

Smartwatch Intervention for Stress Management: A Pilot Study Amongst Undergraduate Students

By

Madissen Sitka

A Thesis submitted to the Faculty of Graduate and Postdoctoral Studies

The University of Manitoba

In the Partial fulfillment of the requirements of the degree of

MASTER OF ARTS

Department of Psychology

University of Manitoba

Abstract

Experiences of stress result in the activation of a variety of psychophysiological systems designed to address the increased demands of the environment that exceed routine demands. While such activation is initially productive in supporting one's ability to manage and address stressors, prolonged activation as a result of exposure to chronic and unpredictable stressors can lead to maladaptive physiological and psychological outcomes. Here, we investigated how a smartwatch psychoeducational intervention grounded in key principles of biofeedback might impact sleep duration and cardiovascular reactivity and recovery from a laboratory acute stressor. Thirty-three undergraduate students were randomized to either a self-guided smartwatch intervention or control group, and completed a longitudinal intervention, with pre- and post-intervention laboratory assessments of cardiovascular responsivity to acute stress. In total, 14 participants had useable data and were included in analysis. Permutation-based mixed-model analyses revealed no significant time-by-condition change in HRV responsivity, phasic HRV, or HR (p 's > 0.05) during the second administration of the acute stressor task. Furthermore, no time by condition specific improvement in sleep duration or negative affect was evident (p 's > 0.05). Feasibility analysis indicated that 54% of participants evidenced improvement in their sleep hygiene knowledge with 27% indicating a perceived improvement in sleep quality. Although results indicate promising effects of the intervention on participants perception and knowledge of sleep, findings suggest more research is needed to explore how to effectively translate key concepts of biofeedback interventions into a light-touch smartwatch intervention.

Keywords: biofeedback, smart-watch intervention, wearable technology, stress prevention, heart rate variability

Acknowledgements

Completing my master's thesis has been an exciting part of my educational journey, and there are several people whom I wish to thank. First, to my advisor Dr. Ryan Giuliano, thank you for your mentorship and support. Your belief in this work and dedication to mentorship has been fundamental to my growth and journey as a researcher. I would like to thank my committee members Dr. Sunmee Kim and Dr. Leslie Roos. Without their invaluable guidance, feedback and expertise across statistical modeling and intervention design this research would not have been possible.

I would also like to thank the participants who enrolled and engaged in this program of work for their efforts and participation. Furthermore, to the dedicated team of research assistants who worked tirelessly to run laboratory visits and collect data, a heartfelt thank you for your dedication, efforts and belief in this work. It was a privilege to work with such an amazing, conscientious group of individuals.

Finally, I am profoundly grateful to my family Kevin, Jen, Katie and late grandparents Mervin and Bernice. What a special privilege it was to grow up surrounded by people who not only taught me how to dream big but showed up every step of the way as my number one supporters. Your unconditional love and belief in the work I pursue has never gone unnoticed and I am forever grateful.

Table of Contents

List of Tables	6
List of Figures	7
Smartwatch Intervention for Stress: A Pilot Study Amongst Undergraduate Students.....	8
Biofeedback Interventions	12
Limitations of Biofeedback Interventions.....	14
Smartwatch Interventions.....	15
Guiding MINC Principles.....	16
Current Study	17
Research Hypotheses	18
Methods.....	20
Participants.....	21
Sample Size: Cardiac Analysis.....	22
Sample Size: Sleep and Negative Affect	22
Sample Size: Feasibility Analysis	23
Laboratory Visits	23
Measures	28
Data Processing	31
Analytic Plan H1:	32
Analytic Plan H2:	32
Analytic Plan H3:	34
Analytic Plan H4:	34
Permutation Approach.....	35
Results.....	36
Research Hypothesis 1: Acceptability and Feasibility	36
Research Hypothesis 2: Heart Rate Variability, Recovery and Average HR.....	39

Research Hypothesis 3: Psychological Response to TSST	44
Research Hypothesis 4: Change in Sleep	44
Discussion	47
Strengths and Limitations	51
Implications	55
Conclusions	56
References	58
Appendix A.....	76
Appendix B.....	77
Appendix C.....	81
Appendix D.....	86
Appendix F	87

List of Tables

Table 1. Cardiac Model Sample Size	22
Table 2. Sample Characteristics	25
Table 3. Reason for Exclusion from Analysis	26
Table 4. Questionnaire Administration Timeline.....	29
Table 5. TSST Administration Timeline.....	31
Table 6. Acceptability Analysis	38
Table 7. Mean Cardiac and Affect Values Across Lab Visits: Control Condition	40
Table 8. Mean Cardiac and Affect Values Across Lab Visits: Intervention Condition.....	40
Table 9. Deceptive Statistics for Total Sleep Duration by Condition and Week	41
Table 10. Permutated Beta Coefficients Cardiac Models.....	41
Table 11. Permutated Beta Coefficients Negative Affect and Sleep Models.....	46

List of Figures

Figure 1. Consort Flow Diagram for Inclusion and Exclusion Criteria	24
Figure 2. Self-Reported Perceived Efficacy of Intervention	37
Figure 3. Permutated Distribution of Interaction Coefficient for Model 1: HRV at Rest.....	42
Figure 4. Permutated Distribution of Interaction Coefficient for Model 2: HRV During Prep Period.....	42
Figure 5. Permutated Distribution of Interaction Coefficient for Model 3: HRV During Speech Task.....	43
Figure 6. Permutated Distribution of Interaction Coefficient for Model 4: HRV During Math Task.....	43
Figure 7. Permutated Distribution of Interaction Coefficient for Model 5: HR During Stressor	44
Figure 8. Permutated Distribution of Interaction Coefficient for Model 6: HRV Recovery.....	44
Figure 9. Permutated Distribution of Interaction Coefficient for Negative Affect.....	45
Figure 10. Permutated Distribution of Interaction Coefficient for Sleep Duration.....	46

Smartwatch Intervention for Stress: A Pilot Study Amongst Undergraduate Students

The human experience of stress involves a series of psychophysiological responses designed to support the body in meeting an increase in challenging demands from the environment (DeLongis et al., 1988; Kemney et al., 2003; Noushad et al., 2021; O'Connor et al., 2021). While the initial activation of such psychophysiological systems (e.g., Autonomic Nervous System, ANS) is adaptive, prolonged activation as a result of chronic or repeated stressors can lead to a variety of adverse health effects including increased susceptibility to physical ailments, anxiety, depression, and disrupted sleep quality (Cohen 2000; Kemeny 2003; O'Connor et al., 2021; Wadsworth et al., 2019). While a variety of populations are at risk for experiencing stress, college students remain vulnerable to experiences of chronic stressors due to increasing school, work, and life demands associated with early adulthood (Beiter et al., 2015; Brougham et al., 2009; Leppink et al., 2016; Mohr et al., 2014). Studies have demonstrated moderate levels of stress in samples of college students with factors such as increased alcohol and internet usage, poor sleep, and unstable romantic relationships contributing to increased susceptibility of stress amongst this age cohort (Ramonón-Arbués et al., 2020).

High rates of perceived stress amongst college students are associated with poor quality of sleep and academic performance, mental and physical health concerns, impulse control issues, and increased risk of drop out (e.g., Leppink et al., 2016; Mohr et al., 2014; Wallace et al., 2017). Left unaddressed, college students facing chronic, unmanaged stressors are at long term risk for a variety of mental and physical health concerns (depression, anxiety, cardiac health issues, inflammation, immune system dysfunction, difficulty with cognition; Das & O'Keefe, 2006; Khan et al., 2025; McEwen, 2017). Furthermore, optimism bias, which can be defined as a tendency to believe one is exempt from negative outcomes, directly influences college students awareness of health risks and likelihood of engaging in risk reducing behaviours (see Weinstein, 1982), and as such the design of cost effective, user-friendly, accessible stress management programs are needed to support this population and buffer long-term maladaptive and costly health outcomes (Lee et al., 2013; Wallace et al., 2017).

While stress management programs and psychoeducation have shown to be effective in reducing stress and accompanying mental health problems (see Amanvermez et al., 2013), few directly address the underlying psychophysiological components of stress, instead focusing on behavioural oriented mindfulness coping strategies (Dvořáková et al., 2016). While such strategies are demonstrated to be effective, without addressing the underlying physiological changes that accompany experiences of stress such interventions may not buffer maladaptive health consequences on cardiovascular health.

Stress is not only a risk factor for cardiac health issues, but a determinate for the course of heart disease and associated clinical outcomes (Esch et al., 2002). The role of stress as a large risk factor for poor cardiac health outcomes is well documented, with research demonstrating psychosocial stress may account for nearly 30% of risk associated with having a heart attack (Esch et al., 2002). Other work utilizing data from the Whitehall cohort study found that those who self-reported negatively being impacted by stress were nearly three times more likely to develop a heart attack (Nabi et al., 2013). Others point to the link between stress and self-deprecating health behaviours (e.g., alcohol and substance use) as supporting the relationship between stress and cardiac disease (Steptoe & Kivimäki, 2012). Regardless of the mechanisms by which stress influences cardiac health outcomes, scholars continue to call for increased attention towards how stress prevention and intervention should be integrated into cardiac care protocols (Esch et al., 2002; Steptoe & Kivimäki, 2012).

Biofeedback interventions are uniquely positioned to be an effective stress management tool through teaching individuals how to monitor and modify their physiological stress response in order to negate adverse health implications associated with prolonged activation (De Witte et al., 2019; Yu et al., 2018). By directly addressing the physiological changes that accompany stress responses, biofeedback interventions are designed to address both the psychological and physiological effects of stress thereby promoting physical and mental wellness. Biofeedback stress management interventions have been successfully applied across a variety of populations (see De Witte et al., 2019), however the long-term efficacy and scalability of such work is hindered by lab-based delivery and use of expensive technology (Hallman et al., 2011; Lemaire et al., 2011).

In an effort to increase the accessibility of biofeedback stress management interventions, we translated key concepts of lab-based biofeedback interventions onto wearable technology modalities. Using a pilot design and guiding principles from the Minimal Intervention Needed for Change (MINC) strategic approach, we examined how a novel smart-watch intervention, grounded in key principles of biofeedback, could modulate college student's perceived and physiological stress responses to repeated administration of the Trier Social Stress Test (TSST), a laboratory induced measure of stress. The college age sample offered an ideal population to pilot this novel intervention given their increased susceptibility to stress and positive responsiveness and engagement with stress management interventions (Amanvermez et al., 2023; Beiter et al., 2015; Brougham et al., 2009; Leppink et al., 2016; Mohr et al., 2014). Following pilot study guidelines as outlined by Abbott (2014), analyses included exploration of feasibility and acceptability outcomes and also examined preliminary effects of the intervention in order to determine whether future full scale efficacy trials are warranted.

Stress Psychophysiology

The human stress response involves the coordination of several systems (e.g., autonomic nervous systems (ANS), hypothalamic-pituitary-adrenal (HPA) axis and immune system; Kim et al., 2018; O'Connor et al., 2021; Yu et al., 2018) that are activated when an organism encounters unpredictable, uncontrollable circumstances resulting in demands that exceed their capacity (Koolhass et al., 2010; Kim et al., 2018; Noushad et al., 2021). One critical system involved in the stress response is the ANS which regulates cardiac activity through its parasympathetic branch (Kim et al., 2018; Rajendra Acharya et al., 2006). High-frequency heart rate variability (HRV; beat to beat changes in heart rate; Steffen et al., 2017) serves as a measure of activation of the parasympathetic nervous system and is often utilized as a proxy for regulatory capacity during stress, with higher resting HRV associated with better emotional and physical regulation to stressors (Kim et al., 2018; Rajendra Acharya et al., 2006; Steffens et al., 2017). While higher average resting HRV may be adaptive, flexibility of HRV during a stressor (e.g., HRV reactivity) is considered the desired response as it represents an adaptive and dynamic interaction between

withdrawal and engagement of the parasympathetic nervous system (Castaldo et al., 2015; Hamilton & Allyo, 2016; Held et al., 2021).

In addition to HRV, the Hypothalamic Pituitary Adrenal Axis (HPA) and Sympatho-Adrenal Medullary Axis (SAM) are integral components of the human stress response (O'Connor et al., 2021). The SAM response is activated in the presence of stressors, mobilizing energy and metabolic processes needed to produce the required response to the perceived threat (Aguilera, 2011; Wadsworth et al., 2019). Following activation of the SAM, the HPA helps to regulate the body back to homeostatic functioning through the release of cortisol (Wadsworth et al., 2019). While the HPA and SAM axis are fundamental neuroendocrine stress responses, they are not conducive to biofeedback interventions given the time dependent nature of cortisol and difficulty measuring catecholamines in real time (Bergquist et al., 2002; Henckens et al., 2012; Nikolajsen et al., 2001). However, the current study focused on physiological systems that can be monitored with wearable technology (e.g., HR, HRV), therefore the presents work focuses on the cardiovascular system.

In addition to the neurobiological systems described above, which are directly implicated in the stress response, exposure to stressors disrupts a variety of other tangential physiological processes such as sleep. High levels of perceived stress are associated with poor sleep quality and duration, particularly amongst college students (Lee et al., 2013; Lund et al., 2010; Wallace et al., 2017). Individuals who experience poor sleep quality also evidence greater amygdala activity and central nervous system activation in responses to adverse or negative stimuli which is thought to contribute to the relationship between poor sleep and stress responses (Williams et al., 2013). Additionally, research has demonstrated that poor sleep quality is associated with increased negative affect in response to daily hassles (Zohar et al., 2005). Disruptions in sleep quality and latency have also been associated with greater physiological arousal and greater negative affect to laboratory induced stressors (Williams et al., 2013).

While the initial activation of such psychophysiological systems is desired, prolonged activation as a result of repeated or chronic stressors increases the risk of adverse mental and physical health outcomes (e.g., anxiety, depression, heart disease; Yu et al., 2018). Given the adverse health implications

of prolonged chronic stress on the systems implicated directly in the stress response (e.g., cardiac systems) and outcomes of such activation (e.g., sleep quality) it remains important to find innovative ways to reduce repeated experiences of chronic, unpredictable stressors.

Biofeedback Interventions

Biofeedback interventions are one such way that individuals can learn to buffer the negative health implications of stress. Biofeedback interventions teach participants how to monitor and modify physiological systems commonly implicated in the stress response in an effort to gain control over such systems and improve overall health and wellness (e.g., HRV, HR, blood pressure, skin conductance; Dillion et al., 2016; De Witte et al., 2019; Steffen et al., 2017; Thabrew et al., 2018). Within these interventions, real time feedback on physiological processes is provided by a trained clinician or research assistant in order to teach participants how to advantageously modify physiological outcomes. While there are a many types of biofeedback interventions, heart rate variability biofeedback (HRV biofeedback) remains among the most prevalent, evidencing improvement in symptoms of depression, anxiety, and stress across diverse populations (Knox et al., 2011; Kudo et al., 2014; Paul et al., 2012).

The exact mechanisms underpinning the effectiveness of HRV biofeedback are not fully understood. Increased relaxation, greater internal locus of control, lowered arousal, and cognitive shifts have been proposed as potential explanations; however, increases in HRV are seen as one prominent reason why such interventions are effective (Leher & Gevritz, 2014; Steffen et al., 2017; Thabrew et al., 2018). HRV can be divided into low and high frequency bands, with changes in low-frequency bands associated with a variety of cardiac mechanisms, including parasympathetic and sympathetic nervous system activation, while high frequency changes are associated exclusively with parasympathetic activity (Schiweck et al., 2018). Biofeedback interventions target the high frequency bands of HRV through the use of breathing techniques designed to slow respiration rate while maintaining or increasing respiratory volume. In the literature, these are often referred to as breathing at resonance frequency, a specific inhalation exhalation rate of approximately six breaths per minute (Steffen et al., 2017). Participants are traditionally given ‘real-time’ physiological feedback on their respiration rate often in the form of a visual

or digital game. Participants are presented with real time data and instructed to slow their respiration until the live feedback indicates their respiration rate matches the desired resonance frequency (Dillon et al., 2016; Prinsloo et al., 2013). Studies suggest that engaging in such breathing exercises optimally activates a feedback loop whereby changes in heart rate as a result of slowed breathing facilitate subsequent increases in HRV, a key indicator of parasympathetic nervous system activity and thus health, implicated in a number of physiological and psychological conditions (e.g., depression, anxiety, chronic stress; emotional reactivity; Chalmers et al., 2014; Hartman et al., 2019; Immanuel et al., 2023; Steffen et al., 2017). Higher average HRV is associated with greater psychological and physiological regulation, whereby a blunted HRV has been implicated in emotional issues and poor response to stressors (Kim et al., 2018; Leher & Gevirtz, 2014; Steffen et al., 2017). As such, HRV biofeedback interventions can improve overall wellness by increasing participant's ability to increase their average HRV, and thereby activate their parasympathetic nervous system, through repeated engagement with breathing exercises. While extant literature has demonstrated the importance of higher average resting HRV and biofeedback's subsequent effect on this tonic physiological measure, what is less explored is biofeedback influence on phasic measures of HRV (Dillon et al., 2016; Prinsloo et al., 2013; Whited et al., 2014). Furthermore, emerging literature posits that in the context of stressors, HRV flexibility indexing the ability to dynamically engage and withdraw the PNS may be most adaptive (Giuliano et al., 2015; Knight et al., 2020; Muhtadie et al., 2015)

Heart rate variability biofeedback interventions have been applied across a number of populations for various presenting concerns (e.g., anxiety, depression, stress, emotional regulation; Karavidas et al., 2007; Knox et al., 2011; Wells et al., 2012) and in combination with evidence-based therapies (Caldwell & Steffen 2018; Gevirtz, 2024; Gevirtz, 2015). While such applications have been successful in improving well-being, an emerging application of HRV biofeedback in the context of stress intervention and prevention work is of particular interest. HRV biofeedback has been identified as being particularly well suited to help individuals regulate physiological responses to stress (DeWitte et al., 2019; Steffen et al., 2017; Yu et al., 2018). Multi and single session HRV biofeedback interventions delivered across a

variety of modalities (e.g., video games, app-based, in-person, online) have been successfully applied to reduce physiological and psychological stress responses (DeWitte et al., 2019; Yu et al., 2018, Dillion et al., 2016). Additionally, evidence suggests that the application of HRV biofeedback techniques immediately following a laboratory stressor can improve physiological and psychological recovery from stress including lowered HR and perceived stress recovery (Dillion et al, 2016; Prinsloo et al., 2013).

Limitations of Biofeedback Interventions

Despite the robust success of HRV biofeedback in buffering physiological and psychological stress responses, there remains several limitations amongst the current applications. Challenges largely center on potential difficulty translating previous intervention research on biofeedback programs into clinical settings and practice (Curran et al., 2021). Across disciplines, intervention research often does not readily translate into feasible clinical applications, with evidence suggesting that across a two-decade span, only 14% of research evidence will translate into meaningful clinical implications and practice (Curran et al., 2021; Green, 2008). Currently, there are three identified areas which may hinder successful translation of biofeedback intervention research into meaningful clinical practice.

First, many studies deliver HRV biofeedback through in person laboratory sessions led by clinicians and senior research staff that provide real time commentary on physiological data (e.g., Berry et al., 2014; Hallman et al., 2011). Such interventions are limited in their scalability and accessibility as their efficacy hinges on clinician availability, a finite and limited resource (DeWitte et al., 2019; Economides et al., 2020). Second, most studies employ expensive laboratory or handheld technology to deliver HRV biofeedback (e.g., the emWave PSR; Lemaire et al., 2011), limiting the long-term efficacy of the programs (Kotozaki et al., 2014; Lemaire et al., 2011; Reiner et al., 2008; Whited et al., 2014). The use of such an expensive modality to deliver HRV biofeedback prevents skills learned in laboratory settings from readily transferring to real world stressors as participants are left without access to the laboratory-grade technology used to implement such skills (Dillon et al., 2016; Reiner et al., 2008). Third, HRV biofeedback interventions focus on the cardiac component of the stress response, neglecting to consider how other physiological processes (e.g., sleep) contribute to and are affected by the experience of stress.

Smart Watch Interventions

Considering the previous success of biofeedback interventions coupled with the need for innovative stress management interventions, the present study created a novel smartwatch intervention for stress in order to address and mitigate the aforementioned limitations of traditional biofeedback. While we acknowledge that it is common for initial efficacy trials to utilize gold-standard, expensive equipment and involve significant clinician involvement we position this novel smartwatch intervention as a way to translate the research knowledge on biofeedback into more feasible, real world, clinical application. The smartwatch intervention for stress incorporated key principles of biofeedback (e.g., tracking and improving physiological processes) and was delivered using smartwatches and psychoeducation materials, making it less reliant on clinician involvement, lab-based technology and real-time physiological feedback.

The use of smartwatches to promote health and wellness has gained increased popularity over the past decade as technological advancement continue to be developed (Jerath et al., 2023; Kapogianni et al., 2025; Triantafyllidis et al., 2024). In the literature, smartwatch interventions are conceptualized as being uniquely positioned to help individuals take action to improve their health and well-being (Triantafyllidis et al., 2024). While biofeedback interventions are designed to teach participants to monitor and modify key physiological indices implicated in the stress response (Steffen et al., 2017), smartwatch interventions have largely focused on creating algorithms to identify and predict the onset of both physiological and mental stress and anxiety (Dai et al., 2021; Ding et al., 2024; Can et al., 2019), with some studies specifically looking at how physiological indices captured by smartwatches may relate to mental health outcomes (e.g., depression, major mental illness; Johnston et al., 2024; Sharma & Jyoti, 2022). Medication compliance, physical activity promotion, and adherence to medical management protocols have also been key focuses of such interventions (Li et al., 2023; Low et al., 2023 ; Marvel et al., 2021; Timurtas et al., 2022; Triantafyllidis et al., 2024). The design of smartwatch interventions varies across studies, with both consumer grade (e.g., Kaplan et al., 2023) and study specific smartwatches used (Li et al., 2023). Interventions are often delivered in tangent with an associated app and health data (e.g.,

personalized workout plans; Timurtas et al., 2022) or paired with app to track health metrics and adherence to protocol (Marve et al., 2021).

The present study combined key principles and ideology of biofeedback (e.g., improving and monitoring key physiological processes) and paired it with traditional smartwatch intervention methodology (e.g., delivered using wrist worn wearable technology independent of clinician involvement, real time feedback, and lab-based technology) to create a smartwatch intervention designed to improve the psychophysiological stress responses of participants. Instead of centering around predictive models of health or a notification-based intervention, this work used smartwatches to draw participant's attention to their own physiological indices, particularly HRV and sleep duration, while associated psychoeducation materials provided participants with tools to improve such metrics. Essentially, this intervention has replaced live biofeedback with smartwatch-based health measures. Through creating a smartwatch intervention that targets improvement of physiological processes we aimed to increase generalizability and long-term accessibility of key biofeedback principles through translating them onto a smartwatch intervention design, extending applications to more diverse contexts and clinical populations.

Guiding MINC Intervention Principles

To create a smartwatch intervention for stress, key principles of the Minimal Intervention Needed for Change (MINC) approach were utilized given the intervention was designed to be light touch, involve limited clinician involvement, and be largely self-directed (Curran et al., 2021; Glasgow et al., 2014). Briefly, the MINC approach aims to minimize the mismatch between intervention research and clinical application by focusing on three key principles namely, reduction in cost and materials, the creation of a succinct intervention protocol that involves only a small number of highly effective strategies, and emphasis on patients evidencing meaningful change that positively impacts their daily functioning over large statistical significance (Curran et al., 2021; Glasgow et al., 2014).

Key principles of the MINC approach were integrated into our intervention through the use of consumer grade affordable technology, focus on teaching participants a handful of effective, easily

applied strategies to improve stress physiology and examining both statistical significance of the intervention alongside participant self-reported improvements.

Current Study

The present study piloted an 8-week self-guided smartwatch intervention that was designed to improve HRV and sleep through biofeedback inspired breathing and sleep hygiene techniques translated onto a wearable modality. Accessibility of the smartwatch interventions was considered by employing cost effective, publicly available technology (e.g. Fitbit watches) allowing participants to more readily apply HRV biofeedback skills to everyday stressors outside a laboratory setting. Additionally, utilizing a wearable technology modality that captures multiple physiological indicators allowed for the creation of a multimodal smartwatch intervention that incorporates physiological processes that are both directly (e.g., HRV) and indirectly (e.g., sleep) related to the human stress response. Through the use of this multimodal design, we effectively target both antecedents and consequences of the stress response.

An open pilot study design was utilized to examine the development of this novel 8-week intervention. Following best practices for pilot studies (see Abbot et al., 2014), acceptability outcomes were examined to determine if future efficacy trials are warranted. Here, acceptability was conceptualized as the degree to which participants found the intervention appropriate across a variety of domains (e.g., cognitive, emotional; Sekhon et al., 2017). Through examining acceptability of the novel, self-guided smart watch intervention for stress we aimed to understand participants' experience and use of the intervention in order to inform future trials. To examine acceptability, a modified theoretical framework of acceptability questionnaire (TFA; see Sekhon 2022, Appendix D) was administered to participants to assess their cognitive and emotional experiences related to participating in the intervention. The questionnaire wording was adapted to reflect the specific self-guided biofeedback intervention they completed. To supplement acceptability analysis as assessed through the modified TFA questionnaire, key feasibility metrics as utilized by previous pilot studies (see Abbot et al., 2014; Signore et al., 2022) were descriptively examined. Specifically, we reported on retention rate, and suitability assessment procedures.

Feasibility metrics were compared to pre-determined success criteria as outlined in previous pilot and protocol studies.

To examine preliminary efficacy of this intervention, we explored how engaging in the eight-week self-guided smart-watch intervention for stress modified participant physiological and psychological (e.g., perceived stress) responses to repeated administration of the Trier Social Stress Test (TSST), a robust laboratory protocol shown to reliably induce an acute stress response (Allen et al., 2014).

Research Hypotheses

To examine the acceptability and preliminary efficacy of the novel eight-week smartwatch intervention for stress, the present study examined four research hypotheses. Specifically, we proposed three hypotheses related to preliminary efficacy situated in the context of participant responsivity to repeated administration of the TSST and one hypothesis related to acceptability of the intervention.

Research Hypothesis 1: Acceptability and Feasibility

Given the pilot nature of this study acceptability and feasibility of the intervention was examined. Utilizing the TFA for healthcare intervention questionnaire, which was modified for the context of the present study, participant reported acceptability of the smartwatch intervention was investigated. Given previous research that demonstrates the feasibility of online interventions to support mental health and well-being amongst college students (e.g., González-García et al., 2021), we hypothesized that this intervention would be acceptable to our sample population. Furthermore, to supplement the acceptability analysis, we examined key feasibility metrics (e.g., participant retention rate as well as suitability of assessment procedures). Feasibility metrics were compared against pre-determined criteria for success. Specifically, participant retention was compared against mean attrition rates for smartwatch interventions (e.g., 73% retention; Triantafyllidis et al, 2024; Signore et al., 2022). Suitability of assessment procedures was assessed through number of adverse events reported (Sitka et al., 2026).

Preliminary Efficacy

Given previous literature that suggests both sleep and stress-reduction techniques modulate participant response to the TSST (e.g., Chen et al., 2016; Domes et al., 2019; Williams et al., 2013), we

hypothesize that participants randomized to the intervention condition will evidence an improved stress response to the second administration of the TSST when compared to the first administration. We propose three specific hypotheses related to preliminary efficacy which may serve as evidence of intervention effect in order to inform future efficacy trials.

Research Hypothesis 2: HRV Responsivity Across TSST Administration

We examined if there existed a condition-specific change in both resting HRV and HRV responsivity (e.g., reactivity and recovery) across the first and second administration of the TSST. While biofeedback research commonly examines change in average (e.g., tonic) HRV following intervention (Boesch et al., 2014; Schuman et al., 2019; Whited et al., 2014), here we examined both the presence of condition-specific change in HRV responsivity and resting average across time. The decision to examine both phasic and tonic measurements of HRV was informed by reactivity and recovery's specific and defined role in adaptive cardiac responses to stress, in addition to resonance breathing exercises well-defined ability to improve resting HRV (Boesch et al., 2014; Castaldo et al., 2015; Hamilton & Allyo, 2016; Held et al., 2021; Schuman et al., 2019; Whited et al., 2014).

Across hypothesis 2, we anticipated that the intervention condition would evidence altered HRV responsivity compared to their first administration of the TSST and that such change would not be evident in the control condition. We specifically postulated the finding of an *altered* rather than directional change in HRV during the TSST for several reasons. While there is evidence to suggest that biofeedback interventions can improve parasympathetic activation (Bornemann et al., 2019), directionality of change in HRV responsivity in the context of TSST is mixed in the literature (see Allen et al., 2014; Fellingner et al., 2025; Immanuel et al., 2023; Roos et al., 2017; Whited et al., 2014). Furthermore, there is growing evidence to suggest that vagal flexibility, that is an adaptive fluctuation in PNS engagement and withdrawal may represent a more adaptive profile to stressors opposed to static withdrawal or engagement (Giuliano et al., 2015; Muhtadie et al., 2015). Considering these findings, we postulated that the smartwatch intervention would enable participants to develop increased vagal flexibility, resulting in a change in cardiac profile during stressors that enables a more adaptive response to the TSST.

Furthermore, considering the relationship between heart rate and HRV, coupled with evidence to suggest that heart rate (HR) can be impacted by the TSST as an index of stress reactivity (Harvie et al., 2021), condition specific change in HR across time was also examined.

Research Hypothesis 3: Change in Negative Affect

We examined changes in psychological distress experienced during the first and second administration of the TSST. Specifically, we examined if there existed a condition specific improvement in negative affect across the first and second administration of the TSST. Given the literature to suggest that HRV is associated with emotional regulation processes (Appelhans et al., 2006; Geisler et al., 2010) we postulate that participants in the intervention group would evidence a statistically significant decrease in negative affect immediately following the second administration of the TSST compared to the first administration, whereas we predict that the control group will maintain similar elevations in negative affect across TSST administrations.

Research Hypothesis 4: Sleep Duration

We examined if there existed a condition specific improvement in sleep duration (e.g., total number of minutes asleep) following the intervention. While sleep is not directly activated during the physiological stress response, there is adequate literature to suggest sleep directly impacts HPA functioning, a key biological system implicated in stress reactivity (Balbo et al., 2010). As such, we conceptualized sleep as a tangential and related health metric that impacts the stress response, and examined the smartwatch intervention's ability to improve sleep. Given the literature to suggest that HRV biofeedback can improve sleep quality (e.g., Ebben et al., 2009) and that sleep quality can intern influence physiological stress responses (e.g., Williams et al., 2013) we hypothesized that participants in the intervention condition would show significant improvement in sleep duration across the 8 week intervention not evidenced in the control condition.

Methods

This pilot study employed a variety of physiological measurements and quantitative survey methodology to assess the acceptability and preliminary efficacy of an eight-week self-guided smartwatch

intervention on buffering physiological stress responses to the TSST. This pilot study also utilized a series of in-person laboratory visits where participants learned how to integrate the sleep and HRV biofeedback centred interventions into their everyday lives, completed electrocardiogram measures, and participated in the TSST laboratory stressor.

Participants

Undergraduate students were recruited from the human subject pool at the University of Manitoba between September 2024 and January 2026. Eligible participants were enrolled in Introductory to Psychology at the University of Manitoba at the time of their participation in the study. Additionally, participants were 18 years of age or older, had no history of adverse cardiac events or neurological conditions, were willing to wear a Fitbit watch for the duration of the intervention, comfortable reading and writing in English, and had no health condition that prevents them from engaging in physical activity. Inclusion criteria were defined to ensure participants were of sufficient age and health status to engage in an biofeedback based intervention strategies.

In total $N = 33$ participants (14 Control, 19 Intervention) enrolled in the study ($M_{age} = 20.70$, $SD_{age} = 2.492$). Seventy-five percent of participants who initial enrolled ($n = 25$; 11 control, 14 Intervention $M_{age} = 20.44$, $SD_{age} = 1.85$) completed all three laboratory visits, making them eligible for inclusion in analysis. Laboratory visits containing the TSST occurred between 9:00am and 2:00pm. While time of day (e.g., morning vs afternoon) has been demonstrated to impact cortisol (see Kudielka et al., 2004; Matsuda et al., 2012), many TSST administrations utilize a standardized window of test administration opposed to single administration time often ranging between 11am and 4pm (Allen et al., 2020; Drake et al., 2017).

Sample size varied across each individual hypothesis due to the longitudinal and multimodal nature of data collection (e.g., longitudinal sleep data, in person cardiac data). Furthermore, the permutation approach utilized to support small sample size (See below, Permutation Approach), requires complete case analysis (e.g., complete longitudinal data) in order to respect the data structure and satisfy exchangeability principles (Domhof et al., 2002). For a detailed breakdown of sample size across

analyses, see Table 1 and Figure 1. Given the high rate of attrition across those who enrolled versus those had useable data, an independent-samples t-test was conducted. There were no statistically significant differences across gender or age between those who were included and excluded in analyses ($p's > 0.05$). See Table 2 for self-reported demographic information.

Sample Size: Cardiac Analysis

Twenty-five participants (11 control, 14 Intervention) were eligible to be included in cardiac analysis. Upon cardiac data processing, $n = 11$ had completely unusable data due to invalid event codes or messy cardiac readings which prevented the study team from accurately identifying the start and stop time for various TSST components. Four participants had missing data from either the speech, math or recovery section. See Table 1 for full list of sample sizes across cardiac models and Table 3 and Figure 1 for exclusion criteria.

Table 1

Cardiac Model Sample Size

Cardiac Hypothesis	Sample Size
Model 1: Tonic HR	N = 14 (5 Intervention, 9 Control)
Model 2: Prep portion TSST	N = 14 (5 Intervention, 9 Control)
Model 3: Speech portion TSST	N = 13 (4 Intervention, 9 Control)
Model 4: Math portion TSST	N = 13 (5 Intervention, 8 Control)
Model 5: HR during TSST	N = 14 (5 Intervention, 9 Control)
Model 6: Recovery following TSST	N = 12 (5 Intervention, 7 Control)

Note. Table 1 describes total sample size included in statistical analysis across cardiac models

Sample Size: Sleep Analysis and Negative Affect

Twenty-five participants (11 Control, 14 Intervention) were eligible to be included in sleep analysis. Upon data analysis, $n = 8$ (4 Control, 4 Intervention) had useable sleep data. While previous studies have used an eighty-percent threshold to define useable physiological data (see Sitka et al., 2024; Wang et al., 2022), here we defined useable sleep data as individuals who had collected sleep for at least 4

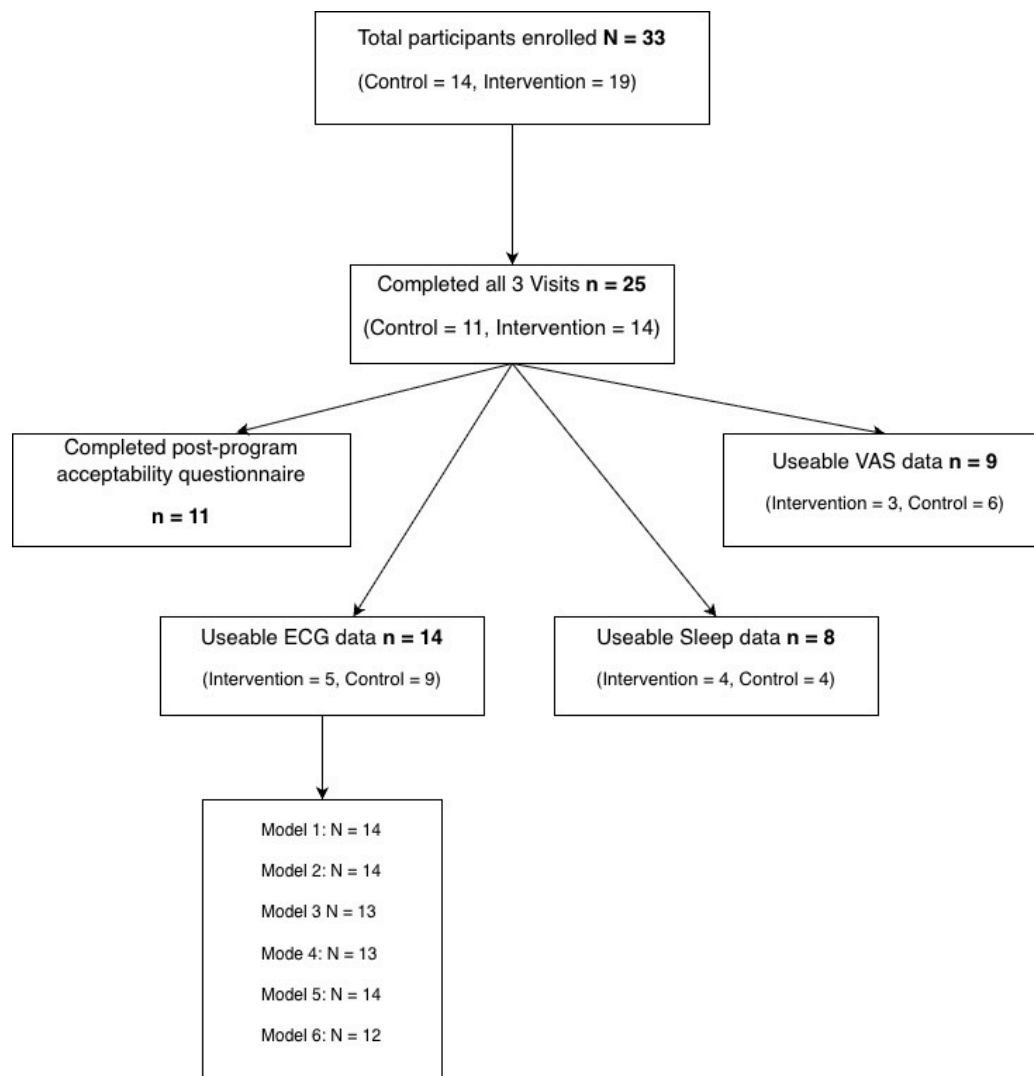
out of 7 nights per week across each of the eight weeks to satisfy complete data set requirement while maximizing sample size. The decision to deviate from the previously defined eighty-percent threshold was a post-hoc data driven decision based on data availability and quality. Similarly, $n = 9$ (3 intervention, 6 control) had useable affect data.

Sample Size: Feasibility Analysis

Participants self-reported measures of acceptability were included regardless of complete cardiac or sleep data so as long as they attended all three laboratory visits and were randomized to the intervention condition. As such all participants who were randomized to the intervention condition and completed the feasibility questionnaire were included, even if they had partial completion of the three laboratory visits (e.g., complete 1/3 or 2/3 visits; $n=11$)

Laboratory Visits

The first laboratory visit was 120 minutes in duration and occurred prior to the start of the eight-week intervention. During this first visit, all participants met with a trained research assistant to set up their Fitbit Inspire 3 watch, which they were instructed to wear for the duration of the intervention. Additionally, participants were fitted with electrodes for ECG measurement, then completed a standardized resting task, a series of pre-questionnaires, and the TSST. This first administration of TSST served as a baseline for participant stress reactivity. During the fourth week of the intervention, all participants were invited to attend a second, 60-minute laboratory visit in which they completed a second resting task with ECG measured throughout. Additionally, all participants received a 'Fitbit Refresher' where a trained research assistant reviewed best practices for wearing and charging the watch. Participants in the intervention condition also reviewed how to engage with the intervention materials.

Figure 1*Consort Flow Diagram for Inclusion and Exclusion*

Note. Figure 1 depicts enrollment and attrition across the eight-week longitudinal study. Useable data was defined as participant data that was complete across all laboratory visits and thus eligible for use in permutation based mixed models.

Table 2*Self-Reported Participant Demographics Total Eligible Sample*

Participant Characteristics	N
Ethnicity	
Arab	2
Black	4
Chinese	1
Filipino	7
Indigenous	2
Latin American	2
Southeast Asian	1
South Asian	2
White	6
Other	1
Sex	
Female	18
Male	6
Prefer not to answer	1

Note. N = 25; Eligible Sample is defined as the 25 participants who completed all three laboratory visits and were eligible to be included in analysis

After the eight-week intervention (e.g., week 9 or 10), participants attended the final 120-minute laboratory visit. During this final visit, participants were fitted with electrodes for ECG measurement, then completed the same resting task as the first visit, post-questionnaires, and completed a final TSST slightly modified to reduce habituation effects. See Table 5 for a full outline of laboratory visits and Appendix F for study timeline.

Intervention condition.

In addition to the procedures outlined above, participants in the intervention condition met with a trained research assistant during the first laboratory visit to receive the smartwatch intervention. Broadly, the intervention centred around taking key skills in biofeedback (e.g., resonance breathing, sleep tracking) and applying them to a smartwatch intervention modality in the absence of direct physiological feedback.

Table 3*Reason for Exclusion from Analysis*

Reason for Exclusion	N
<i>Eligible for cardiac, sleep and affect analysis</i>	25
Cardiac Data Exclusion	
Unusable Event Code	9
Poor data quality	2
Missing segment of TSST (math/speech/recovery)	3
Sleep Data Exclusion	
Did not collect any sleep data	13
Missing 1+ weeks of sleep data	4
Affect Analysis Exclusion	
Did not complete VAS scales at both lab visits	13
<i>Eligible for feasibility analysis</i>	14
Feasibility Analysis Exclusion	
Did not complete TFA Questionnaire	3

Note. N Total = 25 represents the total number of participants eligible for analysis prior to data cleaning and subsequent exclusion criteria were applied.

Smartwatch Intervention: HRV component The self-guided smartwatch intervention consisted of utilizing the RELAX breathing function once daily for five minutes. The RELAX function is a built-in application to Fitbit watches which provides a customized breathing exercise based on an individual's previously collected cardiac data. Research has demonstrated that using the RELAX breathing function once daily for five minutes can improve sleep, reduce anxiety and depression (Garcia et al., 2023). The rate of breathing suggested by the watch is similar to that of individualized breathing exercises used in traditional HRV biofeedback, and as such utilization of the RELAX function has the potential to modulate

participant cardiac responses to laboratory induced stressors by improving HRV functioning (Reiner, 2008; Steffen et al, 2017).

The design and implementation of the smartwatch intervention was informed by best practice guidelines across domains of smartwatch and biofeedback interventions and is consistent with the ‘individual HRV biofeedback’ design as described by Lanza et al. (2023) whereby participants engage in breathing exercises based on previously established resonance frequency rate determined by baseline cardiac data. The duration of the intervention (e.g., 8 weeks) was determined based on evidence suggesting most individual HRV biofeedback interventions are two to eight weeks in duration (Lanza et al., 2023). Because we were interested in examining the cumulative effect of smartwatch intervention on stress physiology (e.g., the repeated application of self-practice of HRV over time), the intervention length was maximized. While most individual HRV biofeedback interventions include one 20-minute laboratory session per week and minimal to no at home practice, our smartwatch intervention relied solely on independent at home practice with one laboratory refresher on week four (Lanza et al., 2023). In this sense, our intervention looked to extend the knowledge on individual HRV biofeedback interventions by focusing on self-guided model facilitated by wearable technology built in breathing function with the aim of increasing accessibility through reducing costs and limitations of laboratory-based interventions.

Smartwatch Intervention: Sleep Hygiene Component. A significant advantage of using consumer facing wearable technology within smartwatch interventions is that such devices capture multiple physiological measures. Given the literature describing the relationship amongst stress and sleep (Han et al., 2012; Lo Martire et al., 2020; Petersen et al., 2013; Verlander et al., 1999), coupled with HRV breathing exercise’s ability to improve sleep outcomes (Ebben et al., 2009), sleep hygiene psychoeducation materials were incorporated within this intervention design. Sleep centred psychoeducational intervention materials were designed following intervention best practices (see Hoffmann & Worrall, 2004). Following the key components of CBT-I (cognitive behavioural therapy for insomnia; Blake et al., 2018; Walker et al., 2022) the intervention taught participants how sleep hygiene, sleep restriction and mindfulness can improve sleep and wellness. Participants met with a trained research

assistant to discuss the importance of sleep on stress management and receive sleep hygiene psychoeducational handouts. Aligned with MINC approach, the sleep hygiene materials were designed to address four effective, straightforward strategies (e.g., sleep hygiene strategies, sleep restriction, mindfulness, stimulus control) that were explicitly linked to intervention outcomes quality (Curran et al., 2022).

Measures

A series of online self-report measures, gold standard laboratory measures, and wearable technology was used to assess participant physiological and psychological functioning throughout the intervention.

Pre-Questionnaires

During the first laboratory visit, participants completed a series of pre-questionnaires, however the present study will consider two questionnaires in the subsequent analysis. For a full list of questionnaires delivered across time points see Table 4. This will include an author generated sociodemographic questionnaire which asks participants to self-report on a variety of demographic variables (e.g., “What gender do you identify with?”).

Visual Analogue Scales

A visual analogue scale was used to assess participants psychological responses to the TSST. Following work by Hellhammer & Schubert (2011), participants were presented with a sliding scale ranging from 0 (*e.g., feeling not in control*) to 100 (*e.g., feeling very in control*) in which they were asked to mark their subjective appraisal of their current state of stress, anxiety, feeling of insecurity and control both immediately preceding and following the administration of the TSST. Participants completed this measure digitally within REDCap using a lab computer. Informed by work done by Harvie et al. (2021), only participant’s responses to the stress and anxiety measure were averaged to create on composite score representing their psychological response to the TSST.

Post Questionnaires

During the third lab visit, participants completed a series of post-questionnaires. The present study considered the acceptability questionnaire for the purpose of analysis. Following work by Sekhon and colleagues (2022), participants were asked 13 questions to assess the acceptability of the novel, multi-modal smartwatch intervention. Specifically, questions assessed participant affective attitudes, perceived burden, ethicality, effectiveness, coherence, self-efficacy and opportunity cost of the intervention. Questions to assess acceptability were modified and based upon Sekhon and colleagues (2022) Theoretical Framework of Acceptability Questionnaire. Questions were modified to better reflect the self-guided biofeedback intervention. See Appendix D for a full list of questions.

Weekly Questionnaires

Participants in the intervention condition were asked to complete an author generated weekly questionnaire that assess compliance and engagement with the intervention materials (e.g., Over the last 7 days, how often have you engaged in deep breathing and sleep hygiene techniques?). Participant response options range from less than once to nearly every day.

Table 4
Questionnaire Administration Timeline

Questionnaire	Timepoint		
	T1	T2	T3
Sociodemographic	X		
Generalized Anxiety Disorders Questionnaire	X		
Patient Health Questionnaire	X		
Theoretical Framework Acceptability Questionnaire			X
Weekly Questionnaire		X	

Note. T1 = pre-intervention lab visit; T2 = Weekly across the eight-week intervention; T3 = post-intervention lab visit

Physiological Measures

Participant physiology was captured using a series of ECG and Fitbit watch measurements. Participant cardiovascular recordings were captured using the Biopac MP160 system and Biopac Acknowledge software. This ECG assessment procedure involved placing 3 electrodes on participants' clavicle, abdomen and rib to assess general cardiac arousal as well as parasympathetic indices of cardiac activity. Specifically, we assessed participant's resting heart rate, resting HRV, and HRV responsivity through the duration of the visit.

Participant sleep duration, defined as the total number of minutes spent asleep, was assessed using the Fitbit Inspire 3 watches. Research suggests that Fitbit watches are able to capture participant sleep data with accuracy similar to that of gold-standard laboratory-based polysomnography measures (Lim et al., 2023). Participant watches were paired to Fitabase, an online data capture platform that allows extraction of participant Fitbit data. Specifically, the total time participants spend asleep in minutes per day was extracted for purpose of analysis.

Resting Task

The resting task is an eight-minute interval task designed to capture participant physiology at rest. Once participants were situated with ECG assessment equipment, they sat at a desk where they alternated in 1-minute periods with their eyes open or closed. Pre-recorded instructions were played for participants on speakers, and they followed the auditory instructions in tangent with visual instructions presented on the computer screen directly in front of them.

Trier Social Stress Test

The Trier Social Stress Test (TSST) has been reliably used to induce laboratory stress for several decades (see Allen et al., 2014) and involves giving a speech under false social evaluative conditions, directly followed by a challenging arithmetic problem. Participants completed the TSST twice, once before and once after the 8-week smartwatch intervention. Within the TSST protocol, there are two components of minor deception used to elicit physiological and psychological stress reactivity. First, participants are not told ahead of time that their speeches will be evaluated by a team of committee

members. Second, they are not told ahead of time that they will be completing the challenging arithmetic problem. Debriefing procedures occurred in person following the second administration of the TSST.

Following work by Boesch and colleagues (2014), the TSST protocol at post-testing was changed slightly so participants could not anticipate the design and structure of the assessment during the second administration. See Table 5 for a detailed description of the TSST administration.

Table 5

TSST Administration Timeline

TSST Component	TSST Time Point	
	Administration 1	Administration 2
Task Anticipation Period	Five minutes to prepare for speech and create notes	Five minutes to prepare for oral speech and create notes
Social Evaluative Committee	Trained research assistants watch participants deliver speech.	Trained research assistants watch participants deliver speech.
Speech Content	Five-minute speech to convince the committee they are the best candidate for a job	Five-minute speech to convince the committee they were wrongly accused of shop lifting
Math Task	Subtract 17 from 2043	Subtract 19 from 3035
Debriefing		Participants debriefed in person or via email

Note. TSST = Trier Social Stress Test. Administration 1 occurs during the first laboratory visit during week 0 of the study. Administration 2 occurs during the third laboratory visit during week 9/10 of the study.

Data Processing

Cardiac data (e.g., HR and HRV) collected during the TSST was processed by a team of trained research assistance using the Mindware IMP software. For a detailed description of cardiac processing procedures, (see Barkman et al., 2020). Briefly, data was distilled down into 30 second epochs for further processing and review of clear C waves and Q-waves before log transformation was applied to convert R-peak interbeat intervals into HRV values (Barkman et al., 2020). In total, 11 participants were excluded

during the data processing phase due to unusable data or unclear time codes that prevented researchers from accurately marking the different onset and cessation of TSST components. After data processing sample size was reduced (see Figure 1 for Consort flow diagram of participant enrollment).

Sleep data was extracted from each Fitbit watch using Fitabase (see Sitka et al., 2024). Sleep duration was extracted nightly for each participant and was defined as the total number of minutes spent asleep. For each participant, a person-centred average for sleep duration was calculated for each of the eight weeks of the intervention.

Analytic Plan: Hypothesis 1

To assess the first research hypothesis, which investigated the acceptability of this intervention, a series of descriptive and frequency statistics were used to investigate participant affective attitudes, perceived burden, ethicality, effectiveness, coherence, self-efficacy and opportunity cost of the intervention (Sekhon et al., 2017). Furthermore, information regarding recruitment, retention and feasibility assessment and protocol procedures is also descriptively provided to supplement understanding of protocol feasibility and acceptability. Within the supplementary feasibility analysis, we defined retention rate as the number of people who completed all three laboratory visits compared to those who dropped out or failed to book all three visits (Signore et al., 2022).

Analytic Plan: Hypothesis 2

To address the second research hypothesis, which assessed for the presence of condition specific improvement in both tonic (e.g., resting) HRV and HRV responsivity across administration of TSST, a permutation-based mixed-model analyses with random intercept was utilized. The first level of these models accounted for within person changes, specifically individual's physiological cardiac responses (e.g., Tonic HRV or HRV Responsivity) across repeated administration of the TSST (e.g., Administration 1, Administration 2). The second level of the model accounted for condition (e.g., control vs intervention). Given the small sample size, a decision was made to exclude a random slope term or additional covariates within the model due to increased likelihood of model instability:

$$HRV_{ti} = (\beta_o + u_{oi}) + \beta_1 Time_{ti} + \beta_2 Condition_i + \beta_3 (Time_{ti} \times Condition) \epsilon_{ti}$$

Where u_{0i} accounts individual random intercept

Where HRV_{ti} is the index of HRV (e.g., reactivity, recovery or resting state)

To examine the exploratory hypothesis which investigated condition specific changes in HR across time, a similar model was used to that of HRV. HR during the stress period was averaged (see Harvie et al., 2021).

$$HR_{ti} = (\beta_0 + u_{0i}) + \beta_1 Time_{ti} + \beta_2 Condition_i + \beta_3 (Time_{ti} \times Condition) \in_{ti}$$

Where u_{0i} accounts individual random intercept

While the measure of HRV is inherently itself a measure of variability in cardiac stress responses, here our models is equipped to assess the presence of directional mean level change in cardiac responses. As such we restrict interpretation of our results to the presence of a mean level change across TSST administrations opposed to the direct assessment of adaptive flexibility.

Calculation of Tonic HRV and HRV Responsivity

A singular HRV average was computed across each segment of the TSST (i.e., speech preparation, speech delivery, math task, and recovery; Hamidovic et al., 2020; Kudielka et al., 2004). Given evidence to suggest that the stressful nature of the TSST builds across tasks (e.g., Allen et al., 2014), HRV across all four periods were examined independently. For each segment of the stressor (e.g., prep, math and speech portion of TSST), reactivity was calculated by subtracting the average HRV during the stressor task from average HRV during the baseline period (Hamilton & Allyo, 2016; Knight et al., 2020). Here, a larger positive HRV reactivity score indicates PNS activation whereas a larger negative score indicates PNS withdrawal (Knight et al., 2020).

Following work by Knight et al. (2020), recovery scores were calculated by subtracting average task HRV from the recovery period immediately following task cessation. Because the TSST is designed to build across task segments, with the final math task involving both the element of uncontrollability and social evaluative threat (Allen et al., 2014), recovery scores were calculated by subtracting average HRV during recovery period from the average HRV during the math segment:

$$HRV_{Recovery} = (HRV_{Math\ Task}) - (HRV_{post})$$

Here, the HRV recovery period is defined as the first five-minutes following the cessation of the stressor. To examine condition by time improvement in resting state HRV, a composite HRV average score was computed for each participant following for the duration of the resting state task.

Analytic Plan: Hypothesis 3

To address the third research hypothesis, which investigated impact of condition on psychological affect during the stressor, permutation-based mixed-model analyses with random intercept was also utilized. The first level of these models accounted for within person changes, specifically individual's VAS score across repeated administration of the TSST (e.g., Administration 1, Administration 2). The second level of the model accounted for condition (e.g., control vs intervention).

To calculate the VAS score across repeated administration of the TSST, a singular composite negative affect score was calculated individually for the first and third laboratory visit. To create a composite score for each TSST administration the within visit pre-VAS was subtracted from the post-VAS score to assess the total change in affect observed after completing a singular administration of the TSST. More specifically, the VAS scores assessing level of anxiety and stress were administered before and after each administration of the TSST. Scores were averaged to create a within visit pre and post VAS score. Then, a change score calculated to represent difference in psychological distress (e.g., anxiety and stress) before and after completing a single administration of the TSST. A larger positive score indicated higher level of psychological distress. Following the creation of a composite score for each visit, condition by time improvement between change scores across first and second administration of the TSST was investigated using a mixed model analyses:

$$NA_{ti} = (\beta_0 + u_{0i}) + \beta_1 Time_{ti} + \beta_2 Condition_i + \beta_3 (Time_{ti} \times Condition_i) \in_{ti}$$

Where u_{0i} accounts individual random intercept

NA = Negative Affect

Analytic Plan: Hypothesis 4

To address the fourth research hypothesis, which investigated how participants sleep duration changed across the intervention, a permutation-based mixed-model analyses with random intercept was

utilized. Similarly, the first level of the model accounted for within person changes, defined here as within subject linear and quadratic growth in sleep duration across the eight-week intervention. Within the first level of the model, *Time* represented each of the eight weeks of the intervention whereas *Sleep* represented the participant centred average in sleep duration for each of the corresponding eight weeks of the intervention. The second level of the model accounted for condition (e.g., control vs intervention). Given the small sample size, a decision was made to exclude a random slope term with the model due to increased likelihood of model instability:

$$Sleep_{ti} = (\beta_o + u_{0i}) + \beta_1 Time_{ti} + \beta_2 Condition_i + \beta_3 (Time_{ti} \times Condition_i) + \epsilon_{ti}$$

Where u_{0i} accounts individual random intercept

Permutation Approach

To help account for the small sample size across models (e.g., $n < 20$) a permutation approach was utilized to support analysis. To preserve the longitudinal data structure, condition labels were shuffled ($n = 10,000$) across participants to build a null distribution to sufficiently examine condition effect on outcome variables. In short, permutation methodology may be described as method that “compares the observed value to a distribution generated by a group of permutations that would not affect the distribution if the null hypothesis were true” (Bůžková et al., 2011, p. 1). Given permutations tests observed suitability and strength in hypothesis testing (see Bonnini et al., 2024), permutation was chosen over bootstrap methodology. Using a permutation test allowed for the calculation of exact p values to better understand uncertainty around predictors values in the model.

It is important to acknowledge that while the permutation approach supports the robustness of our distributional assumption while mitigating type 1 error, it does not increase statistical power of our analysis. Permutation approach was chosen over bootstrap given its suitability to test hypotheses under the null, however as described in Abott et al. (2014), effect sizes are of particular usefulness in pilot studies as they can support future sample size planning and bootstrap methodology is better suited to produce such effect sizes and uncertainty intervals. Considering strengths and limitations of each approach, permutation methodology was selected given its robustness in hypothesis testing, however the

analysis was supplemented with semi-partial R^2 effect size calculations to better understand the unique variance in the model attributable to the interaction term of interest (Stofell et al., 2013).

Results

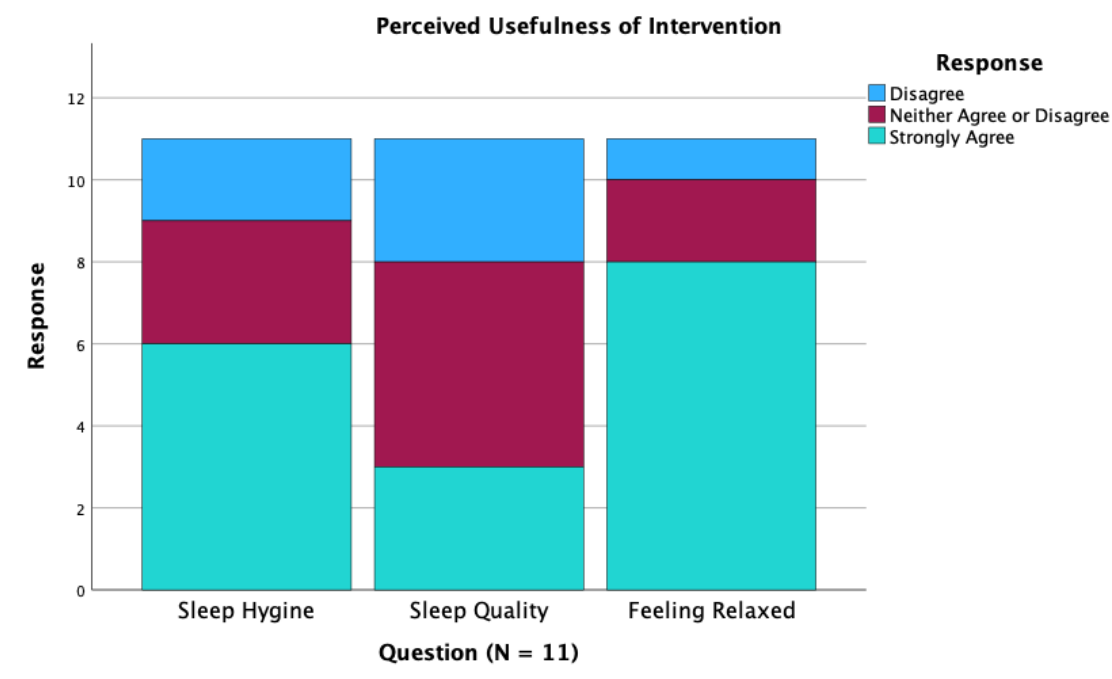
While the permutation approach does not assume normality of outcome, residual diagnostics were still utilized to examine if extreme violations of normality were present within the data set. Gross violation of normality was absent within these residual plots. See Appendix C for residual graphs and diagnostic statistics.

Research Hypothesis 1: Acceptability and Feasibility

To assess acceptability, a series of descriptive statistics were used to investigate participants responses to the Theoretical Framework of Acceptability questionnaire which was modified to assess intervention specific components of the smartwatch intervention. The questionnaire assesses seven constructs of acceptability namely, affective attitude, burned, ethicality, perceived effectiveness, intervention coherence, self-efficacy, opportunity costs and general acceptability. In total, the survey was administered to all $n = 11$ participants who were randomized to the intervention condition and completed the intervention's third laboratory visit and associated acceptability questionnaire. Participants across all four cohorts were included (See Figure 1, consort diagram). Based on participant responses, the intervention was acceptable with 81% of participants reporting the intervention was completely acceptable. See table 6.

Furthermore, 54% of participants self-reported an improvement in their sleep hygiene knowledge, 27% reported improved sleep quality and 72% reported increased feeling of relaxation over the course of the eight-week intervention see Figure 2.

Figure 2
Self-Reported Perceived Efficacy of Intervention



Note. Figure 2 depicts participant reported perceived efficacy of the intervention as it related to improvement in sleep hygiene knowledge, perceived sleep quality and overall feeling of relaxedness

Table 6

Acceptability Analysis

TFA Construct	Question	Participant Responses
Affective Attitude	Did you like or dislike this biofeedback intervention	Like (n=9) Neutral (n=2)
	How comfortable did you feel engaging in the biofeedback intervention	Very comfortable (n=4) Comfortable (n=5) Neutral (n=2)
Burden	How much effort did it take to engage in the biofeedback intervention	A little effort (n=10) A lot of effort (n=1)
Ethicality	There are moral or ethical consequences to engaging in the intervention	Strongly Disagree (n=2) Disagree (n=5) Neutral (n=1) Agree (n=3)
Intervention coherence	It is clear to me how integrating sleep hygiene and deep breathing exercises will improve my health, sleep and well-being	Strongly Agree (n=3) Agree (n=5) Neutral (n=1) Disagree (n=2)
Self-efficacy	How confident did you feel engaging in the sleep hygiene and tracking exercises	Very Confident (n=2) Confident (n=6) Neutral (n=2) Unconfident (n=1)
	How confident did you feel about engaging in deep breathing exercises?	Very Confident (n=2) Confident (n=6) Neutral (n=2) Unconfident (n=1)
Opportunity costs	Engaging in the biofeedback program interfered with my other priorities	Strongly Disagree (n = 4) Disagree (n=6) Neutral (n=1)
General Acceptability	How acceptable was the intervention to you?	Completely Acceptable (n=5) Acceptable (n=4) Neutral (n=2)

Note. N=11

To assess feasibility, participant retention and recruitment as well as protocol success were compared to pre-determined criteria based on similar pilot and smartwatch intervention studies. Retention rate was successfully with a final rate of 76% (e.g., 25 of total 33 participants completed all three visits), which is above the average rate observed in previous smartwatch interventions (e.g., 73% retention; Triantafyllidis et al, 2024). Furthermore, no adverse events were reported within this protocol providing additional evidence of feasibility.

Research Hypothesis 2: Heart Rate Variability, Recovery and Average HR

Preliminary efficacy of the novel intervention was assessed by examining if there existed a condition specific directional mean level change in both tonic HRV and phasic HRV responsivity across the first and second administration of the TSST. A permutation-based mix model analysis (n=10,000 permutation) did not provide evidence of a condition-specific directional mean-level change in HRV responsivity across each phase of the stressor (e.g., prep, speech, math task) or during recovery (p 's > 0.05, See Table 10, Models 2-4, 6 and Figures 4-6,8). Furthermore, the semi-partial R^2 values across models indicated that the observed condition-specific change was very small to small. See Table 7, 8 for descriptive mean level cardiac values by condition during the TSST segments across time one and two.

When running the permutation based mixed model for changes in tonic HRV, the model returned a singular fit with random intercept variance estimated at zero. The singular fit demonstrates insufficient information to distinguish residual variance and between person variability; however, the random intercept was retained to respect repeated measures design and dependence structure of data. In total, 5220 (e.g, 52%) of the permuted fits returned a singular value. Fixed effects within this model did not provide evidence of a condition-specific directional mean-level change in tonic HRV, and semi-partial R^2 indicated that the observed condition specific change was very small ($p = 0.85$, $R^2 < .001$; see Table 10 Model 1 and Figure 3). Given the singularity of fit, generalizability of results is cautioned, and this may be seen as a limitation to current sample size rather than the model fit itself. The semi-partial R^2 values for this model (e.g., Table 10 Model 1) should also be interpreted with additional caution. Furthermore, there was evidence of a condition-specific directional mean-level change in average HR during the entire

stressor task ($p = 0.41$; $R^2 = 0.02$; See Table 10, Model 5 and Figure 7). For individual trajectory plots, see Appendix B.

Table 7

Mean Cardiac and Affect Values Across Lab Visits: Control Condition

TSST Segment	Time 1	Time 2
HRV _{Baseline}	5.99 (0.79)	6.16 (0.67)
HRV _{prep}	5.19 (1.26)	5.61 (1.43)
HRV _{speech}	5.51 (0.90)	5.39 (0.98)
HRV _{math}	5.28 (0.92)	5.66 (0.68)
HRV _{recovery}	5.70 (0.58)	5.89 (0.61)
HR	92.56 (20.11)	90.68 (18.21)
Negative Affect	42.33 (23.45)	17.50 (23.44)

Note. SD appears in brackets; HRV baseline, prep, speech, math and recovery represents the mean average HRV value during that segment of the TSST; HR values represent the mean HR value across the entire duration of the TSST; Negative affect values represent the average change score assessed by the VAS scale

Table 8

Mean Cardiac and Affect Values Across Lab Visits: Intervention Condition

Variable	Time 1	Time 2
HRV _{Baseline}	6.34 (0.52)	6.62 (0.24)
HRV _{prep}	5.43 (1.05)	5.57 (1.48)
HRV _{speech}	6.21 (0.79)	6.51 (1.06)
HRV _{math}	6.33 (0.57)	6.76 (0.58)
HRV _{recovery}	6.11 (0.73)	6.55 (0.54)
HR	85.59 (8.17)	88.31(9.41)
Negative Affect	27.50 (24.04)	22.50 (25.84)

Note. SD appears in brackets; HRV baseline, prep, speech, math and recovery represents the mean average HRV value during that segment of the TSST; HR values represent the mean HR value across the entire duration of the TSST; Negative affect values represent the average change score assessed by the VAS scale

Table 9*Deceptive Statistics for Total Sleep Duration by Condition and Week*

Intervention Week	Sleep Duration in Minutes	
	Control	Intervention
1	476.04 (71.67)	447.03 (70.31)
2	480.30 (127.53)	412.94 (56.59)
3	485.99 (92.50)	417.98 (63.52)
4	429.94 (73.20)	461.54 (83.59)
5	491.50 (85.40)	393.166 (72.88)
6	459.88 (135.37)	409.37 (66.71)
7	466.62 (127.50)	409.37 (66.71)
8	484.74 (87.49)	420.99 (59.57)

Note. SD appears in brackets; Sleep duration represents the average time spent asleep in minutes for the corresponding week of the intervention

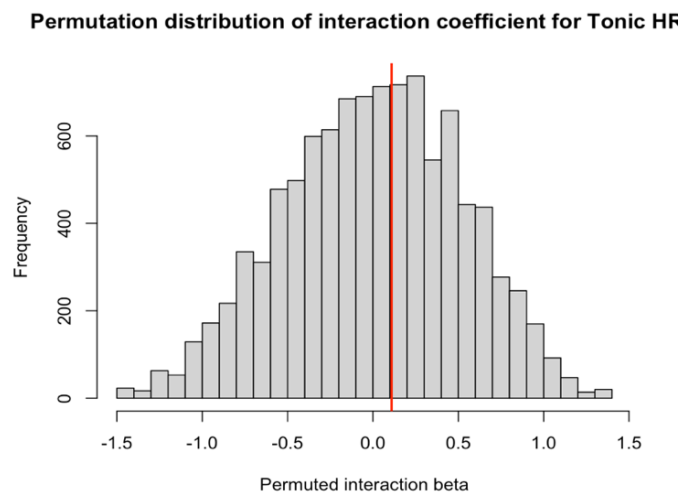
Table 10*Permutated Beta Interaction Coefficients: Cardiac Models*

Model	Observed β	Permutated p value	Interaction semi-partial R^2
Model 1: Tonic HRV	0.11	0.85	<.001
Model 2: HRV during Prep	0.18	0.80	<.001
Model 3: HRV during Speech	-0.14	0.79	.044
Model 4: HRV during Math	0.17	0.76	.003
Model 5: HR during stressor	3.59	0.41	.020
Model 6: HRV Recovery	0.20	0.49	<.001

Note. N = 10,000 permutations; Permutated p value can be conceptualized as the likelihood of seeing a test statistic of equal or greater value under the null hypothesis; Semi-partial R^2 represents the unique variance associated with the Condition X Timepoint interaction after accounting for the remaining fixed effects. Estimates are descriptive and reported without confidence intervals because of the small analytic samples and model instability

Figure 3

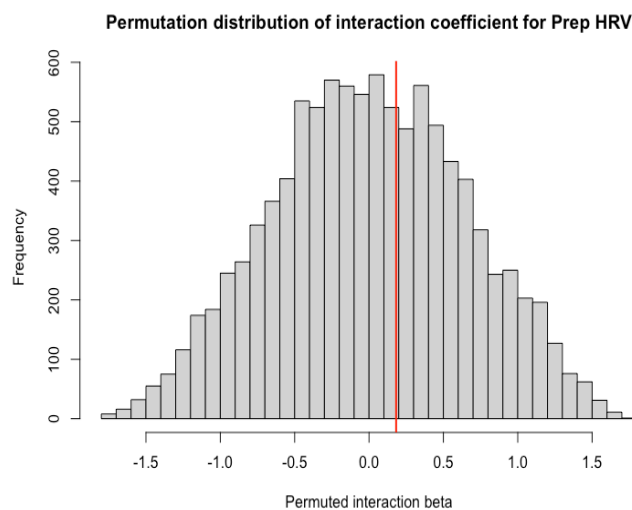
Permuted Distribution of Interaction Coefficient β for Model 1: HRV at Rest



Note. Figure 3 represents the permuted null distribution for Model 1, Tonic HRV. The red line indicates where the observed test-statistics fall relative to the permuted distribution.

Figure 4

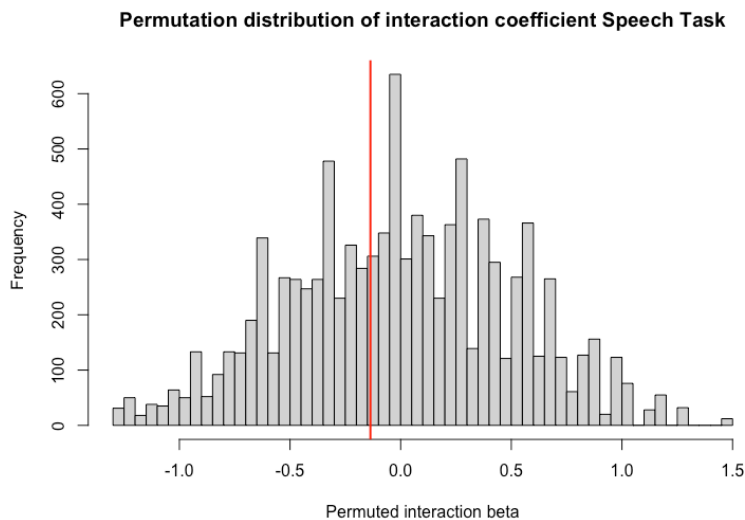
Permuted Distribution of Interaction Coefficient β for Model 2: HRV During “Prep” Period of TSST



Note. Figure 4 represents the permuted null distribution for Model 2, HRV reactivity during the “prep” phase of the TSST. The red line indicates where the observed test-statistics fall relative to the permuted distribution.

Figure 5

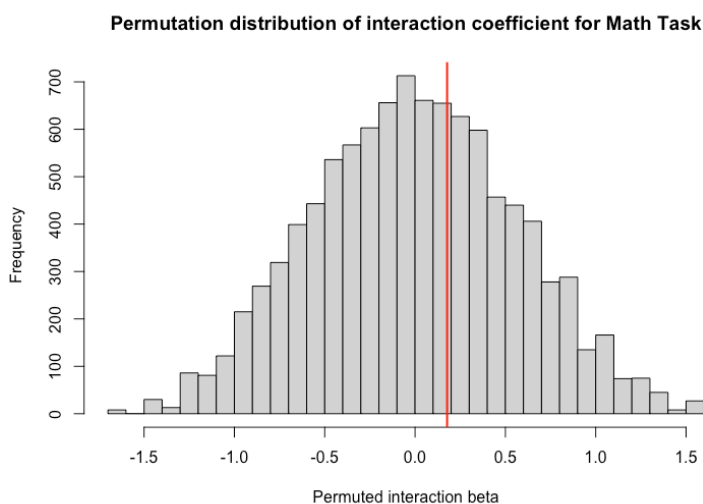
Permuted Distribution of Interaction Coefficient β for Model 3: HRV During Speech Task



Note. Figure 5 represents the permuted null distribution for Model 3, HRV reactivity during the “speech” phase of the TSST. The red line indicates where the observed test-statistics fall relative to the permuted distribution.

Figure 6

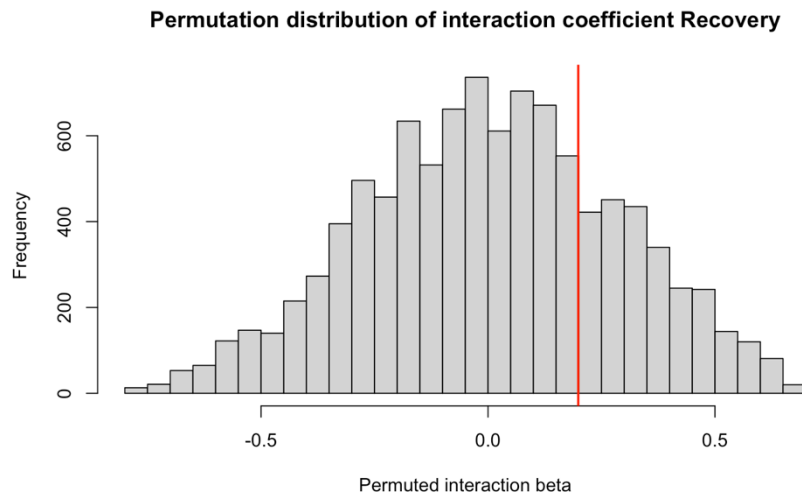
Permuted Distribution of Interaction Coefficient β for Model 4: HRV During Math Task



Note. Figure 6 represents the permuted null distribution for Model 4, HRV reactivity during the “math” phase of the TSST. The red line indicates where the observed test-statistics fall relative to the permuted distribution.

Figure 7

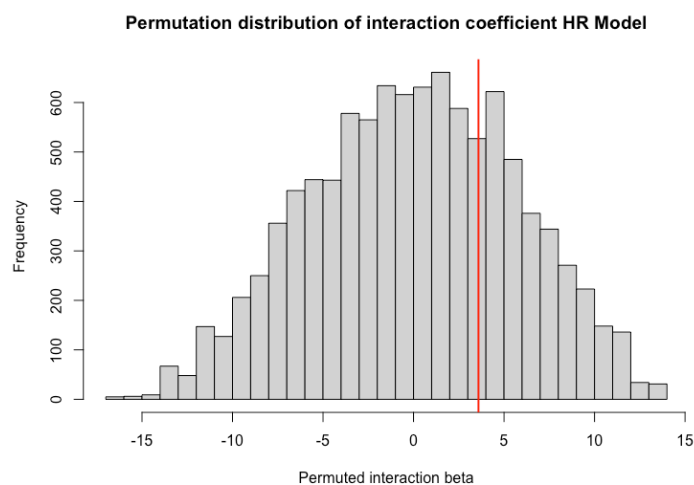
Permuted Distribution of Interaction Coefficient β for Model 5: HR During Stressor Duration



Note. Figure 7 represents the permuted null distribution for Model 5, HRV reactivity during the “math” phase of the TSST. The red line indicates where the observed test-statistics fall relative to the permuted distribution.

Figure 8

Permuted Distribution of Interaction Coefficient β for Model 6: HRV Recovery



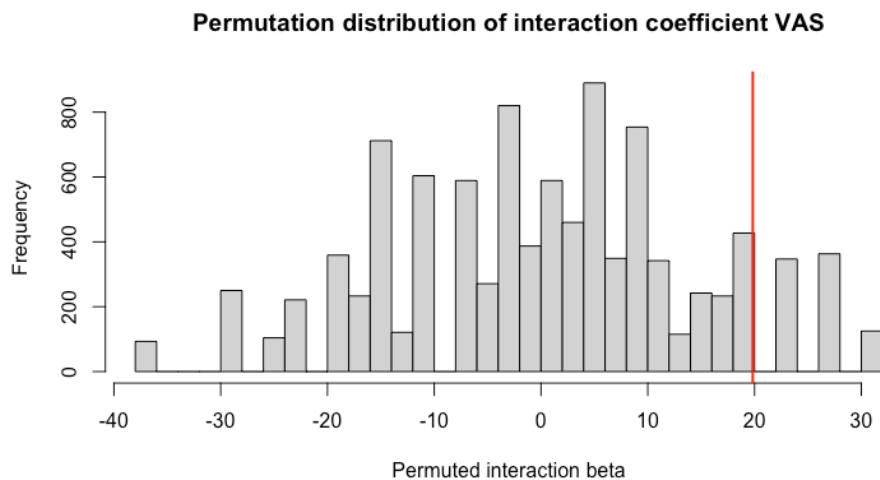
Note. Figure 8 represents the permuted null distribution for Model 6, HRV recovery following the TSST. The red line indicates where the observed test-statistics fall relative to the permuted distribution.

Research Hypothesis 3: Psychological Response to TSST

A permutation-based mix model analysis ($n=10,000$ permutation) revealed no significant interaction effect of Time by Condition on negative affect $\beta_3 = 19.83, p = 0.17$. The interaction accounted for little unique variance in the model ($R^2 = 0.04$; See Figure 9, Table 11, Model 7). See Table 7, 8 for mean change in negative affect during the TSST across time one and two.

Figure 9

Permuted Distribution of Interaction Coefficient β for Negative Affect



Note. Figure 9 represents the permuted null distribution for Model 7, negative affect following administration of the TSST. The red line indicates where the observed test-statistics fall relative to the permuted distribution.

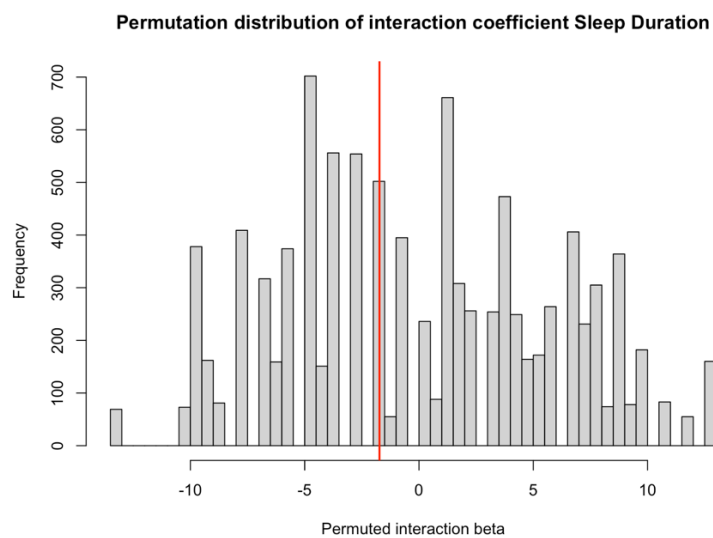
Research Hypothesis 4: Change in Sleep

A permutation-based mix model analysis ($n=10,000$ permutation) revealed no significant interaction effect of Time by Condition on Sleep across the eight-week intervention $\beta_3 = -3.36, p = 0.67$. The interaction accounted for very little unique variance in the model ($R^2 < .001$; See Figure 10, Table 11, Model 8). See Table 9 for descriptive statistics for sleep duration across conditions.

Table 11*Permutated Beta Interaction Coefficients: Affect and Sleep Models*

Model	Observed β	Permutated p value	Interaction semi-partial R^2
Model 7: Negative Affect	19.83	0.17	0.04
Model 8: Sleep Duration	-3.36	0.67	<.001

Note. N = 10,000 permutations; Permutated p value can be conceptualized as the likelihood of seeing a test statistic of equal or greater value under the null hypothesis; Semi-partial R^2 represents the unique variance associated with the Condition X Time-point interaction after accounting for the remaining fixed effects. Estimates are descriptive and reported without confidence intervals because of the small analytic samples and model instability

Figure 10*Permutated Distribution of Interaction Coefficient β for Sleep Duration*

Note. Figure 10 represents the permutated null distribution for Model 8, which examined sleep duration. The red line indicates where the observed test-statistics fall relative to the permutated distribution.

Discussion

Using permutation-based mixed-model analyses, the present study assessed the preliminary efficacy of the novel smartwatch intervention designed to improve stress psychophysiology by examining condition specific improvements in psychophysiology (e.g., sleep, HRV, HR, affect) across repeated administration of the TSST. Cardiac measurements were assessed using lab-based electrocardiogram measures and sleep data was collected using Fitbit Inspire 3 watches and extracted as total sleep duration in minutes through Fitabase. Intervention materials utilized wearable technology (e.g., Fitbit watches), psychoeducational materials and key strategies found in biofeedback (e.g., resonance breathing, tracking physiological measures to improve outcomes). No direct, live clinician biofeedback was administered to participants. Contrary to my initial hypothesis, there was no evidence of a condition specific mean directional change in cardiac measures (e.g., HRV responsivity, resting HRV, HR) during the second administration of the TSST compared to the first. There was also no evidence for condition specific improvements in sleep across the eight-week period or change in psychological responses to the stressor (e.g., negative affect following the TSST). Present findings contradict previous research demonstrating the efficacy of biofeedback related intervention strategies and smartwatch interventions (Dillon et al., 2016; Prinsloo et al., 2013; Thabrew et al., 2018). While several factors may have contributed to null findings, the small sample sized limited our power to adequately detect small and medium effect sizes between groups and as such we position this as an important limitation that may have contributed to such findings in addition to elements discussed further below.

First, we may look to the length of the smartwatch intervention as an explanation for minimal change across sleep and cardiac physiology. While duration of the present intervention was informed based on traditional duration of in-lab biofeedback (see Lalanza et al., 2023), the eight-week period may not have been long enough to account for the self-guided nature of the intervention. While interventions designed to modify HRV in diverse samples commonly range from two to six weeks (Chen et al., 2022), they have direct and frequent clinician involvement (e.g., 3-4 sessions per week with a clinician, real time clinician feedback) and utilize lab technology and real time feedback opposed to smartwatch-based data.

Furthermore CBT-I interventions, which are designed to improve sleep duration and quality, are commonly delivered via three to four sessions a week across a three-month period (e.g., 12 weeks). While single session CBT-I programs exist, long term efficacy is understudied, with research often reliant on short term, self-report measures (see Okun et al., 2023). Extending the protocol may allow individuals more time to become acquainted with the skills, wearable technology, and subsequent application of techniques in their daily lives in the absence of direct clinical involvement allowing for more pronounced physiological effects to be present. While extending the protocol may create more time to habituate to the integration of skill into participants lives, our present study was also limited by the academic semester and bound by student schedules making extension of the intervention not feasible within the current sample. As such, future work may look to not only extend the intervention length but also select differing populations that do not inherently present with time constraints.

Second, the null effect of condition on physiology may elucidate a need to re-evaluate the light-touch, self-guided nature of the smartwatch intervention. While results of the present study conflict with previous research indicating effectiveness of biofeedback strategies on HRV and sleep improvement (De Witte et al., 2019; Yu et al., 2018), we postulate the self-guided nature and deliverance method may account for such discrepant findings. The present study differs from traditional biofeedback interventions across two critical domains namely, the use of wearable technology opposed to real time laboratory feedback and the single session intervention design opposed to multiple in person sessions. In essence, the present study aimed to change the modality in which biofeedback is delivered onto wearable technology while simultaneously reducing clinician involvement in an effort to increase the accessibility of the intervention. The smartwatch intervention aimed to teach key strategies found in biofeedback (e.g., resonance breathing, sleep tracking), without real time feedback from clinicians. When distilled down to its core, this intervention may be conceptualized as a biofeedback informed smartwatch psychoeducation program in which efficacy was assessed through longitudinal change in physiology.

While such aims were informed through a MINC lens, which aims to increase the intervention research application in real world clinical settings through reducing cost and complexity of the intervention (see Curran et al., 2022), the significant change across modality, simultaneous reduction in clinician involvement and removal of real time physiological feedback may represent too stark of a transition to a light-touch design, contributing to such insignificant findings. While light-touch interventions designs can hold significant promise, future studies may look to change one key factor (e.g., deliverance modality OR clinician involvement). Furthermore, a scaffolded approach which involves a stepwise reduction in direct clinical involvement across iterative trials may allow more accurate understanding of the degree to which clinical involvement contributes to the positive changes in physiology evidenced across previous biofeedback interventions (e.g., Dillion et al., 2016; De Witte et al., 2019).

Change in psychophysiological stress responses were assessed by examining changes in self-reported anxiety and stress across the first and second administration of the TSST. In contrast to initial hypothesis, no change in affect was observed across time by condition. Given the vast literature to suggest a relationship between physiological status (e.g., perceived stress, anxiety, depression) and HRV (Dishman et al., 2000; Nako et al., 2018; Martinez et al., 2022; Silva et al., 2015; Pulpulos et al., 2018) the null change in HRV observed across trials may explain the insignificant change in physiological affect across TSST administrations.

Finally, the null findings related to cardiac hypotheses may in part be contributed to by choice of model. While the multi-level models utilized focused on mean level directional change, utilizing more advanced models that explicitly elucidate the presence of vagal flexibility (e.g., intra-individual standard deviation in HRV across task segments; Spangler et al., 2018) may elucidate a person-centred change in cardiac flexibility. Furthermore, the individual trajectory plots (see Appendix B) indicate within-condition heterogeneity. As such, future studies that yield larger sample sizes may wish to utilize more person-

centered approach (e.g., latent class growth analysis, k-mean clustering) in order to explore such heterogeneity and associated intervention implications further.

Acceptability of the intervention was assessed using an online questionnaire (see Sekhon et al., 2022; 2017). Aligned with my initial hypothesis, results demonstrate that participants viewed the intervention as feasible, enjoyable, and personally beneficial. Findings align with previous research demonstrating the acceptance and feasibility of digital health interventions, particularly for college aged populations (González-García et al., 2021; Kanuri et al., 2020). Additionally, aligned with our hypothesis, participants self-reported a perceived improvement in their sleep hygiene knowledge (e.g., factors that contribute to improved sleep quality) as well as sleep quality itself. Such results are aligned with previous research that demonstrates effectiveness of sleep hygiene strategies utilized within the intervention at improving sleep quality, and overall knowledge of sleep hygiene (Blake et al., 2018; Walker et al., 2022). Furthermore, self-reported improvement in sleep duration aligns with previous smartwatch interventions findings whereby 30% of individuals also reported improvement in sleep metrics (see Kaplan et al., 2023).

Following key feasibility studies (see Abbot et al., 2014; Signore et al., 2022), feasibility metrics were also reported including, participant retention and adverse effects. As expected, no adverse effects were reported within the protocol. Retention rate exceeded pre-determined criteria of 73%, the average retention rate in smartwatch interventions, further providing evidence for feasibility of protocol and intervention (Triantafyllidis et al, 2024). However, we acknowledge that significant data loss occurred due to insufficient lab techniques or challenges event coding during the TSST (e.g., the computer system designed to time mark the onset and cessation of TSST elements did not function properly during the assessment). As such, while retention rate exceeds predetermined criteria, refinement of internal lab procedures is warranted to support greater data attainment prior to future iterations or trials. Simplifying the stressor task used, such as the Montreal Imaging Stress Task (see Grodin et al., 2024) may be a way to mitigate data loss in future iterations.

Strengths and Limitations

There are several notable strengths in the present work, including the use of gold-standard laboratory measurements to assess physiology, collection of objective sleep data, and the use of the TSST, a globally validated stress paradigm.

The present studies use of ECG measurements involving three electrodes across the clavicle, abdomen and rib allowed us to accurately capture changes in high frequency heart rate variability, a more insightful measurement of parasympathetic cardiac activity compared to low frequency or average HR measurements. The use of gold-standard procedures to quantify HRV during lab visits is a strength of the work and instills confidence in conclusions drawn and data quality.

The use of Fitbit watches to objectively assess changes in participant sleep data across the intervention period is another notable strength of the present work for two reasons. First, there is a well-documented discrepancy between objective and subjective reports of sleep with only a moderate correlation between these two measures ($r = .45$; Lauderdale et al., 2008a; Lauderdale et al., 2008b) and as such the use of objective data opposed to reliance on self-report may be considered a strength. Second, the modality used to collect sleep data (e.g., Fitbit watches) allows participants to remain in their naturalistic environment while data is captured, allowing examination of how the self-guided intervention improves daily health metrics (e.g., total minutes of sleep) outside of the laboratory context. Finally, the use of repeated measures TSST stress paradigm to explore the preliminary efficacy of this self-guided biofeedback intervention may also be seen as a strength of this work given its robust design, cross-cultural use and reliability in inducing a physiological stress response (Allen et al., 2014; Miller & Kirschbaum, 2019).

While there are many notable strengths to this work there are also several limitations that must be addressed namely, the influence of moderating variables on TSST performance and the small sample size. The present study examined the influence of a smartwatch intervention on stress reactivity in the context of the TSST. Given the complex physiological and psychological systems involved in the human stress response, there exists several potential moderating variables that may influence TSST reactivity. Key

moderating variables that have been previously examined in the literature include the role of biological sex, menstrual cycles, and chronic historic stressors on TSST performance; however, the present study did not measure or control for each of these variables.

To date, research investigating the role of biological sex on TSST reactivity has yielded mixed results. Evidence suggests that gender differences are most pronounced in measurements of cortisol and psychological mood following the TSST, with men evidencing greater cortisol response and women evidencing greater negative mood (Allen et al., 2014; Eisenberger et al., 2007; Kelly et al., 2008; Uhart et al., 2006). Research has also demonstrated a significant gender by time interaction, whereby woman evidenced higher resting heart rate and faster recovery in response to the TSST (Kudielka et al., 2004). However, there is also evidence to suggest that age rather than biological sex may better moderate resting HR (Kudielka et al., 2004). In addition to biological sex, related research has examined the role of menstrual cycles and oral contraceptives on TSST stress responses. Findings suggest that in younger adult samples the impact of sex hormones may not be as pronounced (Allen et al., 2014; Lennartsson et al., 2012). Research also suggests that the stage of menstrual cycle and the use of oral contraceptives more directly influence cortisol opposed to cardiac indices of stress (Liu et al., 2017).

Given the novelty of the proposed intervention coupled with the pilot study design this study recruited both biologically male and female participants. Future studies may consider recruiting homogenous sample of either males or females in order to more stringently assess sex differences in outcome variables. Furthermore, the present study did not assess or control for the use of oral contraceptives or stage of menstrual cycle. While a potential limitation, the young age of the sample coupled with the use of cardiac opposed to salivary based measurements supports our inclusive sample composition (Allen et al., 2014; Eisenberger et al., 2007; Kelly et al., 2008; Kudielka et al., 2004; Liu et al., 2017; Uhart et al., 2006). Future efficacy trials should control for stage of menstrual cycle at time of repeated TSST measurements as well as use and type of oral contraceptive.

Time of day has also been considered a moderating variable that impacts stress response, with effects most pronounced on measures of cortisol (Kudielka et al., 2004). While the present study utilized

a general time window in which TSST was administered (see Allen et al., 2020; Drake et al., 2017; Kudielka et al., 2004; Dunn & Taylor, 2014), future work may look to more explicitly control for and track time of TSST completion within analytic models.

In addition to potential biological and circadian moderators, research has considered the influence of chronic stressors and lifetime history of stress on physiological reactivity to the TSST. Similar to the influence of oral contraceptives, lifetime history of stress (e.g., childhood maltreatment, recent history of trauma) predominately influences cortisol and immune responses to the TSST, opposed to cardiac reactivity, which is the primary physiological variable of interest (Allen et al., 2024; Carpenter et al., 2010). Given the aforementioned evidence, the limited consideration of previous life stress may be seen as a limitation of the current work. This may be particularly relevant for the sample of undergraduate students participating in this study, as the intervention took place over the course of an academic semester, with the last laboratory visit occurring in proximity to final exams. Future studies may consider including a more comprehensive and extensive clinical history or interview in order to better assess and control for chronic stressors in the lives of study participants.

Sample size and associated data exclusion criteria remained a key limitation throughout this work. The present study remains underpowered despite the inferential statistical framework used. Our sample ($n = 14$) is not sufficiently powered to detect moderate or small effects, and as such future studies must look to increase sample size to provide more comprehensive information on the efficacy of this intervention. The post-hoc decision to lower the previously standard of eighty-percent data requirement for inclusion in sleep analysis further reduced sample size and may be seen as an additional limitation. Moreover, cardiac exclusions also reduced sample size and such exclusion were not random, but rather driven by data availability and quality. While an independent samples t-test revealed no sociodemographic differences for those who were included and excluded, there remains the possibility that participants were included or excluded systematically based on physiological profiles as this was not explicitly examined. Future studies may look to compare baseline HRV for those who are and are not

included in analysis to more transparently examine if there exists a difference in physiology amongst groups.

While the smartwatch intervention itself was intentionally designed not to be time intensive (e.g., self-guided, light-touch design, use of consumer grade technology) the nature of repeated laboratory and TSST measurements may have been burdensome and contributed to a low sample size for several reasons. Participants were asked to complete three in person laboratory visits across an academic semester that ranged from 90 to 120 minutes in duration and involved ECG and EEG measurements. The time intensive nature and more invasive physiological measurements used (e.g., 3 electrodes on clavicle, abdomen and rib) likely contributed to attrition rates resulting in lower sample size. Furthermore, the longitudinal design was resource intensive, requiring several study personnel to be present in order to run one laboratory visit, ultimately limiting the study team's capacity to recruit larger samples.

The nature of the chosen population and compensatory structure of this study also contributed to attrition challenges. Students were compensated in course credits for completing various stages of the intervention. Given considerations from our research ethics board, students did not have to commit to the full study trial at time of visit one. For example, students need a set number of course credits and as such someone may achieve the required number of credits part way through the study and subsequently disengage or unenroll. This contributed significantly to longitudinal attrition and may have resulted in limited engagement with study materials once students had reached their required course credits. The use of a permutation analytic method was intentionally chosen to acknowledge the small sample size and mitigate its subsequent effect on analysis, however it is imperative that future work increase sample size in order to be powered to detect small and moderate effect sizes.

Future studies may look to utilize a number of strategies to mitigate attrition and improve total sample size. First, future work may look to provide monetary compensation, opposed to course credit, in order to better incentivise participation and more accurately compensate participants for their time. Additionally, extending recruitment to community samples may also increase total sample size. Finally,

future work might consider the use of alternative stress paradigms other than the TSST that are less research personnel intensive, such as the Montreal Imaging Stress Task (see Grodin et al., 2024).

Finally, while the present study utilized semi-partial R^2 as a descriptive estimate of effect size, we position the findings of small effect sizes as preliminary given the small analytic samples and evidence of model instability in some analyses. Confidence intervals were not calculated which further restricts interpretation of such values. Future fully powered studies should report both effect sizes alongside appropriate uncertainty intervals in order to inform sample size.

Data Loss Implications and Future Directions

The present study encountered significant data attrition which held overarching implications across variables, analysis, implications and future directions. While efforts were made to conduct appropriate statistical analysis given the sample composition, the data attrition limited our ability to detect small and medium effect sizes. The largest factor contributing to data loss was automated timestamping system that failed, resulting in inaccurate or missing timestamps during the TSST. Timestamps are needed in order to accurately analyze different cardiac responses to stressor segments. Such data loss highlights the importance of utilizing multiple methods to track stressor time stamps, including analogue methodology, in order to ensure that digital equipment failure does not equate to significant data loss. Future studies should look to implement such protocols including manually noting the time ECG equipment is turned on as well as the time each component of the TSST (e.g., prep, speech, math) occurs. Additional safeguards may include testing systems every day prior to the start of the assessment to ensure any trouble shooting can occur prior to a participant assessment time.

Implications

Despite the minimal physiological changes in participant stress psychophysiology, there are several key implications of the present study particularly as it relates to perceived efficacy of the intervention. Twenty-seven percent of participants in the intervention condition reported perceived improvements in sleep quality. While changes in objective sleep duration were not evidenced, the improvement in sleep perception is notable for several reasons and aligns with other smartwatch-based interventions that

evidence similar change in sleep (see Kaplan et al., 2023). Sleep perception is conceptualized in the literature as an important health variable, independent of sleep duration, that is predictive of physical health outcomes, quality of life, and subjective cognitive performance (Cha et al., 2018; Matos et al., 2017). Additionally, there is evidence to suggest that sleep interventions and CBT-I strategies utilized within this intervention evidence more pronounced effects in self-reported rather than objective sleep outcomes, especially earlier in treatment (Sun et al., 2022).

Taken together in tangent with the discussion above regarding intervention length, results suggest that the intervention was able to improve participant perception of sleep and increase knowledge of sleep hygiene strategies. Findings may provide evidence for extending the duration of this smartwatch intervention in order to allow sufficient time for objective improvement in sleep duration to become evident. Furthermore, 54% of participants indicated improvement in sleep hygiene knowledge, a particularly relevant variable for college aged students as sleep hygiene has been demonstrated to have both direct and indirect effects on mental health outcomes and well-being in college at populations (Peach et al., 2016)

When analyzing results through a MINC lens, we see that the intervention group was able to evidence personally relevant, albeit non-statistically significant, improvement in their sleep perception. Echoing work by Curran et al. (2022), we recognize that while statistical significance is important, creating an intervention which produces important participant gains is also a critical component to intervention research. The sleep hygiene and related skills taught within the intervention may be readily applied within participants lives outside of the intervention context given their accessibility (e.g., easily digestible, low cost), further augmenting and extending the perceived sleep benefit evidenced within the intervention group.

Conclusion

The present study piloted a novel self-guided biofeedback intervention delivered using wrist worn wearable technology. Through integrating key components of the MINC approach, we aimed to create an intervention that could readily translate key principles of biofeedback interventions onto a self-paced

intervention more adept to long term everyday use and more readily applied in clinical settings. Results demonstrate that while participants self-reported improved sleep quality and hygiene knowledge, more research is needed to examine how we can create a wearable technology biofeedback intervention that more reliably improves both cardiac and sleep profiles of users. Future research should consider increasing direct clinical contact with participants through a scaffolded approach which involves a stepwise reduction in direct clinical involvement across iterative trials to more accurately examine the minimal degree of intervention required for change. Moderators discussed in the limitation's sections should be considered and controlled for carefully within a fully powered sample and alternative models that assess the presence of individual changes in flexibility opposed to mean level change in cardiac responses should be considered.

References

- Abbott J. H. (2014). The distinction between randomized clinical trials (RCTs) and preliminary feasibility and pilot studies: what they are and are not. *The Journal of orthopaedic and sports physical therapy*, 44(8), 555–558. <https://doi.org/10.2519/jospt.2014.0110>
- Aguilera, G. (2011). HPA axis responsiveness to stress: implications for healthy aging. *Experimental gerontology*, 46(2-3), 90-95.
- Allen, A. P., Kennedy, P. J., Cryan, J. F., Dinan, T. G., & Clarke, G. (2014). Biological and psychological markers of stress in humans: Focus on the Trier Social Stress Test. *Neuroscience & Biobehavioral Reviews*, 38, 94-124.
- Amanvermez, Y., Rahmadiana, M., Karyotaki, E., de Wit, L., Ebert, D. D., Kessler, R. C., & Cuijpers, P. (2023). Stress management interventions for college students: A systematic review and meta-analysis. *Clinical Psychology: Science and Practice*, 30(4), 423–444. <https://doi.org/10.1111/cpsp.12342>
- Appelhans, B. M., & Luecken, L. J. (2006). Heart Rate Variability as an Index of Regulated Emotional Responding. *Review of General Psychology*, 10(3), 229–240. <https://doi.org/10.1037/1089-2680.10.3.229>
- Balbo, M., Leproult, R., & Van Cauter, E. (2010). Impact of sleep and its disturbances on hypothalamo-pituitary-adrenal axis activity. *International journal of endocrinology*, 2010, 759234. <https://doi.org/10.1155/2010/759234>
- Barkman, H., Allan, A., Salisbury, M. R., Knight, E. L., Karns, C., Roos, L. E., ... Giuliano, R. J. (2020, August 25). Growth Models of Positive Caregiving Behaviours and Concurrent Autonomic Activity in Caregivers and Children during a Challenging Puzzle Task: Replication and Extension. <https://doi.org/10.31234/osf.io/8uvkc>
- Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The prevalence and correlates of depression, anxiety, and stress in a sample of college students. *Journal of affective disorders*, 173, 90-96.

- Bergquist, J., Ściubisz, A., Kaczor, A., & Silberring, J. (2002). Catecholamines and methods for their identification and quantitation in biological tissues and fluids. *Journal of Neuroscience Methods*, 113(1), 1-13.
- Berry, M. E., Chapple, I. T., Ginsberg, J. P., Gleichauf, K. J., Meyer, J.A., & Nagpal, M. L. (2014). Non-pharmacological Intervention for chronic pain in veterans: A pilot study of heart rate variability biofeedback. *Global Advances in Health and Medicine*, 3(2), 28–33. <https://doi.org/10.7453/gahmj.2013.075>.
- Blake, M. J., Blake, L. M., Schwartz, O., Raniti, M., Waloszek, J. M., Murray, G., Simmons, J. G., Landau, E., Dahl, R. E., McMakin, D. L., Dudgeon, P., Trinder, J., & Allen, N. B. (2018). Who benefits from adolescent sleep interventions? Moderators of treatment efficacy in a randomized controlled trial of a cognitive-behavioral and mindfulness-based group sleep intervention for at-risk adolescents. *Journal of Child Psychology and Psychiatry*, 59(6), 637–649. <https://doi.org/10.1111/jcpp.12842>
- Boesch, M., Sefidan, S., Ehlert, U., Annen, H., Wyss, T., Steptoe, A., & La Marca, R. (2014). Mood and autonomic responses to repeated exposure to the Trier Social Stress Test for Groups (TSST-G). *Psychoneuroendocrinology*, 43, 41-51.
- Bonnini, S., Assegie, G. M., & Trzcinska, K. (2024). Review about the permutation approach in hypothesis testing. *Mathematics*, 12(17), 2617
- Bornemann, B., Kovacs, P., & Singer, T. (2019). Voluntary upregulation of heart rate variability through biofeedback is improved by mental contemplative training. *Scientific Reports*, 9(1), Article 7860. <https://doi.org/10.1038/s41598-019-44201-7>
- Brougham, R.R., Zail, C.M., Mendoza, C.M., et al. (2009). Stress, Sex Differences, and Coping Strategies Among College Students. *Current Psychology*, 28, 85–97. <https://doiorg.uml.idm.oclc.org/10.1007/s12144-009-9047-0>

- Bůžková, P., Lumley, T., & Rice, K. (2011). Permutation and parametric bootstrap tests for gene–gene and gene–environment interactions. *Annals of human genetics*, 75(1), 36-45.
- Caldwell, Y. T., & Steffen, P. R. (2018). Adding HRV biofeedback to psychotherapy increases heart rate variability and improves the treatment of major depressive disorder. *International Journal of Psychophysiology*, 131, 96–101. <https://doi.org/10.1016/j.ijpsycho.2018.01.001>
- Can, Y. S., Arnrich, B., & Ersoy, C. (2019). Stress detection in daily life scenarios using smart phones and wearable sensors: A survey. *Journal of biomedical informatics*, 92, 103139.
- Carpenter, L.L., Gawuga, C.E., Tyrka, A.R., Lee, J.K., Anderson, G.M., Price, L.H., 2010. Association between plasma IL-6 response to acute stress and early-life adversity in healthy adults. *Neuropsychopharmacology* 35, 2617–2623.
- Castaldo, R., Melillo, P., Bracale, U., Caserta, M., Triassi, M., & Pecchia, L. (2015). Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis. *Biomedical Signal Processing and Control*, 18, 370-377.
- Chen, S., Sun, P., Wang, S., Lin, G., & Wang, T. (2016). Effects of heart rate variability biofeedback on cardiovascular responses and autonomic sympathovagal modulation following stressor tasks in prehypertensives. *Journal of human hypertension*, 30(2), 105-111.
- Chen, S., Wang, H., Yue, J., Guan, N., & Wang, X. (2022). Intervention methods for improving reduced heart rate variability in patients with major depressive disorder: A systematic review and meta-analysis. *Comprehensive Psychiatry*, 119, 152347.
- Cha, D. S., Carmona, N., Cha, R. H., Zhou, A. J., Subramaniapillai, M., Mansur, R. B., ... & McIntyre, R. S. (2018). Perceived sleep quality predicts cognitive function in adults with major depressive disorder independent of depression severity. *Annals of Clinical Psychiatry*, 30(4), 1-10.
- Chalmers, J. A., Quintana, D. S., Abbott, M. J. A., & Kemp, A. H. (2014). Anxiety disorders are associated with reduced heart rate variability: a meta-analysis. *Frontiers in Psychiatry*, 5, 80.

- Cohen, J. I. (2000). Stress and mental health: A biobehavioral perspective. *Issues in Mental Health Nursing*, 21(2), 185–202. <https://doi.org/10.1080/016128400248185>
- Curran, M., Komesidou, R., & Hogan, T. P. (2022). Less Is More: Implementing the Minimal Intervention Needed for Change Approach to Increase Contextual Fit of Speech-Language Interventions. *Language, speech, and hearing services in schools*, 53(2), 317–328. https://doi.org/10.1044/2021_LSHSS-21-00050
- Dai, R., Lu, C., Yun, L., Lenze, E., Avidan, M., & Kannampallil, T. (2021). Comparing stress prediction models using smartwatch physiological signals and participant self-reports. *Computer Methods and Programs in Biomedicine*, 208, Article 106207. <https://doi.org/10.1016/j.cmpb.2021.106207>
- Das, S., & O’Keefe, J. H. (2006). Behavioral cardiology: recognizing and addressing the profound impact of psychosocial stress on cardiovascular health. *Current atherosclerosis reports*, 8(2), 111-118.
- DeLongis, A., Folkman, S., & Lazarus, R. S. (1988). The impact of daily stress on health and mood: Psychological and social resources as mediators. *Journal of Personality and Social Psychology*, 54(3), 486–495 <https://doi.org/10.1037/0022-3514.54.3.486>
- De Witte, N. A., Buyck, I., & Van Daele, T. (2019). Combining biofeedback with stress management interventions: A systematic review of physiological and psychological effects. *Applied Psychophysiology and Biofeedback*, 44(2), 71-82.
- Dillon, A., Kelly, M., Robertson, I. H., & Robertson, D. A. (2016). Smartphone applications utilizing biofeedback can aid stress reduction. *Frontiers in psychology*, 7, 832.
- Ding, Y., Tan, K., Sheng, L., Ren, H., Su, Z., Yang, H., Zhang, X., Li, J., & Hu, P. (2024). Integrated mental stress smartwatch based on sweat cortisol and HRV sensors. *Biosensors & Bioelectronics*, 265, Article 116691. <https://doi.org/10.1016/j.bios.2024.116691>
- Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L., & Blair, S. N. (2000). Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International journal of psychophysiology*, 37(2), 121-133.

- Domes, G., Stächele, T., von Dawans, B., & Heinrichs, M. (2019). Effects of internet-based stress management on acute cortisol stress reactivity: Preliminary evidence using the Trier Social Stress Test for Groups (TSST-G). *Psychoneuroendocrinology*, *105*, 117-122.
- Domhof, S., Brunner, E., & Osgood, D. W. (2002). Rank procedures for repeated measures with missing values. *Sociological methods & research*, *30*(3), 367-393.
- Drake, C. L., Cheng, P., Almeida, D. M., & Roth, T. (2017). Familial Risk for Insomnia Is Associated With Abnormal Cortisol Response to Stress. *Sleep*, *40*(10), zsx143.
<https://doi.org/10.1093/sleep/zsx143>
- Dvořáková, K., Kishida, M., Li, J., Elavsky, S., Broderick, P. C., Agrusti, M. R., & Greenberg, M. T. (2017). Promoting healthy transition to college through mindfulness training with first-year college students: Pilot randomized controlled trial. *Journal of American College Health*, *65*(4), 259-267.
- Dunn, J. S., & Taylor, C. E. (2014). Cardiovascular reactivity to stressors: effect of time of day?. *Chronobiology international*, *31*(2), 166-174.
- Ebben, M. R., Kurbatov, V., & Pollak, C. P. (2009). Moderating Laboratory Adaptation with the Use of a Heart-rate Variability Biofeedback Device (StressEraser®). *Applied Psychophysiology and Biofeedback*, *34*(4), 245–249. <https://doi.org/10.1007/s10484-009-9086-1>
- Economides, M., Lehrer, P., Ranta, K., Nazander, A., Hilgert, O., Raevuori, A., Gevirtz, R., Khazan, I., & Forman-Hoffman, V. L. (2020). Feasibility and Efficacy of the Addition of Heart Rate Variability Biofeedback to a Remote Digital Health Intervention for Depression. *Applied Psychophysiology and Biofeedback*, *45*(2), 75–86. <https://doi.org/10.1007/s10484-020-09458-z>
- Eisenberger, N.I., Taylor, S.E., Gable, S.L., Hilmert, C.J., Lieberman, M.D., 2007. Neural pathways link social support to attenuated neuroendocrine stress responses. *Neuroimage* *35*, 1601–1612.
- Esch, T., Stefano, G. B., Fricchione, G. L., & Benson, H. (2002). Stress in cardiovascular diseases. *Signature*, *8*(5), 101.

- Fellinger, E., Brandt, T., Schittenhelm, A., Quarg, E., Pröll, M., Domes, G., & Schmidt, A. (2025). Virtual stress testing: Analyzing endocrine, metabolic, cardiovascular, and psychological responses to the TSST-VR. *Neurobiology of Stress*, 39, Article 100760.
<https://doi.org/10.1016/j.ynstr.2025.100760>
- Garcia, L., Ferguson, S. E., Facio, L., Schary, D., & Guenther, C. H. (2023). Assessment of well-being using Fitbit technology in college students, faculty, and staff completing breathing meditation during COVID-19: A pilot study. *Mental Health & Prevention*, 30, 200280.
- Geisler, F. C. M., Vennwald, N., Kubiak, T., & Weber, H. (2010). The impact of heart rate variability on subjective well-being is mediated by emotion regulation. *Personality and Individual Differences*, 49(7), 723–728. <https://doi.org/10.1016/j.paid.2010.06.015>
- Gevirtz, R. N. (2024). Integrating Heart Rate Variability Biofeedback into Acceptance and Commitment Therapy (ACT). In *Integrating Psychotherapy and Psychophysiology*. Oxford University Press.
<https://doi.org/10.1093/oso/9780198888727.003.0015>
- Gevirtz, R. (2015). Integrating Heart Rate Variability Biofeedback into Mindfulness-Based Therapies. *Biofeedback (Wheat Ridge, Colo.)*, 43(3), 129–132. <https://doi.org/10.5298/1081-5937-43.3.03>
- Giuliano, R. J., Skowron, E. A., & Berkman, E. T. (2015). Growth models of dyadic synchrony and mother–child vagal tone in the context of parenting at-risk. *Biological psychology*, 105, 29–36.
- Glasgow, R. E., Fisher, L., Strycker, L. A., Hessler, D., Toobert, D. J., King, D. K., & Jacobs, T. (2014). Minimal intervention needed for change: definition, use, and value for improving health and health research. *Translational behavioral medicine*, 4(1), 26–33. <https://doi.org/10.1007/s13142-013-0232-1>
- González-García, M., Álvarez, J. C., Pérez, E. Z., Fernandez-Carriba, S., & López, J. G. (2021). Feasibility of a brief online mindfulness and compassion-based intervention to promote mental health among university students during the COVID-19 pandemic. *Mindfulness*, 12(7), 1685–1695.
- Grodin, E. N., Kirsch, D., Belnap, M., & Ray, L. A. (2024). Sex differences in neural

- response to an acute stressor in individuals with an alcohol use disorder. *Alcohol, clinical & experimental research*, 48(5), 843–854. <https://doi.org/10.1111/acer.15301>
- Green L. W. (2008). Making research relevant: if it is an evidence-based practice, where's the practice-based evidence? *Family practice*, 25 Suppl 1, i20–i24. <https://doi.org/10.1093/fampra/cmn055>
- Grodin, E. N., Kirsch, D., Belnap, M., & Ray, L. A. (2024). Sex differences in neural response to an acute stressor in individuals with an alcohol use disorder. *Alcohol: Clinical and Experimental Research*, 48(5), 843-854.
- Hallman, D. M., Olsson, E. M., Von Schéele, B., Melin, L., & Lyskov, E. (2011). Effects of heart rate variability biofeedback in subjects with stress-related chronic neck pain: a pilot study. *Applied psychophysiology and biofeedback*, 36, 71-80.
- Hamilton, J. L., & Alloy, L. B. (2016). Atypical reactivity of heart rate variability to stress and depression across development: Systematic review of the literature and directions for future research. *Clinical psychology review*, 50, 67-79.
- Hamidovic, A., Van Hedger, K., Choi, S. H., Flowers, S., Wardle, M., & Childs, E. (2020). Quantitative meta-analysis of heart rate variability finds reduced parasympathetic cardiac tone in women compared to men during laboratory-based social stress. *Neuroscience and biobehavioral reviews*, 114, 194–200. <https://doi.org/10.1016/j.neubiorev.2020.04.005>
- Han, K. S., Kim, L., & Shim, I. (2012). Stress and sleep disorder. *Experimental neurobiology*, 21(4), 141–150. <https://doi.org/10.5607/en.2012.21.4.141>
- Hartmann, R., Schmidt, F. M., Sander, C., & Hegerl, U. (2019). Heart rate variability as indicator of clinical state in depression. *Frontiers in Psychiatry*, 9, 735.
- Harvie, H. M., Jain, B., Nelson, B. W., Knight, E. L., Roos, L. E., & Giuliano, R. J. (2021). Induction of acute stress through an internet-delivered Trier Social Stress Test as assessed by photoplethysmography on a smartphone. *Stress*, 24(6), 1023-1032.

- Held, J., Vîslă, A., Wolfer, C., Messerli-Bürgy, N., & Flückiger, C. (2021). Heart rate variability change during a stressful cognitive task in individuals with anxiety and control participants. *BMC psychology*, 9(1), 44.
- Hellhammer, J., & Schubert, M. (2012). The physiological response to Trier Social Stress Test relates to subjective measures of stress during but not before or after the test. *Psychoneuroendocrinology*, 37(1), 119–124. <https://doi.org/10.1016/j.psyneuen.2011.05.012>
- Henckens, M. J., van Wingen, G. A., Joëls, M., & Fernández, G. (2012). Time-dependent effects of cortisol on selective attention and emotional interference: A functional MRI study. *Frontiers in Integrative Neuroscience*, 6(66).
- Hoffmann, T., & Worrall, L. (2004). Designing effective written health education materials: Considerations for health professionals. *Disability and Rehabilitation*, 26(19), 1166-1173.
- Immanuel, S., Teferra, M. N., Baumert, M., & Bidargaddi, N. (2023). Heart rate variability for evaluating psychological stress changes in healthy adults: A scoping review. *Neuropsychobiology*, 82(4), 187-202.
- Jerath, R., Syam, M., & Ahmed, S. (2023). The future of stress management: integration of smartwatches and HRV technology. *Sensors*, 23(17), 7314.
- Johnston, D., Foord, R., Casties, A., Viaphay, N., Tohamy, A., Van Leeuwen, N., Sinclair, J., Talbot, D., & Harris, A. (2024). Integrating smartwatches in community mental health services for severe mental illness for detecting relapse and informing future intervention: A case series. *Early Intervention in Psychiatry*, 18(6), 471–477. <https://doi.org/10.1111/eip.13529>
- Kaplan, M., McKniff, M., Holmqvist, S., Simone, S., Rosahl, M., Chaturvedi, R., & Giovannetti, T. (2023). A - 04 Influence of Smartwatch Feedback on Perceived Health Behavior Change in Older Adults. *Archives of Clinical Neuropsychology*, 38(7), 1165–1165. <https://doi.org/10.1093/arclin/acad067.022>
- Kapogianni, N. A., Sideraki, A., & Anagnostopoulos, C. N. (2025). Using smartwatches in stress management, mental health, and well-being: A systematic review. *Algorithms*, 18(7), 419.

- Karavidas, M. K., Lehrer, P. M., Vaschillo, E., Vaschillo, B., Marin, H., Buyske, S., Malinovsky, I., Radvanski, D., & Hassett, A. (2007). Preliminary Results of an Open Label Study of Heart Rate Variability Biofeedback for the Treatment of Major Depression. *Applied Psychophysiology and Biofeedback*, 32(1), 19–30. <https://doi.org/10.1007/s10484-006-9029-z>
- Kanuri, N., Arora, P., Talluru, S., Colaco, B., Dutta, R., Rawat, A., ... & Newman, M. G. (2020). Examining the initial usability, acceptability and feasibility of a digital mental health intervention for college students in India. *International Journal of Psychology*, 55(4), 657-673.
- Kelly, M. M., Tyrka, A. R., Anderson, G. M., Price, L. H., & Carpenter, L. L. (2008). Sex differences in emotional and physiological responses to the Trier Social Stress Test. *Journal of behavior therapy and experimental psychiatry*, 39(1), 87-98
- Kemeny, M. E. (2003). The psychobiology of stress. *Current Directions in Psychological Science*, 12(4), 124-129.
- Khan, A., Song, M., & Dong, Z. (2025). Chronic stress: a fourth etiology in tumorigenesis?. *Molecular cancer*, 24(1), 196.
- Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018). Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry investigation*, 15(3), 235.
- Knight, E. L., Giuliano, R. J., Shank, S. W., Clarke, M. M., & Almeida, D. M. (2020). Parasympathetic and sympathetic nervous systems interactively predict change in cognitive functioning in midlife adults. *Psychophysiology*, 57(10), e13622.
- Koolhaas, J. M., Bartolomucci, A., Buwalda, B., de Boer, S. F., Flügge, G., Korte, S. M., ... & Fuchs, E. (2011). Stress revisited: a critical evaluation of the stress concept. *Neuroscience & Biobehavioral Reviews*, 35(5), 1291-1301.
- Kotozaki, Y., Takeuchi, H., Sekiguchi, A., Yamamoto, Y., Shinada, T., Araki, T., Takahashi, K., Taki, Y., Ogino, T., Kiguchi, M., & Kawashima, R. (2014). Biofeedback-based training for stress management in daily hassles: an intervention study. *Brain and Behavior*, 4(4), 566–579. <https://doi.org/10.1002/brb3.241>

- Kudielka, B. M., Buske-Kirschbaum, A., Hellhammer, D. H., & Kirschbaum, C. (2004). Differential heart rate reactivity and recovery after psychosocial stress (TSST) in healthy children, younger adults, and elderly adults: the impact of age and gender. *International Journal of Behavioral Medicine*, 11(2), 116–121. https://doi.org/10.1207/s15327558ijbm1102_8
- Kudielka, B. M., Schommer, N. C., Hellhammer, D. H., & Kirschbaum, C. (2004). Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. *Psychoneuroendocrinology*, 29(8), 983–992.
<https://doi.org/10.1016/j.psyneuen.2003.08.009>
- Kudo, N., Shinohara, H., & Kodama, H. (2014). Heart rate variability biofeedback intervention for reduction of psychological stress during the early postpartum period. *Applied Psychophysiology and Biofeedback*, 39(3–4), 203–211. <https://doi.org/10.1007/s10484-014-9259-4>
- Knox, M., Lentini, J., Cummings, T., McGrady, A., Whearty, K., & Sancrant, L. (2011). Game- based biofeedback for paediatric anxiety and depression. *Mental Health in Family Medicine*, 8(3), 195–203.
- Lalanza, J. F., Lorente, S., Bullich, R., García, C., Losilla, J.-M., & Capdevila, L. (2023). Methods for heart rate variability biofeedback (HRVB): A systematic review and guidelines. *Applied Psychophysiology and Biofeedback*, 48(3), 275–297.
- Lauderdale, D. S., Knutson, K. L., Yan, L. L., Liu, K., & Rathouz, P. J. (2008a). Self-reported and measured sleep duration: how similar are they?. *Epidemiology*, 19(6), 838-845.
- Lauderdale, D. S., Knutson, K. L., Yan, L. L., Liu, K., & Rathouz, P. J. (2008b). Sleep duration: how well do self-reports reflect objective measures? The CARDIA Sleep Study. *Epidemiology (Cambridge, Mass.)*, 19(6), 838.
- Lee, S. Y., Wuertz, C., Rogers, R., & Chen, Y. P. (2013). Stress and sleep disturbances in female college students. *American journal of health behavior*, 37(6), 851-858.

- Lehrer, P. M., & Gevirtz, R. (2014). Heart rate variability biofeedback: How and why does it work? *Frontiers in Psychology, 5*, 756–756.
- Lemaire, J. B., Wallace, J. E., Lewin, A. M., de Grood, J., & Schaefer, J. P. (2011). The effect of a biofeedback-based stress management tool on physician stress: A randomized controlled clinical trial. *Open Medicine, 5*(4), e154.
- Lennartsson, A.-K., Kushnir, M. M., Bergquist, J., Billig, H., & Jonsdottir, I. H. (2012). Sex steroid levels temporarily increase in response to acute psychosocial stress in healthy men and women. *International Journal of Psychophysiology, 84*(3), 246–253.
<https://doi.org/10.1016/j.ijpsycho.2012.03.001>
- Leppink, E. W., Odlaug, B. L., Lust, K., Christenson, G., & Grant, J. E. (2016). The young and the stressed: Stress, impulse control, and health in college students. *The Journal of Nervous and Mental Disease, 204* (12), 931–938. <https://doi.org/10.1097/NMD.0000000000000586>
- Li, L., Wang, Z., Cui, L., Xu, Y., Lee, H., & Guan, K. (2023). The efficacy of a novel smart watch on medicine adherence and symptom control of allergic rhinitis patients: Pilot study. *The World Allergy Organization journal, 16*(1), 100739. <https://doi.org/10.1016/j.waojou.2022.100739>
- Lim, S. E., Kim, H. S., Lee, S. W., Bae, K. H., & Baek, Y. H. (2023). Validation of Fitbit Inspire 2™ against polysomnography in adults considering adaptation for use. *Nature and Science of Sleep, 15*, 59-68. <https://doi.org/10.2147/NSS.S365724>
- Liu, J. J., Ein, N., Peck, K., Huang, V., Pruessner, J. C., & Vickers, K. (2017). Sex differences in salivary cortisol reactivity to the Trier Social Stress Test (TSST): A meta-analysis. *Psychoneuroendocrinology, 82*, 26-37.
- Lund, H. G., Reider, B. D., Whiting, A. B., & Prichard, J. R. (2010). Sleep patterns and predictors of disturbed sleep in a large population of college students. *Journal of adolescent health, 46*(2), 124-132.

- Lo Martire, V., Caruso, D., Palagini, L., Zoccoli, G., & Bastianini, S. (2020). Stress & sleep: A relationship lasting a lifetime. *Neuroscience and Biobehavioral Reviews*, 117, 65–77.
<https://doi.org/10.1016/j.neubiorev.2019.08.024>
- Low, C. A., Danko, M., Durica, K. C., Vega, J., Li, M., Kunta, A. R., Mulukutla, R., Ren, Y., Sereika, S. M., Bartlett, D. L., Bovbjerg, D. H., Dey, A. K., & Jakicic, J. M. (2023). A Real-Time Mobile Intervention to Reduce Sedentary Behavior Before and After Cancer Surgery: Pilot Randomized Controlled Trial. *JMIR perioperative medicine*, 6, e41425.
- Matos, M. G., Marques, A., Gaspar, T., & Paiva, T. (2017). Perception of quantity and quality of sleep and their association with health related quality of life and life satisfaction during adolescence.
- Matsuda, S., Yamaguchi, T., Okada, K., Gotouda, A., & Mikami, S. (2012). Day-to-day variations in salivary cortisol measurements. *Journal of prosthodontic research*, 56(1), 37–41. <https://doi.org/10.1016/j.jpor.2011.04.004>
- Martinez, G. J., Grover, T., Mattingly, S. M., Mark, G., D’Mello, S., Aledavood, T., ... & Striegel, A. (2022). Alignment between heart rate variability from fitness trackers and perceived stress: Perspectives from a large-scale in situ longitudinal study of information workers. *JMIR human factors*, 9(3), e33754.
- Marvel, F. A., Spaulding, E. M., Lee, M. A., Yang, W. E., Demo, R., Ding, J., Wang, J., Xun, H., Shah, L. M., Weng, D., Carter, J., Majmudar, M., Elgin, E., Sheidy, J., McLin, R., Flowers, J., Vilarino, V., Lumelsky, D. N., Bhardwaj, V., ... Martin, S. S. (2021). Digital Health Intervention in Acute Myocardial Infarction. *Circulation Cardiovascular Quality and Outcomes*, 14(7), e007741–e007741. <https://doi.org/10.1161/CIRCOUTCOMES.121.007741>
- McEwen B. S. (2017). Neurobiological and Systemic Effects of Chronic Stress. *Chronic stress (Thousand Oaks, Calif.)*, 1, 2470547017692328. <https://doi.org/10.1177/2470547017692328>

- Miller, R., & Kirschbaum, C. (2019). Cultures under stress: A cross-national meta-analysis of cortisol responses to the Trier Social Stress Test and their association with anxiety-related value orientations and internalizing mental disorders. *Psychoneuroendocrinology*, 105, 147-154.
- Mohr, C., Braun, S., Bridler, R., Chmetz, F., Delfino, J. P., Kluckner, V. J., ... & Stassen, H. H. (2014). Insufficient coping behavior under chronic stress and vulnerability to psychiatric disorders. *Psychopathology*, 47(4), 235-243.
- Muhtadie, L., Koslov, K., Akinola, M., & Mendes, W. B. (2015). Vagal flexibility: A physiological predictor of social sensitivity. *Journal of personality and social psychology*, 109(1), 106.
- Nabi, H., Kivimäki, M., Batty, G. D., Shipley, M. J., Britton, A., Brunner, E. J., ... & Singh-Manoux, A. (2013). Increased risk of coronary heart disease among individuals reporting adverse impact of stress on their health: the Whitehall II prospective cohort study. *European heart journal*, 34(34), 2697-2705.
- Nakagawa, S., & Schielzeth, H. (2013). *A general and simple method for obtaining R^2 from generalized linear mixed-effects models*. *Methods in Ecology and Evolution*, 4(2), 133–142. <https://doi.org/10.1111/j.2041-210X.2012.00261.x>
- Nakao, M. (2019). Heart rate variability and perceived stress as measurements of relaxation response. *Journal of Clinical Medicine*, 8(10), 1704.
- Nikolajsen, R. P., & Hansen, Å. M. (2001). Analytical methods for determining urinary catecholamines in healthy subjects. *Analytica chimica acta*, 449(1-2), 1-15.
- Noushad, S., Ahmed, S., Ansari, B., Mustafa, U. H., Saleem, Y., & Hazrat, H. (2021). Physiological biomarkers of chronic stress: A systematic review. *International journal of health sciences*, 15(5), 46–59.
- O'Connor, D. B., Thayer, J. F., & Vedhara, K. (2021). Stress and Health: A Review of Psychobiological Processes. *Annual Review of Psychology*, 72(1), 663–688. <https://doi.org/10.1146/annurev-psych-062520-122331>

- Okun, M. L., & Glidewell, R. N. (2023). Improvement of Insomnia Symptoms following a Single 4-Hour CBT-I Workshop. *Behavioral sleep medicine*, 21(1), 72–83.
- Paul, M., & Garg, K. (2012). The effect of heart rate variability biofeedback on performance psychology of basketball players. *Applied Psychophysiology And Biofeedback*, 37(2), 131– 144.
<https://doi.org/10.1007/s10484-012-9185-2>
- Peach, H., Gaultney, J. F., & Gray, D. D. (2016). Sleep hygiene and sleep quality as predictors of positive and negative dimensions of mental health in college students. *Cogent Psychology*, 3(1), 1168768.
- Petersen, H., Kecklund, G., D'onofrio, P. A. O. L. O., Nilsson, J., & Åkerstedt, T. (2013). Stress vulnerability and the effects of moderate daily stress on sleep polysomnography and subjective sleepiness. *Journal of sleep research*, 22(1), 50-57.
- Prinsloo, G. E., Derman, W. E., Lambert, M. I., & Laurie Rauch, H. G. (2013). The effect of a single session of short duration biofeedback-induced deep breathing on measures of heart rate variability during laboratory-induced cognitive stress: a pilot study. *Applied psychophysiology and biofeedback*, 38(2), 81-90.
- Pulopulos, M. M., Vanderhasselt, M.-A., & De Raedt, R. (2018). Association between changes in heart rate variability during the anticipation of a stressful situation and the stress-induced cortisol response. *Psychoneuroendocrinology*, 94, 63–71. <https://doi.org/10.1016/j.psyneuen.2018.05.004>
- Ramón-Arbués, E., Gea-Caballero, V., Granada-López, J. M., Juárez-Vela, R., Pellicer-García, B., & Antón-Solanas, I. (2020). The prevalence of depression, anxiety and stress and their associated factors in college students. *International journal of environmental research and public health*, 17(19), 7001.
- Rajendra Acharya, U., Paul Joseph, K., Kannathal, N., Lim, C. M., & Suri, J. S. (2006). Heart rate variability: a review. *Medical & Biological Engineering & Computing*, 44(12), 1031–1051.
<https://doi.org/10.1007/s11517-006-0119-0>
- Reiner, R. (2008). Integrating a portable biofeedback device into clinical practice for patients with anxiety disorders: Results of a pilot study. *Applied Psychophysiology and Biofeedback*, 33(1), 55–61.

- Roos, L. E., Knight, E. L., Beauchamp, K. G., Berkman, E. T., Faraday, K., Hyslop, K., & Fisher, P. A. (2017). Acute stress impairs inhibitory control based on individual differences in parasympathetic nervous system activity. *Biological Psychology*, 125, 58–63.
<https://doi.org/10.1016/j.biopsycho.2017.03.004>
- Schiweck, C., Piette, D., Berckmans, D., Claes, S., & Vrieze, E. (2019). Heart rate and high frequency heart rate variability during stress as biomarker for clinical depression. A systematic review. *Psychological Medicine*, 49(2), 200–211. <https://doi.org/10.1017/S0033291718001988>
- Schumann, A., Köhler, S., Brotte, L., & Bär, K. J. (2019). Effect of an eight-week smartphone-guided HRV-biofeedback intervention on autonomic function and impulsivity in healthy controls. *Physiological measurement*, 40(6), 064001.
- Sekhon, M., Cartwright, M., & Francis, J. J. (2022). Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. *BMC Health Services Research*, 22(1), Article 279. <https://doi.org/10.1186/s12913-022-07577-3>
- Sekhon, M., Cartwright, M., & Francis, J. J. (2017). Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC health services research*, 17(1), 88.
- Sharma, P., & Jyoti, J. (2022). Role of IoT and data analysis in determining mental wellbeing. *International Journal of Health Sciences*, 6(S3), 8140–8144.
<https://doi.org/10.53730/ijhs.v6nS3.7945>
- Signore, A. K., Jung, M. E., Semenchuk, B., Kullman, S. M., Tefft, O., Webber, S., Ferguson, L. J., Kowalski, K., Fortier, M., McGavock, J., Ahmed, R., Orr, M., & Strachan, S. (2022). A pilot and feasibility study of a randomized clinical trial testing a self-compassion intervention aimed to increase physical activity behaviour among people with prediabetes. *Pilot and Feasibility Studies*, 8(1), Article 111. <https://doi.org/10.1186/s40814-022-01072-6>
- Silva, A. G. C. B. da, Araujo, D. N., Costa, A. C. M., Dias, B. A. L., Fregonezi, G. A. de F., & Dias, F. A. L. (2015). Increase in perceived stress is correlated to lower heart rate variability in healthy

- young subjects. *Acta Scientiarum. Health Sciences*, 37(1), 7–10.
<https://doi.org/10.4025/actascihealthsci.v37i1.21676>
- Sitka, M., Simpson, K. M., Paton, A., MacKinnon, A., Afifi, T. O., Cameron, E. E., ... & Roos, L. E. (2026). Advancing family wellbeing through a Massive Online Open Intervention: the LightBEAM program protocol for randomized waitlist control trial {1a}. *BMC psychology*.
- Sitka, M., McHardy, R. J., Watts, D., Roos, L. E., Tomfohr-Madsen, L., Turner, H., ... & Giuliano, R. (2024). Smartwatch photoplethysmography heart rate measures are associated with parenting behaviors in mothers with elevated mental health symptoms. *Translational Issues in Psychological Science*, 10(2), 232.
- Spangler, D. P., Gamble, K. R., McGinley, J. J., Thayer, J. F., & Brooks, J. R. (2018). Intra-individual variability in vagal control is associated with response inhibition under stress. *Frontiers in human neuroscience*, 12, 475.
- Steffen, P. R., Austin, T., DeBarros, A., & Brown, T. (2017). The impact of resonance frequency breathing on measures of heart rate variability, blood pressure, and mood. *Frontiers in public health*, 5, 222.
- Steptoe, A., & Kivimäki, M. (2012). Stress and cardiovascular disease. *Nature reviews cardiology*, 9(6), 360-370.
- Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2021). partR2: partitioning R^2 in generalized linear mixed models. *PeerJ*, 9, e11414. <https://doi.org/10.7717/peerj.11414>
- Sun, Q., Dai, Y., Chen, B., Vgontzas, A. N., Basta, M., Tang, X., Zhang, S., & Li, Y. (2022). The underestimation of sleep duration phenotype is associated with better treatment response to cognitive behavior therapy for insomnia in patients with chronic insomnia: a preliminary study. *Journal of clinical sleep medicine : JCSM : official publication of the American Academy of Sleep Medicine*, 18(10), 2443–2450. <https://doi.org/10.5664/jcsm.10130>
- Thabrew, H., Ruppeldt, P., & Sollers, J. J. (2018). Systematic Review of Biofeedback Interventions for Addressing Anxiety and Depression in Children and Adolescents with Long-Term Physical

- Conditions. *Applied Psychophysiology and Biofeedback*, 43(3), 179–192.
<https://doi.org/10.1007/s10484-018-9399-z>
- Timurtas, E., Inceer, M., Mayo, N., Karabacak, N., Sertbas, Y., & Polat, M. G. (2022). Technology-based and supervised exercise interventions for individuals with type 2 diabetes: Randomized controlled trial. *Primary care diabetes*, 16(1), 49–56. <https://doi.org/10.1016/j.pcd.2021.12.005>
- Triantafyllidis, A., Kondylakis, H., Katehakis, D., Kouroubali, A., Alexiadis, A., Segkouli, S., ... & Tzovaras, D. (2024). Smartwatch interventions in healthcare: A systematic review of the literature. *International journal of medical informatics*, 190, 105560.
- Uhart, M., Chong, R. Y., Oswald, L., Lin, P. I., & Wand, G. S. (2006). Gender differences in hypothalamic-pituitary-adrenal (HPA) axis reactivity. *Psychoneuroendocrinology*, 31(5), 642–652. <https://doi.org/10.1016/j.psyneuen.2006.02.003>
- Verlander, L. A., Benedict, J. O., & Hanson, D. P. (1999). Stress and Sleep Patterns of College Students. *Perceptual and Motor Skills*, 88(3), 893–898.
<https://doi.org/10.2466/pms.1999.88.3.893>
- Wadsworth, M. E., Broderick, A. V., Loughlin-Presnal, J. E., Bendezu, J. J., Joos, C. M., Ahlkvist, J. A., ... & McDonald, A. (2019). Co-activation of SAM and HPA responses to acute stress: A review of the literature and test of differential associations with preadolescents' internalizing and externalizing. *Developmental psychobiology*, 61(7), 1079-1093.
- Wallace, D. D., Boynton, M. H., & Lytle, L. A. (2017). Multilevel analysis exploring the links between stress, depression, and sleep problems among two-year college students. *Journal of American college health*, 65(3), 187-196.
- Walker, J., Muench, A., Perlis, M. L., & Vargas, I. (2022). Cognitive Behavioral Therapy for Insomnia (CBT-I): A Primer. *Klinicheskaia i Spetsial'naia Psikhologiia*, 11(2), 123–137.
<https://doi.org/10.17759/cpse.2022110208>
- Wang, C., Lizardo, O., & Hachen, D. S. (2022). Using Fitbit data to monitor the heart rate evolution patterns of college students. *Journal of American College Health*, 70(3), 875–882.

- Wells, R., Outhred, T., Heathers, J. A. J., Quintana, D. S., & Kemp, A. H. (2012). Matter Over Mind: A Randomised-Controlled Trial of Single-Session Biofeedback Training on Performance Anxiety and Heart Rate Variability in Musicians. *PloS One*, 7(10), e46597.
<https://doi.org/10.1371/journal.pone.0046597>
- Weinstein, N.D. Unrealistic optimism about susceptibility to health problems. *J Behav Med* 5, 441–460 (1982). <https://doi.org/10.1007/BF00845372>
- Whited, A., Larkin, K.T. & Whited, M. Effectiveness of *emWave* Biofeedback in Improving Heart Rate Variability Reactivity to and Recovery from Stress. *Appl Psychophysiol Biofeedback* 39, 75–88 (2014). <https://doi.org/10.1007/s10484-014-9243-z>
- Williams, P. G., Cribbet, M. R., Rau, H. K., Gunn, H. E., & Czajkowski, L. A. (2013). The Effects of Poor Sleep on Cognitive, Affective, and Physiological Responses to a Laboratory Stressor. *Annals of Behavioral Medicine*, 46(1), 40–51. <https://doi.org/10.1007/s12160-013-9482-x>
- Yu, B., Funk, M., Hu, J., Wang, Q., & Feijs, L. (2018). Biofeedback for everyday stress management: A systematic review. *Frontiers in ICT*, 5, 23.
- Zohar, D., Tzischinsky, O., Epstein, R., & Lavie, P. (2005). The effects of sleep loss on medical residents' emotional reactions to work events: a cognitive-energy model. *Sleep*, 28(1), 47-54

Appendix A

Lay Statement

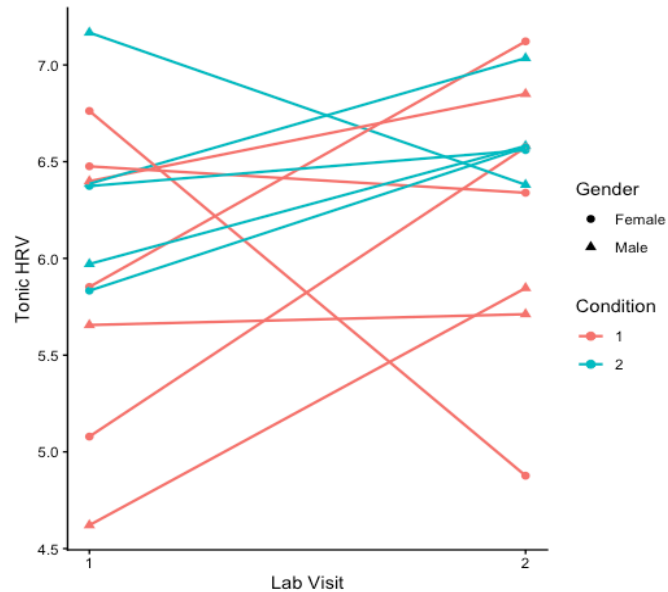
College students experience lots of stress, which can have negative effects on their health and well-being. Getting good sleep and using deep breathing strategies can help reduce stress and stop these negative health outcomes. This study recruited 14 young adults to learn about their sleep and deep breathing strategies using wearable technology (e.g., Fitbit watches). We then looked to see if these strategies lowered the level of stress they experienced during an in-person lab visit where they had to give a speech and do math tasks that were stressful. While measures of heart rate did not change while people completed the math and speech task, people reported having better sleep.

Appendix B

Individual Trajectory Plots

Figure 1

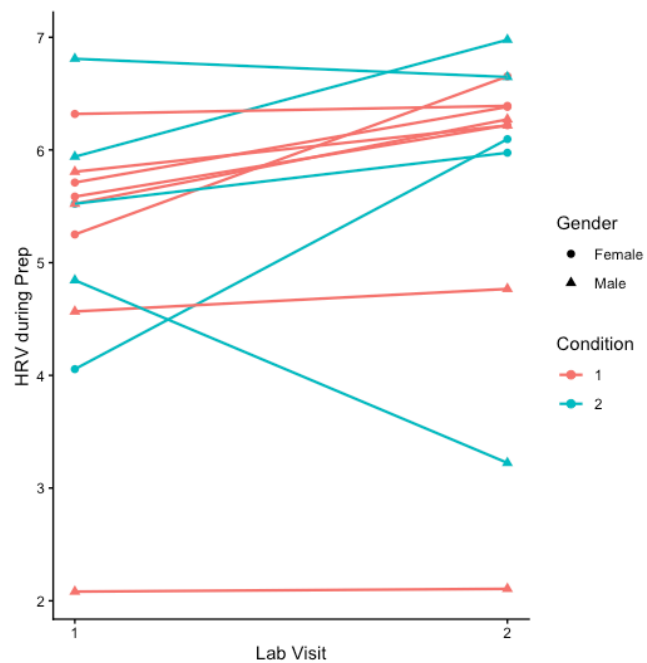
Individual Trajectory Plots Tonic HRV Recovery



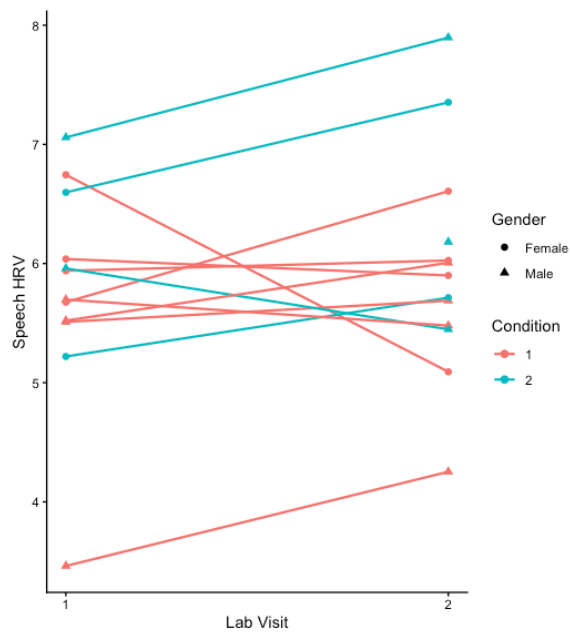
Note. N = 14; Condition 1 = Control; Condition 2 = Intervention

Figure 2

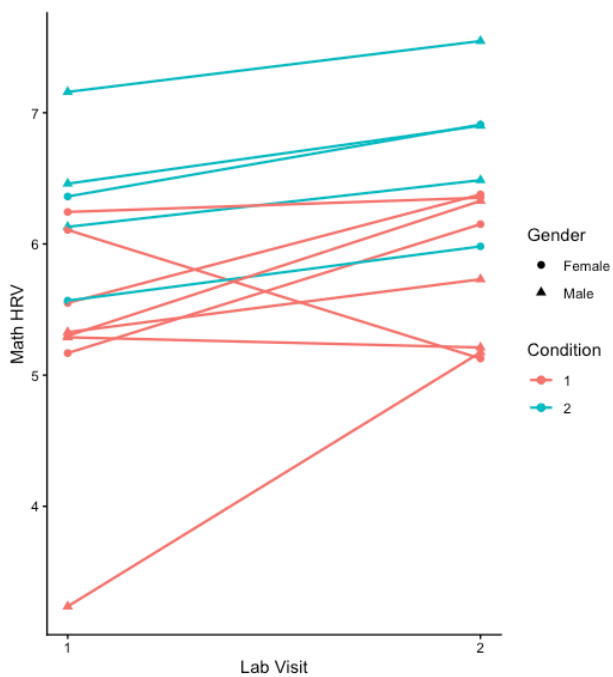
Individual Trajectory Plots HRV Prep Portion TSST



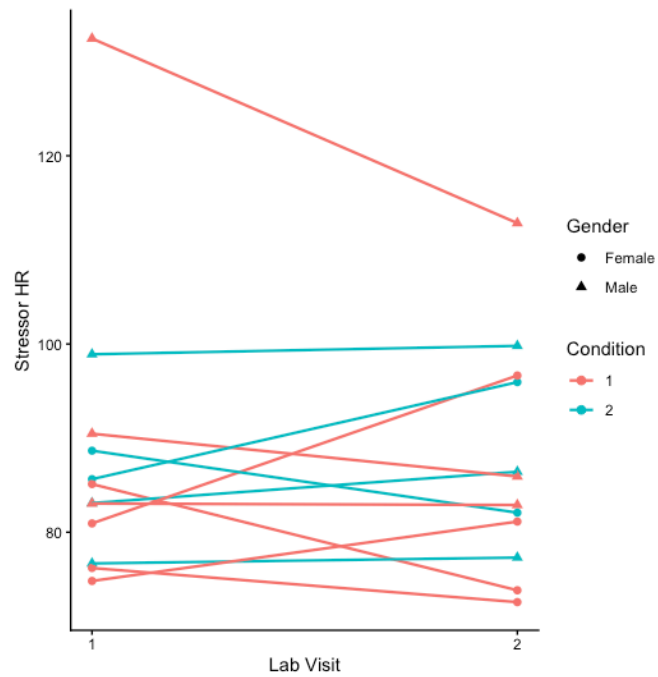
Note. N = 14; Condition 1 = Control; Condition 2 = Intervention

Figure 3*Individual Trajectory Plots HRV During Speech Portion TSST*

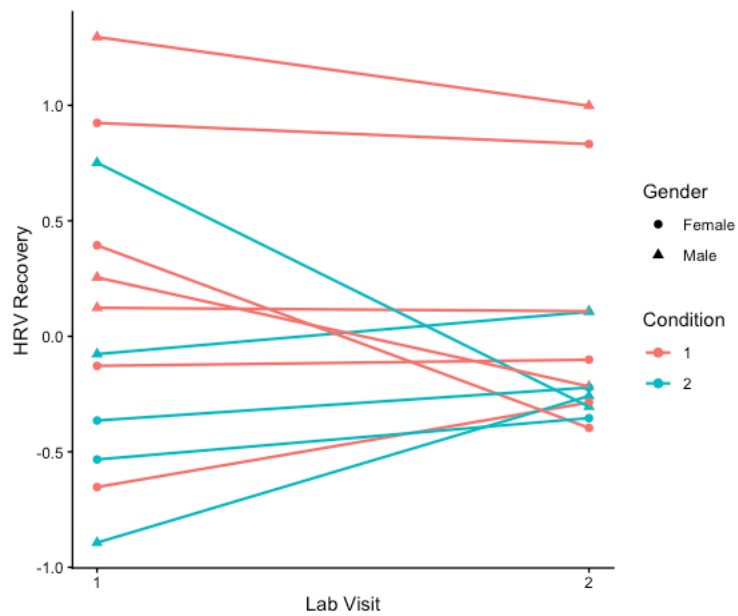
Note. N = 13; Condition 1 = Control; Condition 2 = Intervention

Figure 4*Individual Trajectory Plots HRV During Math Portion TSST*

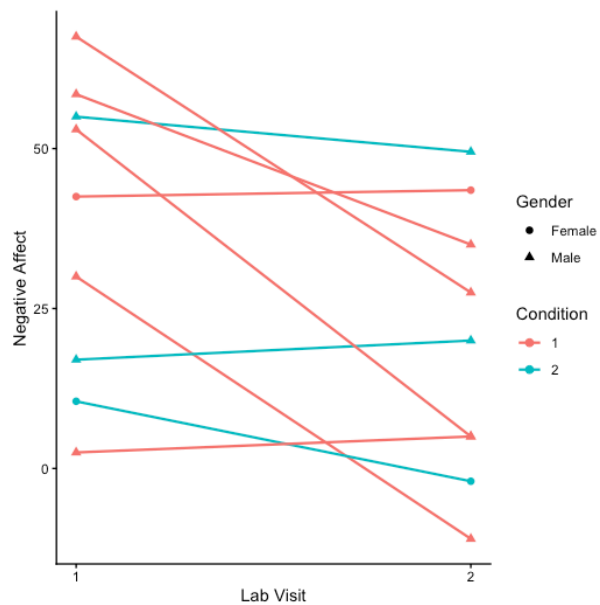
Note. N = 13; Condition 1 = Control; Condition 2 = Intervention

Figure 5*Individual Trajectory Plots HR During Stressor Portion TSST*

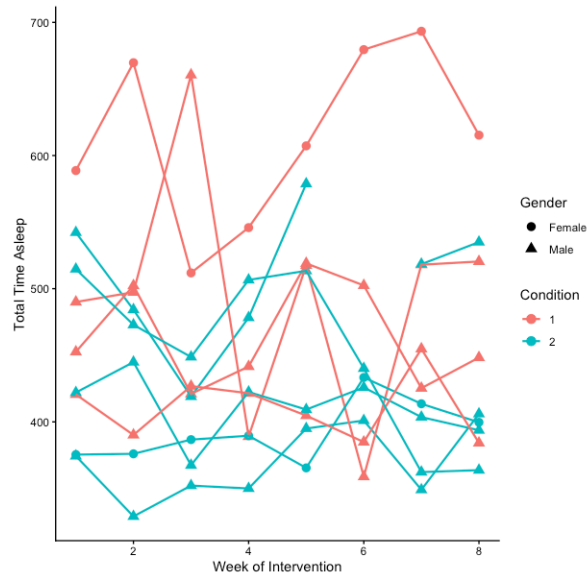
Note. N = 14; Condition 1 = Control; Condition 2 = Intervention

Figure 6*Individual Trajectory Plots HRV Recovery*

Note. N = 12; Condition 1 = Control; Condition 2 = Intervention

Figure 7*Individual Trajectory Plots Negative Affect across repeated TSST*

Note. N = 9; Condition 1 = Control; Condition 2 = Intervention

Figure 8*Individual Trajectory Plots Sleep across repeated TSST*

Note. N = 8; Condition 1 = Control; Condition 2 = Intervention

Appendix C**Residual vs Fitted Plots for Permutation Mixed Models****Figure 1**

Residual vs Fitted Plot for Model 1: Tonic HRV

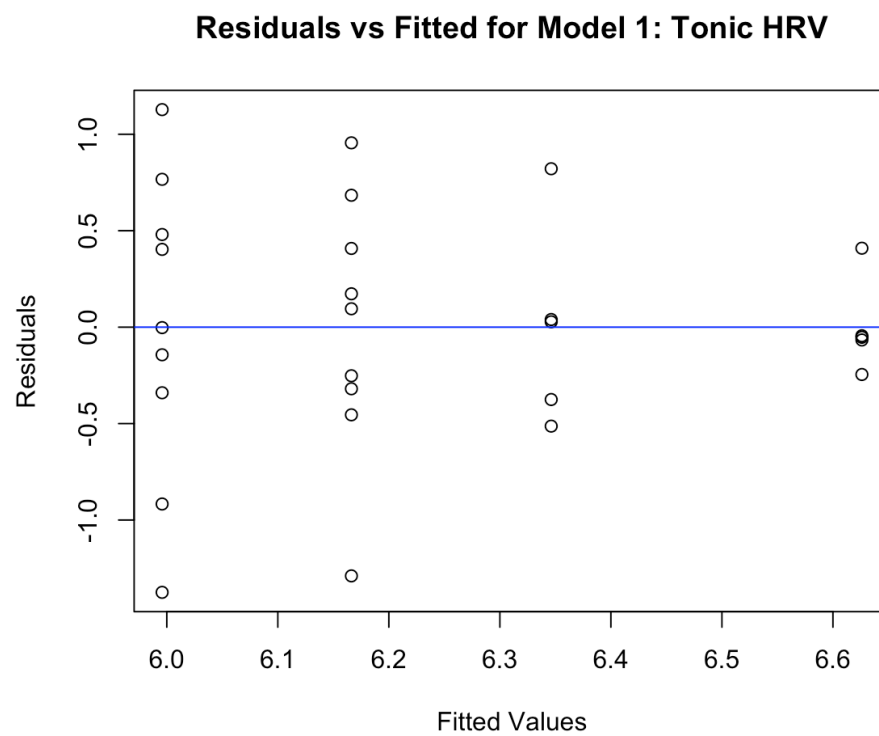
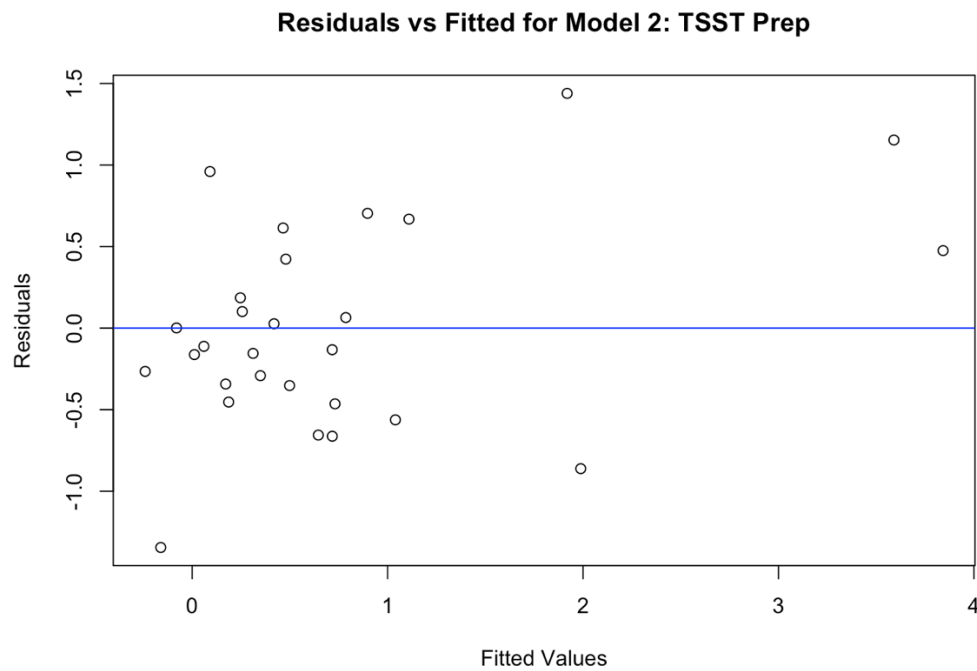


Figure 2

Residual vs Fitted Plot for Model 2: HRV during “Prep” task of TSST

**Figure 3**

Residual vs Fitted Plot for Model 3: HRV during “Speech” task of TSST

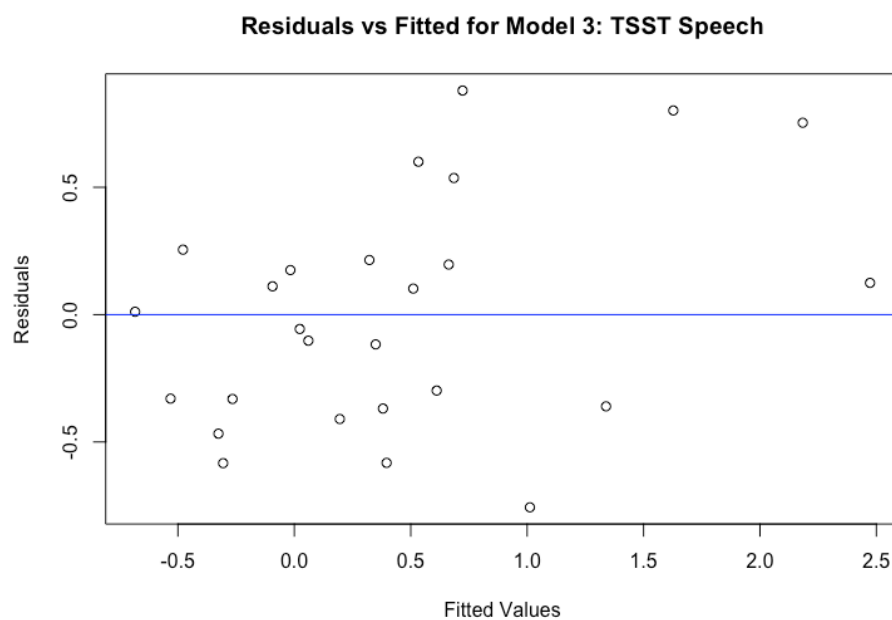
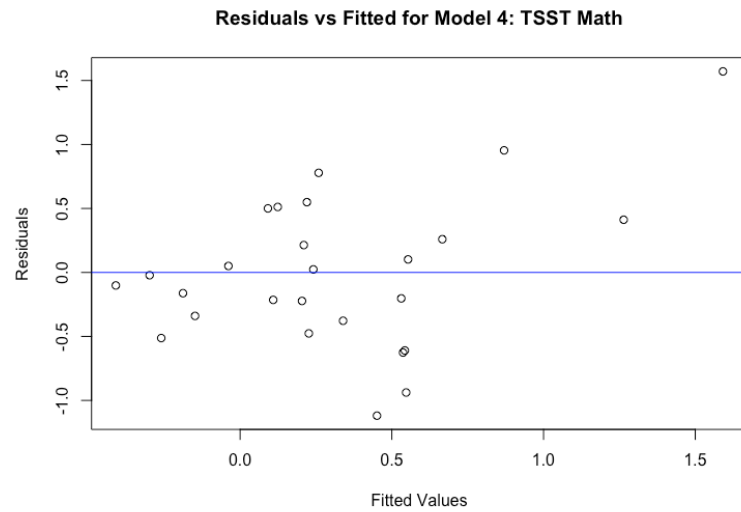


Figure 4

Residual vs Fitted Plot for Model 4: HRV during “Math” task of TSST

**Figure 5**

Residual vs Fitted Plot for Model 5: HR during TSST

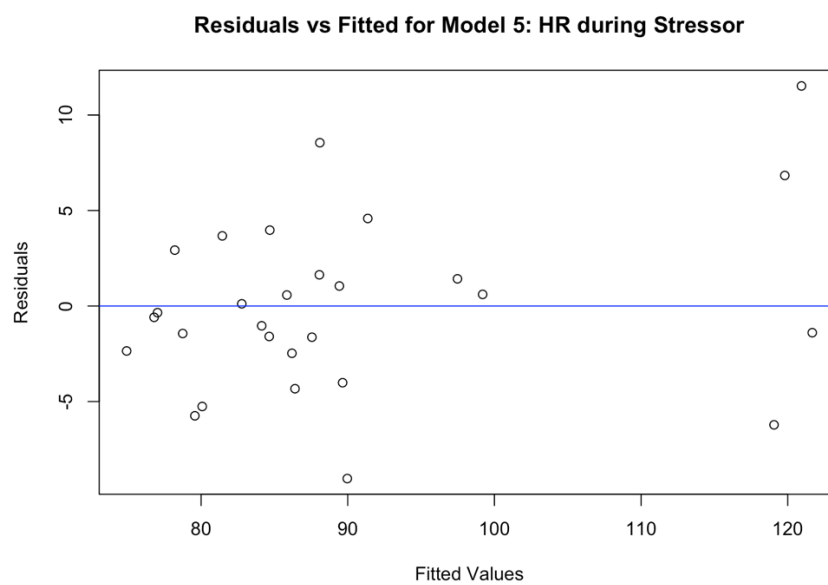
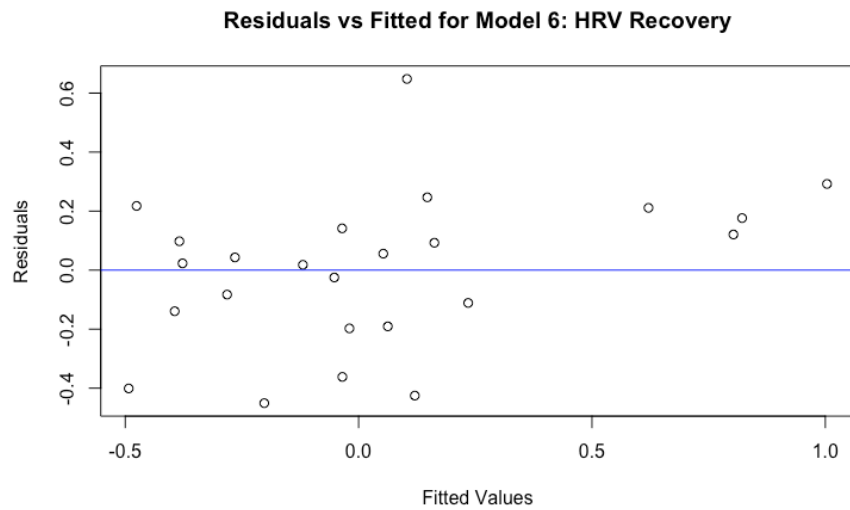


Figure 6

Residual vs Fitted Plot for Model 6: HRV during Recovery

**Figure 7**

Residual vs Fitted Plot for Model 7: Negative Affect

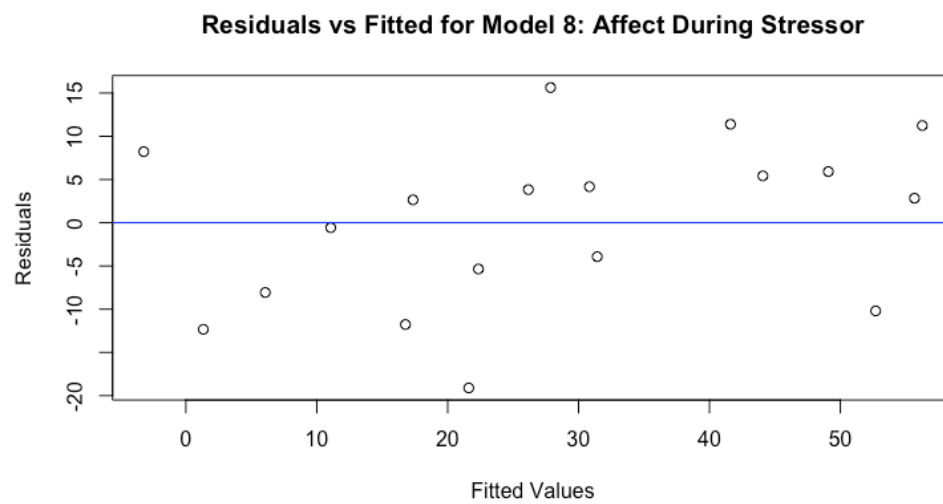
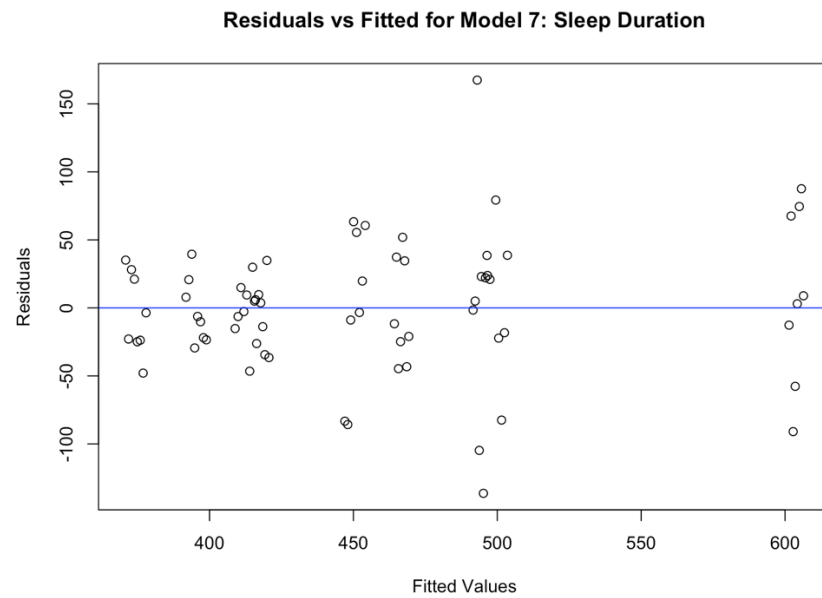


Figure 8

Residual vs Fitted Plot for Model 8: Change in Sleep



Appendix D

Feasibility and Acceptability Questionnaire

Question	Response Option
Did you like or dislike this biofeedback intervention?	1. Strongly dislike 2. Dislike 3. No opinion 4. Like 5. Strongly like
How comfortable did you feel engaging in the biofeedback intervention?	1. Very uncomfortable 2. Uncomfortable 3. No opinion 4. Comfortable 5. Very comfortable
How much effort did it take to engage in the biofeedback intervention	1. No effort at all 2. A little effort 3. No opinion 4. A lot of effort 5. Huge effort
There are moral or ethical consequences to engaging in the intervention?	1. Strongly disagree 2. Disagree 3. No opinion 4. Agree 5. Strongly agree
The biofeedback program helped improve my sleep hygiene	1. Strongly disagree 2. Disagree 3. No opinion 4. Agree 5. Strongly agree
The biofeedback program helped improve my sleep quality	1. Strongly disagree 2. Disagree 3. No opinion 4. Agree 5. Strongly agree

The biofeedback program helped improve me feel more relaxed	1. Strongly disagree 2. Disagree 3. No opinion 4. Agree 5. Strongly agree
It is clear to me how integrating sleep hygiene and deep breathing exercises will improve my health, sleep and well-being	1. Strongly disagree 2. Disagree 3. No opinion 4. Agree 5. Strongly agree
Engaging in the biofeedback program interfered with my other priorities	1. Strongly disagree 2. Disagree 3. No opinion 4. Agree 5. Strongly agree
How confident did you feel engaging in sleep hygiene and tracking exercise?	1. Very unconfident 2. Unconfident 3. No opinion 4. Confident 5. Very confident
How confident did you feel about engaging in deep breathing exercises?	1. Very unconfident 2. Unconfident 3. No opinion 4. Confident 5. Very confident
How acceptable was the intervention to you?	1. Completely unacceptable 2. Unacceptable 3. No opinion 4. Acceptable 5. Completely acceptable

Appendix F
Study Timeline

	Visit 1	Intervention	Visit 2	Visit 3
	Week 0	Week 1-8	Week 4	Week 9
Fitbit Set Up	X			
Sociodemographic Questionnaires	X			
Pre-Questionnaires	X			
ECG Measures	X		X	X
TSST	X			X
VAS	X			X
Resting Task	X		X	X
Intervention Materials Delivered	X			
Post-Questionnaires				X
Fitbit Refresher			X	
Weekly Questionnaires		X		

Note. ECG = electrocardiogram measures TSST = Trier Social Stress Test VAS = visual analogue scale used to assess psychological distress during the TSST