical Software. Stata Press; Release 17. 2021.

R Core Team. R Statistical Software. R Foundation; 2021.

Querstret D, Morison L, Dickinson S et al. Mindfulness-based stress reduction and mindfulness-based cognitive therapy for psychological health and well-being in nonclinical samples: a systematic review and meta-analysis. Int J Stress Manag 2020; 27:394– 411. https://doi.org/10.1037/str0000165

Alsubaie M, Abbott R, Dunn B et al. Mechanisms of action in mindfulness-based cognitive therapy (MBCT) and mindfulness-based stress reduction (MBSR) in people with physical and/or psychological conditions: a systematic review. Clin Psychol Rev 2017; 55:74–91. https://doi.org/10.1016/j.cpr.2017.04.008

Curlee MS, Ahrens AH. An exploratory analysis of the Ignatian Examen: impact on self-transcendent positive emotions and eudaimonic motivation. J Posit Psychol 2022; 18:733–42. https://doi.org/10.1080/17439760.2022.2109197

Balban MY, Neri E, Kogon MM et al. Brief structured respiration practices enhance mood and reduce physiological arousal. Cell Rep Med 2023; 4:100895. https://doi.org/10.1016/j.xcrm.2022.100895

Hopper SI, Murray SL, Ferrara LR et al. Effectiveness of diaphragmatic breathing for reducing physiological and psychological stress in adults: a quantitative systematic review. JBI Database Syst Rev Implement Rep 2019; 17:1855–76. https://doi.org/10.11124/JBISRIR-2017-003848

Berntson GG, Thomas Bigger J Jr, Eckberg DL et al. Heart rate variability: origins, methods, and interpretive caveats. *Psychophysiology* 1997; 34:623–48. https://doi.org/10.1111/j.1469.8986.1997. tb02140.x

Umetani K, Singer DH, McCraty R et al. Twenty-four hour time domain heart rate variability and heart rate: relations to age and gender over nine decades. *J Am Coll Cardiol* 1998; 31:593–601. https://doi.org/10.1016/s0735-1097(97)00554-8

Yi SH, Lee K, Shin D et al. Differential association of adiposity measures with heart rate variability measures in Koreans. *Yonsei Med J* 2013; 54:55–61. https://doi.org/10.3349/ymj.2013.54.1.55

Felber DD, Ackermann-Liebrich U, Schindler C et al.; Sapaldia Team. Effect of physical activity on heart rate variability in normal weight, overweight and obese subjects: results from the SAPALDIA study. *Eur J Appl Physiol* 2008; 104:557–65. https://link.springer.com/article/10.1007/s00421-008-0800-0

Walter SD, Bian M. Relative efficiencies of alternative preference-based designs for randomised trials. *Stat Methods Med Res* 2020; 29:3783– 803. https://doi.org/10.1177/0962280220941874

Gemmell I, Dunn G. The statistical pitfalls of the partially randomized preference design in non-blinded trials of psychological interventions. *Int J Methods Psychiatr Res* 2011; 20:1–9. https://doi.org/10.1002/mpr.326