Research Article

Slaying the Hydra: Integrative Wellness Training Affects Ten Complementary Measures of Stress

Alissa J. Mrazek,1 Michael D. Mrazek,2 Brianna K. Morseth,2 Kaita L. Mrazek,2 Benjamin W. Mooneyham,2 Steve Cole,3 and Jonathan W. Schooler2

1Department of Psychology, The University of Texas at Austin, Austin, TX 78712, USA
2Department of Psychological & Brain Sciences, University of California, Santa Barbara, CA 93106, USA
3David Geffen School of Medicine, University of California, Los Angeles, CA 90095, USA

Received 22 December 2020; Revised 17 April 2021; Accepted 8 June 2021

Abstract The stress someone experiences is influenced by their health, relationships, emotions, habitual patterns of thought, and dozens of other factors. Stress also has wide-ranging consequences on a person’s body, mind, and relationships. This complexity suggests that dramatic and sustainable reductions in stress may require an approach that effectively addresses numerous underlying causes simultaneously. Similarly, reductions in stress may be most meaningfully characterized when several complementary markers of stress are measured in parallel. The present randomized, controlled investigation examined changes in ten complementary measures of stress reduction among healthy young adults completing an integrative mindfulness-based health and wellness program. Substantial reductions in stress were observed across subjective, behavioral, neuroendocrine, neural, genomic, and third-party ratings of stress. Additionally, perceived stress and social anxiety were examined at a six-week longitudinal follow-up and revealed that the reductions in stress were not only enduring but continued to improve after the intervention ended.

Keywords stress; integrative medicine; cortisol; amygdala; gene expression

1. Introduction

Most people have an intuitive understanding of what stress is because they experience it on a regular basis. Yet precisely defining stress is less straightforward. Many describe stress as a state of physiological arousal and psychological distress, but this conventional understanding belies the complexity of stress as it is expressed in our bodies and minds. For scientists tasked with providing a more definitive conceptual framework for stress, a precise and widely accepted definition remains surprisingly elusive [1]. Many existing frameworks cast stress in dramatically broad terms, such as any real or perceived threat to an individual’s homeostatic equilibrium.

The challenge of defining stress is somewhat exacerbated by the existence of dozens of stress measures—including (i) validated scales of numerous kinds of stress, (ii) neuroendocrine markers like cortisol and adrenaline, (iii) neural markers like the functional connectivity (FC) patterns of the amygdala, (iv) behavioral markers like facial expressions or gestures, and (v) molecular markers like patterns of gene expression. In a simpler world, these diverse measures might all be expressions of a single underlying construct called stress. In reality, these measures converge enough to be meaningfully related but they also diverge enough to be complementary windows into how humans respond to the varied challenges of life.

Counterintuitively, meta-analytic evidence indicates that subjective and physiological markers of acute stress are significantly correlated in less than a quarter of studies [2]. There is therefore an important complementarity rather than redundancy to stress measures, and their combination provides the best opportunity to accurately characterize changes in the complex physical and psychological dimensions of stress.

Perhaps the multifaceted nature of stress is precisely the reason why it is a risk factor for so many diseases. Although intermittent stress is inevitable and even adaptive, prolonged stress is problematic. Chronic stress interferes with the downregulation of inflammation and is associated with increased risk of a wide number of illnesses including viral infections, heart disease, hypertension, type 2 diabetes, depression, and anxiety [3,4,5]. Unfortunately, high levels of chronic stress are currently widespread. Even as efforts to reduce stress among the public have increased over the last thirty years, levels of chronic stress continue to rise in the United States [6].

Although chronic stress is widespread, it can be effectively managed. Many people believe that stress originates from external circumstances, but stress is primarily a consequence of how events are interpreted or appraised [7]. When a flight is delayed, some passengers’ blood fills with cortisol while others remain calm and trust that they will work things out. Fortunately, individuals
can learn to proactively direct their attention and shape their appraisals in order to reduce stress [8]. Meta-analytic evidence suggests that mindfulness training—which entails the systematic training of attention—is a particularly promising strategy for reducing stress [9, 10]. Because attention acts as a filter that strongly influences what we experience, mindfulness training gives individuals greater choice over how they relate to the thoughts and evaluations that perpetuate emotional stress. Theoretically, mindfulness training may equip individuals with tools to down-regulate their stress response to repeated acute stressors, which, over time, may help to mitigate levels of chronic stress.

Although mindfulness interventions reliably reduce subjective reports of stress, their impact on physiological markers of stress is less consistent [11]. Across studies, mindfulness has been linked to decreased, unchanged, and even increased salivary cortisol in response to acute stress [12, 13, 14]. Mindfulness training can even lead to increased physiological markers of stress while simultaneously reducing subjective markers [13]. So although mindfulness is a highly promising strategy for reducing stress, an optimal solution to reducing the multifaceted expression of stress in our lives may benefit from an approach that goes beyond mindfulness alone [15].

Recognizing that there are multiple promising strategies for reducing stress, some research has evaluated the impact of multifaceted interventions that combine several potent stress-reducing strategies like mindfulness, exercise, and social support [16, 17]. This research has deviated—although not always explicitly—from the central logic of experimental design, which is to isolate the effect of a targeted manipulation through either control or randomization of all extraneous factors. Although isolating the effect of a targeted manipulation has indisputable value in establishing causality, this approach inherently leads to the study of variables in relative isolation. Given that most phenomena are the result of many interacting causes, there is a risk of neglecting how multiple influences combine to have greater effects than when they are studied in isolation. This risk is of course even greater for complex phenomena like stress. Accordingly, the present study aimed to evaluate to what extent an intensive multifaceted integrative wellness program could reduce both the acute stress response and chronic stress levels of healthy adults. To effectively characterize the intervention’s effects, we employed ten complementary measures of stress spanning genomic, neural, hormonal, behavioral, and self-report levels of analysis.

2. Materials and methods

2.1. Overview of methods

Prior research using a similar integrative intervention found numerous outcomes with large effect sizes, including reductions in perceived stress with a Cohen’s $D$ of 1.6 [15]. Yet given the lack of precedent for estimating the effect size for many of the present investigation’s outcomes, we aimed to power the study to be able to detect a medium effect size. An estimated total sample size of 38 was needed to detect a between-group interaction from pre-test to post-test for a medium effect size ($D = 0.5$) with the conventional standard of 80% power, two-tailed $P < .05$, two assessment points, and 0.43 test-retest measure reliability drawn from the previous intervention study that also assessed perceived stress (computed with G’Power software) [15]. Although this sample size was sufficiently powered to detect changes resulting from the intensive training, it was not large enough for correlational analyses examining the relationships among these changes. Accordingly, those correlational analyses were not conducted and are not included in this report.

Thirty-eight college undergraduates (16 male and 22 female; mean age: 20.38 with SD: 2.28) from the University of California Santa Barbara were recruited to participate in what was described as an intensive lifestyle change program focused on mindfulness, relationships, physical exercise, nutrition, sleep, and compassion. The intervention ($N = 19$) and waitlist control ($N = 19$) conditions were balanced for age and gender using adaptive covariate randomization. Inclusion criteria were (1) availability for all training and testing sessions, (2) a capacity to engage in physical exercise, and (3) no contraindications for MRI scanning. One participant in the waitlist condition withdrew from the study before the second testing session. Following the first six-week intervention, the remaining 37 participants completed a second round of testing. Two female participants from the intervention condition were excluded from stress reactivity analyses due to an adverse response during the Trier Social Stress Test that prevented completion of the task. With minor exceptions (e.g., temporary illness), all participants attended every session of the intervention. Participants received financial compensation at the rate of $10/hour for the research testing. The study protocol was approved by the Human Subject Committee at the University of California Santa Barbara. Prior to the experiment, informed consent was obtained from all participants.

2.2. Intervention

The training program was modeled after an intervention that has previously been shown to reduce subjective stress [15]. An existing report of the current training program’s effects on mindfulness and dynamic FC during fMRI is reported elsewhere [18]. The intervention convened for five and a half hours each weekday over a period of six weeks. This high level of intensity was intended to ensure that each aspect of the intervention could be implemented with sufficient frequency and duration to have a meaningful effect.
The majority of integrative wellness programs involve physical exercise [19]. In the present intervention, each day included 150 min of exercise. This included 60 min of stretching in the morning and 90 min of either pilates, yoga, or circuit training in the afternoon. This frequency and intensity of exercise has been shown to improve muscular endurance and flexibility as well as cardiovascular endurance over six-weeks [15]. In addition, each day included 60 min of formal mindfulness training where participants practiced focused attention meditation by directing their attention to a single aspect of sensory experience (e.g., the sensations of breathing or walking). Each day also included 90 min of lecture on topics related to stress, sleep, nutrition, compassion, relationships, or well-being, as well as 30 min of structured small group discussion on these topics. Each lecture contained evidence-based strategies for improving that area of life (e.g., how much sleep young adults really need and how to adjust one’s environment and lifestyle for optimal rest). Participants were given prompts for interpersonal exchanges, often encouraging them to discuss in small groups how to apply these strategies in their everyday lives. These elements were explicitly chosen because, theoretically, they may have synergistic effects (e.g., becoming more mindful is useful for stress management, since you become increasingly self-aware and notice opportunities for applying emotion regulation strategies; see [16]). Finally, participants were encouraged to limit alcohol intake to no more than one drink a day, to eat a diet of primarily whole foods, and to consistently sleep at least eight hours each night.

2.3. Waitlist control
An inevitable limitation of a multifaceted intervention is the inability to specify which aspects of the intervention produced the observed effects and by what mechanism. For this reason, strictly controlled experiments with extremely well-matched active controls are essential, even though they neglect the complex interactions that often underlie how changes occur in people’s lives. Yet for the present research question of determining the extent to which stress can be reduced using an intensive and multifaceted intervention, a waitlist control is the most appropriate choice.

A waitlist control effectively addresses effects due to developmental maturation, repeated exposure to assessments, and self-selection of participants based on pre-existing characteristics. A waitlist condition therefore controls for factors that are unrelated to the intervention, providing an accurate estimate of the effect size of the intervention as a whole. Given that the theoretical motivation for this investigation was to explore how multiple influences combine to reduce stress, effects due to expectation of improvement or therapeutic alliance—which are sometimes considered confounding effects—represent meaningful elements of the intervention. It would not only be impossible to create an active control condition precisely matched in participants’ expectations of change or interaction with an effective teacher, but doing so would misguidedly attempt to control for a meaningful element of the intervention and thereby bias the effect size estimate. To address genuinely confounding explanations for our findings, we included objective measures—third-party ratings, neuroimaging, and physiological measures like salivary cortisol and immune cell gene expression—that are not susceptible to demand characteristics.

Participants in the waitlist control condition continued their life as usual. They did not receive any training, nor did they receive any recommendations to change their alcohol intake, food consumption or sleep habits.

2.4. Overview of measures
At both pre-test and post-test, participants completed four separate testing sessions in a counterbalanced order (see Figure 1). Each testing session assessed a different dimension of stress: (1) validated scales of subjective stress during daily life, (2) acute stress reactivity during a public speaking task, (3) resting-state fMRI, and (4) gene expression of circulating immune cells. Because the measures in each of these four testing sessions require a different form of analysis, the specific analytic approach taken for each measure is described below. All data collection occurred within four days immediately before and after the intervention. Each participant completed each test at the same time of day at pre-test and post-test. Additional measures not pertinent to the present investigation will be reported in full within separate articles.

2.5. Subjective stress during daily life—measures & analysis
Validated scales measuring stress (Perceived Stress Scale) and trait level social anxiety (Social Interaction Anxiety Scale) were administered at pre-test, post-test, and the six-week follow-up [20,21]. A mixed-model analysis of variance (ANOVA) was used to examine the interaction of condition and testing session (pre/post). Paired samples t-tests examined the persistence of stress reduction from baseline to the six-week longitudinal follow-up.

2.6. Stress reactivity during acute stress—measures & analysis
At pre-test and post-test, acute stress reactivity was assessed using an adapted version of the Trier Social Stress Test (TSST). This test is a public speaking and mental arithmetic task known to reliably elicit a strong stress response [22]. Participants gave speeches and solved difficult subtraction problems in front of a video camera and two judges dressed in semi-formal business attire who remained
neutral in expression and maintained mostly unbroken eye contact with the subject throughout the task. To minimize habituation to the paradigm, participants were given different prompts for the public speaking task at pre-test and post-test (counterbalanced by participant within condition). At pre-test, two participants found the acute stress of this paradigm emotionally overwhelming. The experimenter excused them from the procedure and they were not asked to complete it again at post-test. Full details for the TSST procedure are included in Supplementary Material.

Acute stress reactivity—participant reported subjective stress
Participants also reported the degree to which they felt stressed and anxious on Visual Analog Scales (VASs) at 0 min (at arrival), 30 min (10 min after receiving the prompt for their speech and immediately before the TSST), 40 min (immediately after the TSST), and 80 min (before departure). A composite score of the mean average on the stressed and anxious items across all time points was used in a mixed-model ANOVA.

Acute stress reactivity—third-party ratings by judges & coders
Every minute throughout the eight-minute speaking task, the judges indicated the extent to which the participants appeared uncomfortable on a 1–10 scale. For each judge, a composite score of the mean rating across all time points was calculated. Intraclass correlation coefficients (ICCs) were then computed as a measure of inter-rater reliability. A “consistency” criterion was used for judges’ ratings, allowing raters to vary in their ratings as long as the variation was consistent between them. Two-way mixed-model ICC was utilized in which judges were fixed and subjects were random. Finally, a composite score of the mean rating across both judges was used in a mixed-model ANOVA.

Two separate research assistants who were blind to condition served as video coders in subsequent analysis of the footage recorded as participants underwent the TSST. Coders were asked to indicate on a scale from 1 (not at all) to 5 (extremely) whether the speaker seemed anxious throughout the video. ICC was computed as described
Acute stress reactivity—behavioral markers of stress

A set of stress-related behaviors were derived and adapted from the Ethological Coding System for Interviews [23, 24]: (1) touch hair/face/mouth/neck, (2) mouth-bite lips/twist/lips in, (3) face-furrow brow/grimace/contort, (4) fingers-twirl/flick/tense interlock, (5) armsfolded/crossed, (6) readjust clothing, (7) legs-shift/fidget/sway/bounce. Videos were cut into 15-second segments and rated on whether each category of behavior occurred within the segment. A composite score of the sum of stress behavior occurrences across all categories was calculated. ICC was computed using an “absolute agreement” criterion to assess 1-1 correspondence between coders in the rating of frequency of stress behaviors. Following confirmation of high inter-rater reliability, the mean average of stress behavior occurrence was calculated across coders and analyzed using mixed-model ANOVA.

Acute stress reactivity—cortisol

Salivary cortisol samples were collected using Sarstedt Salivettes following manufacturers recommendations. Samples were collected at 0 min (at arrival), 20 min (after a resting period to establish a baseline measure), 30 min (after 10 min of preparation for the speaking task and immediately before beginning the TSST), 40 min (immediately after the TSST), 50 min, 60 min, 70 min, and 80 min (10-minute intervals during the recovery period). Samples were stored at −20 degrees Celsius prior to shipment to Dresden LabService in Dresden, Germany, for analysis. Each participant completed the TSST at the same time of day at pre-test and post-test. The lower limit of detection for the cortisol measure was 0.4 nmol/L.

To examine the effect of the intervention on cortisol levels throughout the TSST paradigm, we used mixed-effects multilevel regression using full-information maximum likelihood estimation. Cortisol levels at post-test were predicted by condition, time of sample (treated as a categorical variable with eight levels), and cortisol levels at pre-test. We predicted that the intervention would lead to an adaptive response to the challenge of public speaking characterized by a significant cortisol response during the activity that quickly returned to baseline when the task was complete [7].

2.7. FC MRI—measures & analysis

At pre-test and post-test, participants completed a scan to examine resting-state FC for which they received the following instructions: “Please close your eyes. You do not have to think of anything in particular.” MRI images were obtained at pre- and post-testing sessions using a Siemens 3.0T Magnetom TIM TRIO (SYNGO MR B17) MRI scanner. A high-resolution T1-weighted anatomical scan was first acquired for each subject according to FreeSurfer’s recommended MPRAGE specifications for cortical thickness analyses (acquisition time = 6:03; TR = 2530 ms; TE = 3.50 ms; TI = 1100 ms; flip angle = 7°; FOV = 256 mm; acquisition voxel size = 1 × 1 × 1 mm). This was followed by a T2*-weighted echo-planar imaging (EPI) sequence resting-state scan (TR = 1200 ms; TE = 30 ms; flip angle = 90°; acquisition matrix = 64 × 64; FOV = 192 mm; acquisition voxel size = 3 × 3 × 5 mm; 22 interleaved slices; 480 volumes).

Structural (TI) data processing

Cortical surface reconstruction was performed on T1 scans using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/). For each subject, nonlinear transformation from T1 to the 2 mm MNI152 template was calculated using ANTs (http://stnava.github.io/ANTS/).

Resting-state fMRI (EPI) data processing

The first four volumes of each EPI sequence were removed to eliminate potential effects of scanner instability. Slice timing of the EPI images was performed using AFNI’s 3dTshift, followed by motion correction of the images using AFNI’s 3dvolreg. Affine coregistration of the mean EPI image and T1 volume was then calculated using FreeSurfer’s BBRegister. Brain, cerebrospinal fluid (CSF), and white matter masks were extracted after FreeSurfer parcellation and transformed into EPI space using BBRegister. Coregistered EPI images were then masked using the brain mask. Principal components of physiological noise were estimated using CompCor [25], where a joined white matter and CSF mask and voxels of highest temporal variance were used to extract two sets of principal components (i.e., aCompCor and tCompCor); motion and intensity outliers in the EPI sequence were also discovered based on intensity and motion parameters using ArtDetect (http://www.nitrc.org/projects/artifact_detect). All time series data were then denoised using a GLM model with the motion parameters, CompCor components, and intensity outliers used as regressors. Finally, resultant images were smoothed using a 5 mm full width half minimum (FWHM) kernel, high-pass (0.01 Hz) and low-pass (0.1 Hz) filters were applied, and nonlinear normalization warping from subject functional/anatomical space to 2 mm MNI space was computed using Advanced Normalization Tools (ANTS).
Preprocessed fMRI images were imported into CONN, a statistical parametric mapping (SPM) toolbox implemented within MATLAB. rs-FC maps were calculated for each subject and session, separately for each amygdala seed region. The individual FC maps were then used within group-level analyses.

fMRI statistical analyses

Whole-brain analyses employing analysis of variance (ANOVA) on data from randomized controlled trials carry a risk of producing findings that are not driven by the experimental manipulation but rather by a combination of chance baseline differences and chance divergence between conditions over time [30]. This risk is reduced by using stringent corrections for multiple comparisons, but can be further minimized using a statistical procedure that first detects regions showing significant changes over time in the intervention condition and then subsequently examines these regions at pre-test and post-test across both conditions [15, 18, 31]. We employed this approach and then examined the correlation of our rs-FC results with behavioral data to provide converging evidence that the neuroimaging findings reflected meaningful changes in brain function.

For each of the amygdala-seeded FC maps (i.e., using either the left or right hemisphere amygdala seed regions), a paired t-test was performed comparing participants from the intervention condition across the pre-testing and post-testing scanning sessions. Whole-brain analyses were conducted on the seed-based FC maps, correcting for multiple comparisons using topological false discovery rate (FDR) [32]. The voxelwise significance threshold was set at \( P < .001 \) and cluster forming threshold was set at \( P < .05 \) (FDR corrected). Mean values for clusters reaching significance in the initial paired t-test were then extracted for all subjects from both the pre-test and post-test scans. These values were subjected to a mixed model ANOVA to determine whether there was a statistically significant interaction between condition and testing session (\( P < .05 \)). Finally, for clusters that demonstrated a significant interaction within the ANOVA, follow-up paired t-tests were performed to confirm that there was neither a difference in cluster values between the intervention and waitlist-control groups at pre-testing nor a difference within the waitlist-control group from pre-testing and post-testing. Subsequently reported results passed all of the described criteria for statistical significance.

2.8. Gene expression—measures & analysis

The biological effects of stress can affect the expression of genes in immune cells, providing an additional biological measure of intervention impact. There are two broad innate immune response gene expression programs in immune cells: a proinflammatory gene program (e.g., effective against extracellular pathogens such as bacteria) and an anti-viral gene program (effective against intracellular pathogens such as viruses). Activation of the sympathetic nervous system (SNS) during stress leads to the upregulation of proinflammatory immune response genes and the downregulation of antiviral immune response genes [33].

To quantify these effects, we assessed proinflammatory and antiviral gene expression in 4 mL whole blood samples collected pre- and post-intervention. Each participant provided the blood sample at the same time at pre-test and post-test between 9:00 am and 11:15 am. We used the TELiS bioinformatics system to quantify activity of proinflammatory transcription factors (NF-κB) and antiviral transcription factors (Interferon Response Factors; IRFs), as well as transcription factors mediating the biological impacts of the body’s two major physiological stress response pathways involving the SNS (CREB) and the hypothalamus-pituitary-adrenal (HPA)-axis (glucocorticoid receptor). Genome-wide transcriptional profiling was carried out as previously described [34]. Briefly, total RNA was extracted from whole blood samples drawn into PAXgene RNA tubes (Qiagen RNeasy), tested for suitable mass (Nanodrop ND1000) and integrity (Agilent TapeStation), converted to fluorescence-tagged cRNA (Ambion TotalPrep), and hybridized to Illumina HT-12 v4 bead arrays in the UCLA Neuroscience Genomics Core Laboratory, following the manufacturers’ standard protocols. Raw gene expression data were quantile normalized and log2 transformed for analysis using linear statistical models to estimate the magnitude of differential change from pre-test to post-test among participants in the intervention versus waitlist control conditions while controlling for age, sex, smoking, and alcohol consumption. Immune cell genomics analyses excluded anyone reporting infectious disease symptoms at either time point and anyone taking medications indicating potential infection (e.g., antibiotics) or immunosuppressive/anti-inflammatory medications. TELiS promoter sequence-based bioinformatics analysis was used to test three genomic hypotheses [35] by comparing the prevalence of transcription factor-binding motifs (TFBMs) for targeted transcription factors in the core promoter sequences of genes found to be upregulated in association with the intervention versus upregulated in the control group (i.e., relatively downregulated in the intervention group).

We predicted that intervention participants would show (i) reduced activity of transcription control pathways involved in neuroendocrine stress responses (i.e., CREB/ATF transcription factors mediating responses to SNS activation and the glucocorticoid receptor mediating responses to cortisol), (ii) reduced activity of the proinflammatory transcription factor family, NF-κB/Rel, and (iii) increased activity of the antiviral IRF family of
transcription factors (see [33] for background on these stress-responsive patterns of gene transcription). TELiS analyses took as input a list of all genes showing > 1.25-fold differential change in intervention versus control groups, with up- and downregulated genes tested for differential prevalence of transcription factor-binding motifs (TFBMs) for CREB/ATF (TRANSFAC V$CREB_Q2 position-specific weight matrix), the glucocorticoid receptor (V$GR), NF-$κB/Rel (V$CREL), and IRF family factors (V$ISRE). As in a previous research, log2-transformed ratios of TFBM prevalence in up- versus downregulated promoters were computed for nine combinations of 3 core promoter lengths (−2500 bp, −1000 bp, and −1000 to +2000 bp relative to the RefSeq gene transcription start site) and 3 TFBM detection stringencies (TRANSFAC MatSim values of, .80, .90, and .96), with log-ratios averaged over all nine parametric combinations and tested for statistical significance using a standard error of the mean derived by bootstrap resampling of linear model residual vectors across genes (i.e., accounting for any potential correlation among genes) [36].

3. Results

3.1. Baseline equivalence of conditions
Uncorrected independent samples t-tests were used to assess the equivalence of conditions across all measures at pre-test. No significant differences between conditions were observed (all $P$'s > .13).

3.2. Subjective stress during daily life
Relative to the control, the intervention led to a significant reduction in both perceived stress and trait level social anxiety of very large effect sizes from pre-test to post-test (Table 1). The reduction in perceived stress remained significant from pre-test to the six-week longitudinal follow-up ($t(18) = 5.16$, $P < .001$, $d = 1.68$). The reduction in social anxiety not only remained significant from pre-test to follow-up ($t(18) = 4.85$, $P < .001$, $d = 1.57$), but also continued to significantly improve from post-test to follow-up ($t(18) = 2.08$, $P = .05$, $d = 0.67$).

3.3. Stress reactivity during acute stress
Despite habituation to the TSST paradigm in both conditions from pre-test to post-test, the intervention led to significantly greater reduction in subjective stress during the TSST (Table 2).

The two judges physically present during the speaking task also rated participants on their discomfort while presenting. There was high inter-rater reliability, as indicated by intraclass correlation coefficients, for the judge’s ratings of discomfort at pre-test (ICC = 0.95, $P < .001$) and at post-test (ICC = 0.99, $P < .001$). Relative to the control, the intervention condition led to significantly less discomfort while speaking (Table 2).

Two separate coders also rated the video recordings of participants’ speeches for anxiety. There was high inter-rater reliability for these ratings at pre-test (ICC = 0.81, $P < .001$) and post-test (ICC = 0.89, $P < .001$). The maximum discrepancy between raters on global ratings was one point on the 5-point Likert scale. Relative to the control, the intervention led to significantly decreased ratings of perceived anxiety (Table 2).

The two coders also assessed the occurrence of stress behaviors during the speaking task. There was high

---

Table 1: Reductions in subjective stress during daily life.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Intervention Pre-test</th>
<th>Intervention Post-test</th>
<th>Waitlist control Pre-test</th>
<th>Waitlist control Post-test</th>
<th>Paired t-test</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$ (sd)</td>
<td>$M$ (sd)</td>
<td>$M$ (sd)</td>
<td>$M$ (sd)</td>
<td>$P$</td>
<td>$F$</td>
</tr>
<tr>
<td>Perceived stress (1–5 scale)</td>
<td>3.25 (0.65)</td>
<td>3.10 (0.92)</td>
<td>3.01 (0.89)</td>
<td>.67</td>
<td>12.94</td>
<td>.001</td>
</tr>
<tr>
<td>Social anxiety (1–5 scale)</td>
<td>2.45 (0.76)</td>
<td>2.10 (0.82)</td>
<td>2.10 (.070)</td>
<td>1.0</td>
<td>8.55</td>
<td>.006</td>
</tr>
</tbody>
</table>

$F$ and $P$ values are computed from a mixed-model ANOVA with condition and pre-test/post-test as factors.

Table 2: Reductions in stress reactivity during acute stress.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Intervention Pre-test</th>
<th>Intervention Post-test</th>
<th>Waitlist control Pre-test</th>
<th>Waitlist control Post-test</th>
<th>Paired t-test</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$ (sd)</td>
<td>$M$ (sd)</td>
<td>$M$ (sd)</td>
<td>$M$ (sd)</td>
<td>$P$</td>
<td>$F$</td>
</tr>
<tr>
<td>Subjective stress &amp; anxiety (1–140 scale)</td>
<td>54.10 (31.98)</td>
<td>56.55 (26.61)</td>
<td>45.83 (24.06)</td>
<td>.04</td>
<td>4.07</td>
<td>.05</td>
</tr>
<tr>
<td>3rd party rating of discomfort (judges) (1–10 scale)</td>
<td>3.70 (0.88)</td>
<td>4.78 (1.27)</td>
<td>.67</td>
<td>5.98</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>3rd party rating of anxiety (coders) (1–10 scale)</td>
<td>2.24 (0.73)</td>
<td>3.08 (1.53)</td>
<td>.85</td>
<td>5.15</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>3rd party rating of stress behaviors (# of occurrences)</td>
<td>24.47 (14.33)</td>
<td>40.70 (14.45)</td>
<td>.43</td>
<td>7.83</td>
<td>.008</td>
<td></td>
</tr>
</tbody>
</table>

$F$ and $P$ values are computed from a mixed-model ANOVA with condition and pre-test/post-test as factors.
inter-rater reliability for this measure at pre-test (ICC = 0.86, \( P < .001 \)) and post-test (ICC = 0.86, \( P < .001 \)). Relative to the control, the intervention led to significantly reduced stress behaviors during the public speaking task (Table 2).

Habituation to the TSST paradigm is common, particularly when there is relatively little time between subsequent exposures [37]. Despite the six-week delay between testing sessions, there was considerable habituation to the stress paradigm from pre-test to post-test (Figure 2). Nevertheless, there was a significant increase in cortisol at post-test across both conditions from the first sample taken upon arrival to peak cortisol levels at 50 min (\( z = 2.01, P = .044 \)). This indicates that despite habituation to the TSST paradigm, participants still experienced a significant degree of acute stress reactivity at post-test. The timing of this peak cortisol level—corresponding to 10 min into the recovery period—is consistent with the well-established delay between the experience of acute stress and the corresponding increase in salivary cortisol [38].

There was no main effect of condition on cortisol levels across all time points (\( z = 0.69, P = .491 \)). There was also no difference between conditions in cortisol levels upon participants’ arrival to the laboratory (\( z = 1.38, P = .169 \)) or throughout the public speaking task (all \( P’s > .327 \)) (Figure 2). However, a significant interaction emerged between condition and time of sample for cortisol levels at 20 min into the recovery period (\( z = 2.58, P = .010 \)). The intervention led to faster recovery of cortisol levels following the public speaking task (Figure 2). Although no longer statistically significant, this difference between conditions continued to 30 min (\( z = 1.56, P = .119 \)) and 40 min into the recovery period (\( z = 1.65, P = .099 \)).

3.4. FC MRI
We examined neural plasticity using rs-FC, which examines the coactivation of functionally integrated brain regions [39]. We used a seed-based approach to assess changes in rs-FC. The rs-FC of the left and right amygdala was examined in separate analyses. The intervention led to reduced rs-FC between the right amygdala and the right hemisphere mPFC, \( d = 0.93 \) (Figure 3). No significant changes were observed in rs-FC of the left amygdala.

3.5. Gene expression
Chronic activation of stress responses by the SNS activates a gene expression profile in circulating immune cells known as the “conserved transcriptional response to adversity” (CTRA; [33]). This genomic defense program is characterized by increased expression of genes involved in inflammation and decreased expression of genes involved in Type I interferon-mediated antiviral defenses. To determine whether the intervention might reduce stress-mediated CTRA activity, we followed previous research [34] in using TELiS bioinformatics analysis of empirical differences in gene expression to test whether the intervention might (i) decrease activity of the CREB/ATF family of transcription factors involved in beta-adrenergic signaling by the SNS, (ii) decrease activity of the proinflammatory NF-κB/Rel family of transcription factors, and (iii) increase activity of the antiviral IRF family of transcription factors.

![Figure 2](image-url): Cortisol levels during the Trier Social Stress Test. (a) Cortisol levels at pre-test. (b) Cortisol levels at post-test. Error bars represent standard error.
Figure 3: FC results using the right amygdala seed region. (a) The location of a significant cluster in the mPFC demonstrating decreased FC with the right amygdala seed region; MNI peak-coordinate: (2, 60, 6). (b) Mean Fisher r-to-z transformed FC values for the intervention and waitlist groups at pre-testing and post-testing. *P < .05, **P < .001. Error bars represent 95% confidence intervals.

Figure 4: Changes in gene expression of circulating immune cells. Data represent changes in gene expression over time for intervention participants relative to controls among the 182 genes that showed > 1.25-fold difference in the magnitude of change from pre-test to post-test in intervention participants relative to controls. *P < .05. Error bars represent standard error.

Genome-wide transcriptional profiling identified 182 genes that showed > 1.25-fold difference in the magnitude of change from pre-test to post-test in intervention participants relative to controls (1.25-fold set as an arbitrary criterion a priori; Table S1). Of these genes, 85 showed relatively greater increase over time, and 97 showed relatively greater decrease (Table S1). TELiS promoter-based bioinformatics analyses of these gene sets confirmed the expected reduction in CREB/ATF activity (mean = −1.45 ± standard error .43 log2 TFBM ratio, P = .0008) (Figure 4). These analyses also confirmed the predicted increase in activity of antiviral IRF family transcription factors (+1.40 ± .61, P = .0236). There was a nonsignificant trend towards the predicted decrease in activity of proinflammatory NF-κB/Rel transcription factors (−.51 ± .32, P = .1116). Finally, results showed no indication of differential activity of the glucocorticoid receptor transcription factor involved in mediating the genomic effects of cortisol (+.03 ± .25, P = .9042).

4. Discussion
The present findings indicate that six weeks of integrative wellness training can elicit substantial and simultaneous reductions in stress across ten complementary measures. The reductions were of predominantly large effect size and observed across both acute and chronic stress measurements. Participants exhibited both less subjective stress and fewer stress-related mannerisms during public speaking.
Third-party raters also perceived participants as less stressed and uncomfortable. As predicted, these reductions in stress were accompanied by improved cortisol recovery after the stressful event but not by reductions in cortisol response to the event itself. This pattern of salivary cortisol is consistent with the notion of positive stress in which adaptive physiological responses to challenging tasks quickly return to baseline when the task is complete [7]. Additionally, participants had decreased rs-FC between the amygdala and mPFC, as well as changes in gene expression suggesting reduced stress-related biology as indexed by immune cell biomarkers. Remarkably, self-reported stress continued to significantly decline during the six weeks following the intervention despite participants receiving no additional instruction or support. Cumulatively, these findings indicate the possibility of dramatic and enduring stress reduction even among those with nonclinical levels of stress.

A growing literature reveals that stress can impact the FC of the amygdala not only when a person is engaged in active emotional regulation but also while a person is at rest. Brief psychological stress of moderate intensity can lead to immediate changes in the amygdala’s rs-FC [26]. High levels of stress in the past—such as those experienced by individuals with post-traumatic stress disorder—can lead to more enduring changes in the amygdala’s activation patterns and rs-FC [27,28,29]. While prior work has primarily investigated the impact of high levels of stress on amygdala rs-FC, the present investigation reveals that an intervention that significantly reduces stress can also change the rs-FC of the amygdala.

Relative to the control group that did not change, the intervention led to decreased rs-FC between the amygdala and mPFC. Notably, this pattern diverges from the greater amygdala-PFC FC observed when a person engages in active and effective emotional regulation [40,41,42]. Given that stress can elicit prolonged changes in the FC patterns of the amygdala, the relative absence of stress could lead to patterns of rs-FC quite different from the FC observed during intensive emotional regulation. Indeed, healthy individuals low in stress show reduced rs-FC between the amygdala and ventral medial PFC [43]. The present study similarly observed reduced rs-FC between the right amygdala and mPFC.

Prior work has revealed important hemispheric differences between the left and right amygdalae [44]. The present study examined the left and right amygdala separately and only found changes in the rs-FC of the right amygdala. Notably, this is consistent with existing research by Kim et al. [45], which also found that low levels of stress are associated with reduced amygdala-mPFC rs-FC in the right hemisphere but not the left. One theory of hemispheric differences in the amygdala is that the left amygdala is preferentially involved in intentional control of mood whereas the right amygdala is more involved in automatic emotional regulation that relies less on explicit reflection processes [46]. The observed reductions in rs-FC between the amygdala and mPFC might therefore represent a reduction in automatic processes involved in mood regulation following a period of markedly reduced stress.

Reductions in stress are associated with not only changes in brain function but also the downregulation of genes involved in inflammation and upregulation of genes involved in antiviral defenses. These effects may be mediated by reduced SNS influences (indexed by the activity of the CREB signaling pathway) and/or reduced HPA axis influences (indexed by glucocorticoid receptor activity). As predicted, the present findings did reveal reduced CREB/ATF activity (signifying reduced SNS influences) and increased antiviral gene expression. By increasing these antiviral defenses, the stress-reducing intervention would be expected to result in a more robust immune response to viruses.

However, the intervention led to only a nonsignificant reduction in expression of proinflammatory genes. This partial manifestation of the expected pattern could be explained by the intensive exercise component of the intervention, which involved 2.5 h of daily exercise. Exercise leads to a transient increase in the expression of proinflammatory genes [47], and may therefore have canceled out the propensity for stress reduction to decrease inflammatory signaling. Stated differently, the stress-reducing intervention allowed participants to engage in a physically demanding training program while still experiencing a nonsignificant reduction in inflammation.

Although the observed changes in gene expression verify that the intervention succeeded in reducing stress-related biology as manifest in immune cell biomarkers, the health significance of these findings remains to be documented in future research (e.g., examining host resistance to disease, vulnerability to infection, etc.).

Limitations

Whether narrow or broad in focus, all intervention studies face a similar set of methodological challenges. We sought to address alternative explanations for our findings through numerous methodological controls. When possible, participants were tested in mixed groups by experimenters who were blind to condition, thereby minimizing expectancy effects arising from unconscious experimenter bias. Additionally, the inclusion of third-party ratings, neuroimaging, and physiological measures like salivary cortisol and immune cell gene expression provide compelling evidence of changes that are not susceptible to motivation effects or demand characteristics. Nevertheless, an inevitable limitation of this study—and any investigation that utilizes a multifaceted intervention—is the inability
to specify which aspects of the intervention produced the observed effects and by what mechanism. Strictly controlled experiments are therefore essential even though they must inevitably neglect the complex interactions that often underlie how changes occur in people’s lives. Multifaceted and narrow interventions must therefore complement each other to provide a comprehensive understanding of how to best reduce stress. Due to the intensity of this intervention that would make it unrealistic for most people to complete, generalizability cannot be assumed. The goal of the present research was not to design an intervention that could be utilized at scale, although this type of work is laudable. Instead, the goal was to assess the magnitude of effect sizes from an intensive intervention to better characterize the extent of change possible among healthy young adults from this particular training.

Future directions
Future research could improve and extend this work in several ways. Although we observed substantial reductions in stress among a sample of healthy young adults, studies with larger sample sizes would have the statistical power to identify factors that predispose certain individuals to benefit more than others from the intervention. Additional research is also needed to determine if similar benefits could be observed among older adults given varying levels of both stress and plasticity across the lifespan [6, 48, 49]. Future research should also compare this full multifaceted intervention with simplified versions to identify which elements are most crucial for driving the large effect sizes observed. For example, on top of the daily training, participants were asked to limit their alcohol intake, eat a diet of primarily whole foods, and sleep at least eight hours a night. Each of these recommendations is a self-regulatory challenge for participants. Future research could examine which of these health recommendations are necessary to dramatically improve stress responses. Finally, even more extensive reductions in stress may be possible with clinical samples who are experiencing high levels of chronic stress. Although there is a precedent for clinical populations to receive highly intensive and multifaceted treatments (e.g., inpatient addiction programs), challenges of feasibility may limit the investigation and application of the intensive approach used in the present research. Accordingly, an important direction for future research is to leverage technology to create scalable and affordable versions of evidence-based stress-reducing interventions that can be made available to the millions of people who struggle with chronic stress.

Conclusion
In the present investigation, an intervention combining numerous stress-reducing strategies led to substantial reductions in stress across ten complementary measures.

Like the mythical Hydra that could not be killed by cutting off just one of its many heads, stress may be most effectively defeated by reducing all of its expressions at once. As advances in technology and scientific understanding usher in even more effective strategies for stress reduction, it will be important that these interventions are evaluated using multiple complementary measures that combine to provide a clearer picture of an individual’s level of stress.

Acknowledgments This research was supported by the Institute of Education Sciences grant R305A110277, the John Templeton Foundation grant 52071, the Evolve Foundation, and the National Institute of Aging grant P30AG017265. The content does not necessarily reflect the position or policy of the U.S. government.

Authors’ contributions All authors contributed to the study design. Testing and data collection were performed by A. J. Mrazek, M. D. Mrazek, B. W. Mooneyham, and B. K. Morseth. The intervention was designed by A. J. Mrazek, M. D. Mrazek, and K. L. Mrazek. Data analysis was performed by A. J. Mrazek, M. D. Mrazek, B. W. Mooneyham, B. K. Morseth, and S. Cole. A. J. Mrazek, M. D. Mrazek, B. W. Mooneyham, and S. Cole contributed to writing the manuscript. A. J. Mrazek, M. D. Mrazek, B. W. Mooneyham, K. L. Mrazek, and J. W. Schooler contributed to editing the manuscript.

Conflict of interest The authors declare that they have no conflict of interest.

References
[12] K. W. Brown, N. Weinstein, and J. D. Creswell, Trait mindfulness modulates neuroendocrine and affective responses to social

Brief mindfulness meditation training alters physiological and neuroendocrine responses to social evaluative stress, Psychoneuroendocrinology, 44 (2014), 1–12.


Pushing the limits: cognitive, affective, and neural plasticity revealed by an intensive multifaceted intervention, Front Hum Neurosci, 10 (2016), 117.


Anxiety dissociates dorsal and ventral medial prefrontal cortex functional connectivity with the amygdala at rest, Cereb Cortex, 21 (2011), 1667–1673.


