Evaluate Closed-Loop, Mindless Intervention in-the-Wild: A Micro-Randomized Trial on Offset Heart Rate Biofeedback

Yiran Zhao
yz2647@cornell.edu
Cornell University
New York, NY, USA

Tanzeem Choudhury
tanzeem.choudhury@cornell.edu
Cornell University
New York, NY, USA

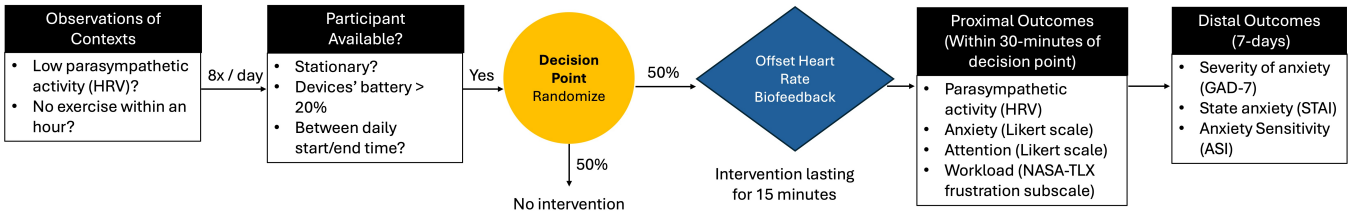

Figure 1: Visual representation of the micro-randomized trial on offset heart rate biofeedback.

ABSTRACT

The vision of closed-loop intervention systems for behavioral health is growing with the flourishing of mobile sensors and multimodal data. There has been abundant work on identifying symptoms, diagnosis, and progression monitoring. However, there has been limited effort in intervention research, tailoring suitable interventions for closed-loop systems. About a decade ago, researchers began exploring **mindless interventions**—subtle interventions to change behavior, cognition, or affect with minimal attention and effort. Despite their success in controlled laboratory settings, few mindless interventions have been deployed in the real world, and none have been integrated into closed-loop systems. Thus, it remains unclear how well these low-effort, low-attention interventions integrate with sensing systems, how their effectiveness varies over time and context, and their overall impact on behavioral health management.

This study is the first to deploy mindless interventions in a closed-loop system in real-world settings. We developed a closed-loop intervention for individuals with moderate to severe anxiety, delivering offset heart rate biofeedback when stress symptoms are detected. This paper presents our work-in-progress, detailing the system and study design, and highlighting this research's methodological and empirical contributions.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in ubiquitous and mobile computing**; **Ubiquitous and mobile computing design and evaluation methods**; • **Applied computing** → **Consumer health**; **Psychology**.

UbiComp Companion '24, October 5–9, 2024, Melbourne, VIC, Australia.

ACM ISBN 979-8-4007-1058-2/24/10
https://doi.org/10.1145/3675094.3678373

KEYWORDS

Mindless Computing; Mindless Intervention; Ubiquitous and Mobile Computing; Wearable Computing; Health and Behavior Change Technology; Behavioral Health; Micro-Randomized Trial; Stress; Anxiety; General Anxiety Disorder; Offset Heart Rate Biofeedback; Closed-Loop System; Closed-Loop Intervention

ACM Reference Format:
Yiran Zhao and Tanzeem Choudhury. 2024. Evaluate Closed-Loop, Mindless Intervention in-the-Wild: A Micro-Randomized Trial on Offset Heart Rate Biofeedback. In *Companion of the 2024 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp Companion '24), October 5–9, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3675094.3678373

1 PROBLEM STATEMENT AND RELATED WORK

The need for effective interventions in behavioral health—the domain focused on mental and substance use disorders—is urgent [37]. In the US, approximately one in five adults live with mental health conditions [35]. Mobile health (mHealth) interventions have the potential to revolutionize behavioral health care through the use of mobile sensing, machine and deep learning, and personalized, context-aware interventions [23]. Closed-loop intervention systems exemplify the advantages of mHealth by providing a system that identifies when a user needs intervention and automatically delivers the most suitable one.

Despite large investments in sensing systems, the impact of closed-loop interventions has been limited. This limitation arises from the mismatch between the capability of the sensing systems and the nature of existing mobile interventions. Most current interventions, even when paired with advanced sensing systems, are digital versions of traditional methods. These interventions, such as mobile therapy [36], behavior change messages [5], and mindfulness exercises [43], demand high levels of attention and effort from users, relying heavily on their motivation to be effective. While effective for various behavioral health conditions [6, 11, 47], they are unsuitable as immediate countermeasures to acute health

needs detected by the sensing systems. Such acute needs (e.g., when the users experience elevated anxiety) often arise when users are preoccupied, unmotivated to cope, or fall to maladaptive coping strategies.

Researchers have recently investigated mindless interventions that require minimal attention and effort to be effective. This notion was first introduced by Adams and Costa et al. in 2015, described as interventions that "subtly integrate themselves into the daily lives of users by **influencing users' behavior**, **requiring little effort and attention** from them" [2]. Since then, many interventions have emerged, translating neural mechanisms into wearable devices [9, 10, 46, 50, 51] and subtly guiding slow-breathing [7, 8]. However, these efforts have predominantly remained in controlled evaluations. The few tested in-the-wild did not couple with a sensing system, requiring the users to be motivated to manually control the intervention [7, 46].

Evaluating mindless interventions in-the-wild is crucial for realizing the potential of closed-loop intervention systems. Without deployment on free-living individuals, it is impossible to determine whether the laboratory effects are reliably repeatable or are merely experimental tricks. Additionally, there are concerns about whether these subtle interventions are perceivable and effective in the noisy real world, and whether their effects persist over time. These questions can only be answered through real-world deployments.

To evaluate mindless interventions in-the-wild, we borrow methods from the most widely-studied type of closed-loop system in behavioral health: just-in-time adaptive intervention (JITAI) [30]. The current state-of-the-art to study JITAI systems is through micro-randomized trials (MRT) [12, 33, 38]. MRT evaluates different intervention components (e.g., different interventions in different contexts, moderating variables) and time-varying factors (e.g., how does the effectiveness of an intervention change over time). Several closed-loop systems have adapted MRT to understand what intervention components are the most effective in each context [3, 4, 21, 29, 31].

This work is the first to evaluate closed-loop, mindless interventions in-the-wild. Specifically, we assess offset heart rate biofeedback, a well-validated mindless intervention delivered as a subtle wrist vibration. Previous research has shown that offset heart rate biofeedback can regulate the parasympathetic nervous system and mitigate anxiety [9, 10]. We developed a proof-of-concept closed-loop system that detects when the participants experience low parasympathetic activity and automatically delivers offset heart rate biofeedback. The study is structured as an MRT, evaluating the effect of offset heart rate biofeedback on parasympathetic activity and anxiety while understanding how intervention components (e.g., time, contexts, attention, workload) impact the intervention's strength. Our contributions are twofold. First, **empirical contribution**: we provide the first evidence of mindless interventions' effectiveness in the wild. Second, **methodological contribution**: we offer a set of considerations for intervention researchers in ubiquitous computing to help translate interventions from laboratory settings to real-world use and explore if MRT is suitable for evaluating mindless interventions.

2 METHODOLOGY

2.1 Intervention Mechanism

Offset heart rate biofeedback was first developed in 2016 [9]. Offset heart rate biofeedback is a subtle vibration on the wrist with a frequency 30% slower than the momentary heart rate. Controlled lab experiments showed that receiving offset heart rate biofeedback during high-intensity events (e.g., Trier Social Stress Test [9], mathematical tests [10], and alcohol craving [50]) leads to stabilized parasympathetic activity, decreased self-reported anxiety and decreased cravings. This effect is placebo-tested, because a placebo vibration of 60 bpm or vibration at the heart rate frequency did not achieve similar effects [9]. These prior works indicated that offset heart rate biofeedback is a reliable, placebo-validated mechanism to mitigate autonomic nervous activity in-the-moment.

Offset heart rate biofeedback can be delivered via custom hardware [9] or an Apple Watch [10]. This study utilized Apple Watches to deliver offset heart rate biofeedback for scalability and repeatability. The vibration we chose is the "click" haptic in the Haptic Engine of the Apple Watch, consistent with prior laboratory studies.

2.2 Closed-Loop System Design

The closed-loop system consists of an iPhone, an Apple Watch, and a Polar H10 heart rate monitor [40]. The Polar H10 senses beat-to-beat intervals (IBI) and transmits IBI to the iPhone application at approximately 1Hz. The Apple Watch measures activity and delivers offset heart rate biofeedback. The iPhone application collects geolocation, delivers ecological momentary assessments (EMA), and determines when to start and stop the intervention (Figure 2).

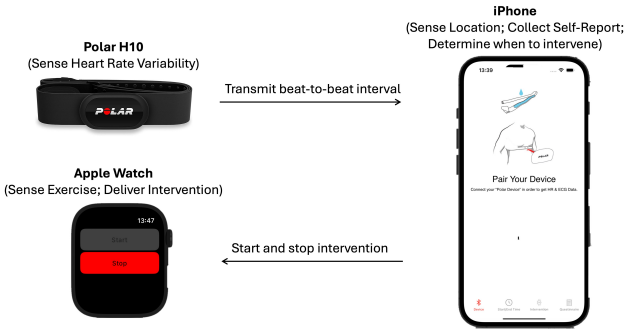

Figure 2: Closed-loop system. The system consists of an iPhone, an Apple Watch, and a Polar H10 heart rate monitor.

We describe the system design following the convention of MRT (Figure 1):

2.2.1 Observations of Contexts. Ideally, the system should detect when the user is stressed in-the-moment. However, despite years of effort in stress detection, sensing stress in real-time in free-living individuals remains a challenge. First, most stress detection algorithms are based on data collected in-lab. Lab-simulated scenarios do not fully represent the complexity of the real world, and algorithms built on these datasets are rarely evaluated in-the-wild. Second, it is impossible to confidently predict stress in free-living

individuals using only physiological sensors [28]. Physiological signals such as heart rate variability (HRV) are indicators of autonomic nervous activity, resulting from "a complex and non-linear system best described by mathematical chaos" [39]. Various factors beyond stress and anxiety influence these signals. Adding context sensing increases hardware requirements and privacy concerns, affecting study adherence.

Given the current state of the art, our solution to balancing the validity of the sensing system and study feasibility is to detect the **decreased in parasympathetic activity**. Although a decrease in parasympathetic activity can be caused by both physiological factors (e.g., exercise [27], caffeine [15]) and psychological states (e.g., alertness [26]), it is still a valid indicator of psychological stress [16], as long as the individual is not under physical load. Parasympathetic activity can be reliably detected by HRV metrics such as the root mean square of successive differences (RMSSD) [18]. Detecting HRV requires only one additional device, a heart rate monitoring band, which benefits study adherence. Furthermore, detecting parasympathetic activity aligns with the intervention mechanism: offset heart rate biofeedback entrains the body's autonomic system and stabilizes parasympathetic activity against stressors.

To ensure the sensed decrease in parasympathetic activity is not caused by physiological load, the system also observes when the user exercises using the Activity feature in the Apple Watch. If the user experiences reduced parasympathetic activity and has not exercised within the recent hour, the system proceeds to the next-step logic, checking if the user is available.

2.2.2 User Availability. The system determines if the user is available using two criteria: (1) whether the user is stationary (because offset heart rate biofeedback has only been evaluated on stationary individuals), (2) whether the Apple Watch and the iPhone have more than 20% battery.

2.2.3 Tailoring Variables. The only tailoring variable in this study is the user's momentary heart rate because offset heart rate biofeedback dynamically adjusts the biofeedback frequency based on the user's momentary heart rate [9].

2.2.4 Randomization and Intervention Delivery. If the user experiences a decrease in parasympathetic activity and is available, this moment is referred to as the **decision point**. There is a 50% likelihood that the user will receive the intervention. Once activated, the intervention will be on for 15 minutes, consistent with prior work [9, 10, 50].

2.2.5 Proximal and Distal Outcomes. The proximal outcomes are the parasympathetic activity and self-reported anxiety 30 minutes after the decision points. The distal outcome is the self-reported anxiety and anxiety sensitivity post-study. The moderating variables are how much attention the participants allocate to the intervention, how much workload the intervention requires, and what contexts the participant is in. To quantify these outcomes and moderating variables, we use measurements detailed in Section 3.3.

2.3 Research Questions

It's important to note that MRTs are not "confirmatory studies designed to evaluate an intervention package; instead, they are focused on selecting and optimizing intervention components" [12]. With the independent variable as whether the participants received the intervention, we answer two types of research questions: one primary, hypothesis-testing research question that determines the statistical power and a set of exploratory analyses [33]. We detail our research questions in the format following the field convention, referring to offset heart rate biofeedback as "the intervention":

(1) Primary Question:
   (a) Is there an effect of the intervention on parasympathetic activity? On average, across time, does receiving the intervention lead to a higher parasympathetic activity 30 minutes after the decision point?
   (b) If so, does the effect deteriorate with time (day in study)?
(2) Secondary Questions:
   (a) Is there an effect of the intervention on anxiety 30 minutes after the decision point and on daily anxiety? If so, does the effect deteriorate with time (day in study)?
   (b) Is there an effect of the intervention on anxiety and anxiety sensitivity before and after the study?
   (c) Is the intervention more effective when the participant is out of home or at home?
   (d) Is the intervention more effective when the participant is less aware of its sensation?
   (e) Is the intervention more effective when the participant experiences less frustration toward its sensation?
   (f) Does the effectiveness of the intervention decrease as the participants receive more of it within a day?

3 EVALUATION

3.1 Study Design

The study will recruit 32 adult participants with moderate to severe anxiety, assessed by the Generalized Anxiety Disorder 7-item (GAD-7) [42]. Participants will recruited throughout the United States. All researcher-participant interaction will happen remotely, including device sendoff, study onboarding, and interviews. This study is reviewed by the Cornell University Institutional Review Board.

The study has three stages: (1) one-hour onboarding, (2) seven-day deployment, and (3) one and half hour offboarding. During the onboarding, we will collect pre-study self-reports (Section 3.3), conduct interviews, and collect baseline parasympathetic activity.

During the seven-day deployment, participants will indicate a daily study start and end time that fits their routine (e.g., start at 9 am and end at 10 pm daily, which would be a 13-hour duration). The system will collect data and intervene during this duration. At the start time, participants will receive a notification to wear the Polar H10 device and their Apple Watch. Participants can go through their normal daily routine while wearing the closed-loop system. Once the participant puts on the heart rate monitor, the system will classify every 30-minute window to identify whether the participant is in a low parasympathetic activity state. This window will be randomly allocated to treatment (micro-randomization) at a 50% likelihood (i.e., intervention or no intervention). There will be, at maximum, eight decision points over the day (i.e., on each day, the participant will receive 0-8 interventions), such that the number of decision points has sufficient power for statistical analysis while not burdening the participants with excessive EMA [48]. After each

decision point, the participants will receive a set of EMA, probing for anxiety (for all decision points), attention, and workload associated with the intervention (for decision points with intervention). At the end of each day and before the study's end time, the participants will receive a set of daily questionnaires asking about their anxiety. Participants can take off the devices at the end of the day after the study end time.

During offboarding, we will collect the participants' post-study self-reports and conduct interviews.

3.2 Population

We plan to conduct experiments with 32 participants to achieve a 0.8 desired power (over 7 days, with 8 decision points per day, 0.5 randomization probability, 0.2 average proximal effect) [24]. Participants are eligible if they are between 18-65 years old. Participants are eligible if they score 10 or higher in GAD-7, which has a 0.82 specificity and 0.89 sensitivity in diagnosing GAD [22]. Other inclusion criteria include residing in the United States, fluency in English, and having consistent internet access. Participants are excluded if they (1) have a current or previous diagnosis of bipolar disorder, schizophrenia, or psychosis, (2) are at risk of self-harm or suicide, and (3) are dependent on any substances (including but not limited to opioid, nicotine, and alcohol).

3.3 Measures

3.3.1 Parasympathetic Activity. The parasympathetic activity is measured by HRV. For every 5-minute IBI, the system computes a root mean square deviation (RMSSD), which is highly correlated with the activity of the parasympathetic branch while not affected by breathing changes [44]. The 5-minute window is the field convention to obtain physiologically meaningful calculations [39]. To capture the temporal dynamics of heart rate, we apply a sliding window with a 1-minute increment, recommended for continuous monitoring [49]. The first RMSSD will be computed once the participant wears the heart rate monitor for the first 5 minutes of the day. Then, a new RMSSD will be calculated with every 1 minute additional data.

RMSSD will be collected throughout the day as the observations of contexts and as proximal outcomes. The system makes a decision on parasympathetic activity using 30-minute of RMSSD because anticipatory stress is detectable in RMSSD 30-60 minutes before the event [25], while stress due to fatigue, burnout, and emotional exhaustion is detectable in the window with a minimal 20 minutes [17, 34]. As proximal outcomes, we take the 30-minute window immediately after the decision point and use this window to determine the effect of the intervention on parasympathetic activity.

3.3.2 Anxiety. We collect self-report anxiety in three ways: after each decision point, at the end of each day, and pre/post-study. At 30 minutes after each decision point, we collect a single-item Likert scale self-report "To what extent are you experiencing anxiety right now?" rated from 0 (no anxiety) and 10 (extreme anxiety) [19, 45]. At the end of each study day, we collect the daily anxiety level via the state version of the State-Trait Anxiety Inventory (STAI). We adjust the language in the original questionnaire to "Throughout today, I feel..." [41]. During the onboarding and offboarding session, we collect STAI, General Anxiety Disorder-7 Item (GAD-7) [42], and Anxiety Sensitivity Index (ASI) [32] to probe the pre-/post-study differences.

3.3.3 Attention and Workload. We collect attention and workload to understand how obtrusive the intervention is during the study. HCI and behavioral science have many ways to measure attention, including visual attention, vigilance, or task performance [46]. We are interested in how unaware participants are of the intervention. At 30 minutes after the decision point, if the participant has received the intervention, we collect a single-item Likert scale question of "How much are you aware of the sensation on your wrist?" rated from 0 (unaware of the sensation) to 100 (extremely aware). We access workload by adapting the frustration subscale in NASA-TLX [14] "How irritated, stressed, and annoyed were you when the sensation was on?" from 0 (not at all) to 100 (extremely).

3.3.4 Context. We infer context through geolocation using GPS. To be considerate of power, we collect one set of longitude and latitude at the end of the 30-minute processing window of RMSSD. We will first label the routine locations for each participant's geolocation, including home, school, and work. Then, we will label the rest of the points of interest using the Google Places API [13].

3.3.5 Interview. We will interview the participants pre and post-study. Both interviews will be conducted in a semi-structured manner. During the onboarding interview, we will probe for (1) how participants manage their anxiety before the study, (2) what challenges they face in managing their anxiety. During the offboarding interview, we will probe how participants perceive the intervention's usefulness, effectiveness, and annoyance during true and false positives of stress detection. In addition, we will probe how having the closed-loop system impacts how participants manage their anxiety and ask the participants to design their ideal closed-loop systems.

4 CONTRIBUTION

This study pioneers the deployment of mindless interventions in-the-wild in a closed-loop system. Adapting the MRT study design used by JITAI systems, this study makes twofold contributions. The first is an **empirical contribution** to behavioral health researchers. The results will reveal the effectiveness of mindless interventions in real-world settings, potentially opening new approaches for intervention in intense contexts such as high-stress scenarios and substance cravings. The second is a **methodological contribution**. Ubiquitous computing and HCI researchers have long struggled to carry out research with longevity – many ideas were evaluated in controlled settings, published, and abandoned. Despite calls for change a decade ago [1, 20], transitioning lab-evaluated interventions to the real world remains challenging. This work details a list of considerations for intervention researchers to conduct early work with translation in mind. It also explores the suitability of MRT for mindless interventions and provides practical guidelines to conduct efficacy evaluation in the wild.

ACKNOWLEDGMENTS

We thank National Science Foundation for supporting our study with Award No.2212352.

REFERENCES

[1] Gregory D Abowd. 2012. What next, ubicomp? Celebrating an intellectual disappearing act. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. 31–40.

[2] Alexander T Adams, Jean Costa, Malte F Jung, and Tanzeem Choudhury. 2015. Mindless computing: designing technologies to subtly influence behavior. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. 719–730.

[3] Samuel L Battalio, David E Conroy, Walter Dempsey, Peng Liao, Marianne Menictas, Susan Murphy, Inbal Nahum-Shani, Tianchen Qian, Santosh Kumar, and Bonnie Spring. 2021. Sense2Stop: a micro-randomized trial using wearable sensors to optimize a just-in-time-adaptive stress management intervention for smoking relapse prevention. *Contemporary Clinical Trials* 109 (2021), 106534.

[4] Lauren Bell, Claire Garnett, Yihan Bao, Zhaoxi Cheng, Tianchen Qian, Olga Perski, Henry WW Potts, Elizabeth Williamson, et al. 2023. How notifications affect engagement with a behavior change app: Results from a micro-randomized trial. *JMIR mHealth and uHealth* 11, 1 (2023), e38342.

[5] Sofian Berrouiguet, Enrique Baca-García, Sara Brandt, Michel Walter, and Philippe Courtet. 2016. Fundamentals for future mobile-health (mHealth): a systematic review of mobile phone and web-based text messaging in mental health. *Journal of medical Internet research* 18, 6 (2016), e135.

[6] Laura Carrasco-Hernandez, Francisco Jódar-Sánchez, Francisco Núñez-Benjumea, Jesús Moreno Conde, Marco Mesa González, Antón Civit-Balcells, Santiago Hors-Fraile, Carlos Luis Parra-Calderón, Panagiotis D Bamidis, Francisco Ortega-Ruiz, et al. 2020. A mobile health solution complementing psychopharmacology-supported smoking cessation: randomized controlled trial. *JMIR mHealth and uHealth* 8, 4 (2020), e17530.

[7] Kyung Yun Choi, Neska ElHaouij, Jinmo Lee, Rosalind W Picard, and Hiroshi Ishii. 2022. Design and evaluation of a clippable and personalizable pneumatic-haptic feedback device for breathing guidance. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–36.

[8] Kyung Yun Choi, Jinmo Lee, Neska ElHaouij, Rosalind Picard, and Hiroshi Ishii. 2021. Aspire: clippable, mobile pneumatic-haptic device for breathing rate regulation via personalizable tactile feedback. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–8.

[9] Jean Costa, Alexander T Adams, Malte F Jung, François Guimbretière, and Tanzeem Choudhury. 2016. EmotionCheck: leveraging bodily signals and false feedback to regulate our emotions. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*. 758–769.

[10] Jean Costa, François Guimbretière, Malte F Jung, and Tanzeem Choudhury. 2019. Boostmeup: Improving cognitive performance in the moment by unobtrusively regulating emotions with a smartwatch. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 1–23.

[11] Sandra Garrido, Chris Millington, Daniel Cheers, Katherine Boydell, Emery Schubert, Tanya Meade, and Quang Vinh Nguyen. 2019. What works and what doesn't work? A systematic review of digital mental health interventions for depression and anxiety in young people. *Frontiers in psychiatry* 10 (2019), 759.

[12] Jessica R Golbus, Walter Dempsey, Elizabeth A Jackson, Brahmajee K Nallamothu, and Predrag Klasnja. 2021. Microrandomized trial design for evaluating just-in-time adaptive interventions through mobile health technologies for cardiovascular disease. *Circulation: Cardiovascular Quality and Outcomes* 14, 2 (2021), e006760.

[13] Clay E Gregory. 2013. *Location-targeted measurement of perceived stress*. Ph. D. Dissertation. University of Illinois at Urbana-Champaign.

[14] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[15] Gaku Hibino, Toshio Moritani, Teruo Kawada, and Tohru Fushiki. 1997. Caffeine enhances modulation of parasympathetic nerve activity in humans: quantification using power spectral analysis. *The Journal of nutrition* 127, 7 (1997), 1422–1427.

[16] Sarah Immanuel, Meseret N Teferra, Mathias Baumert, and Niranjan Bidargaddi. 2023. Heart rate variability for evaluating psychological stress changes in healthy adults: a scoping review. *Neuropsychobiology* 82, 4 (2023), 187–202.

[17] Magdalena K Kanthak, Tobias Stalder, LaBarron K Hill, Julian F Thayer, Marlene Penz, and Clemens Kirschbaum. 2017. Autonomic dysregulation in burnout and depression: evidence for the central role of exhaustion. *Scandinavian journal of work, environment & health* 43, 5 (2017), 475.

[18] Hye-Geum Kim, Eun-Jin Cheon, Dai-Seg Bai, Young Hwan Lee, and Bon-Hoon Koo. 2018. Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry investigation* 15, 3 (2018), 235.

[19] Zachary D King, Judith Moskowitz, Begum Egilmez, Shibo Zhang, Lida Zhang, Michael Bass, John Rogers, Roozbeh Ghaffari, Laurie Wakschlag, and Nabil Alshurafa. 2019. Micro-stress EMA: A passive sensing framework for detecting in-the-wild stress in pregnant mothers. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 3, 3 (2019), 1–22.

[20] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. 2011. How to evaluate technologies for health behavior change in HCI research. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3063–3072.

[21] Andrea C Kramer, Andreas B Neubauer, and Florian Schmiedek. 2023. The effectiveness of a slow-paced diaphragmatic breathing exercise in children's daily life: A micro-randomized trial. *Journal of Clinical Child & Adolescent Psychology* 52, 6 (2023), 797–810.

[22] Kurt Kroenke, Robert L Spitzer, Janet BW Williams, Patrick O Monahan, and Bernd Löwe. 2007. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Annals of internal medicine* 146, 5 (2007), 317–325.

[23] Clemens Scott Kruse, Jose A Betancourt, Matthew Gonzales, Kennedy Dickerson, and Miah Neer. 2022. Leveraging mobile health to manage mental health/behavioral health disorders: systematic literature review. *JMIR Mental Health* 9, 12 (2022), e42301.

[24] Peng Liao, Predrag Klasnja, Ambuj Tewari, and Susan A Murphy. 2016. Sample size calculations for micro-randomized trials in mHealth. *Statistics in medicine* 35, 12 (2016), 1944–1971.

[25] Daniela Lucini, Guido Norbiato, Mario Clerici, and Massimo Pagani. 2002. Hemodynamic and autonomic adjustments to real life stress conditions in humans. *Hypertension* 39, 1 (2002), 184–188.

[26] Antonio Luque-Casado, José C Perales, David Cárdenas, and Daniel Sanabria. 2016. Heart rate variability and cognitive processing: The autonomic response to task demands. *Biological psychology* 113 (2016), 83–90.

[27] Scott Michael, Kenneth S Graham, and Glen M Davis. 2017. Cardiac autonomic responses during exercise and post-exercise recovery using heart rate variability and systolic time intervals—a review. *Frontiers in physiology* 8 (2017), 259883.

[28] Varun Mishra, Sougata Sen, Grace Chen, Tian Hao, Jeffrey Rogers, Ching-Hua Chen, and David Kotz. 2020. Evaluating the reproducibility of physiological stress detection models. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 4, 4 (2020), 1–29.

[29] Inbal Nahum-Shani, Zara M Greer, Anna L Trella, Kelly W Zhang, Stephanie M Carpenter, Dennis Ruenger, David Elashoff, Susan A Murphy, and Vivek Shetty. 2024. Optimizing an adaptive digital oral health intervention for promoting oral self-care behaviors: Micro-randomized trial protocol. *Contemporary Clinical Trials* 139 (2024), 107464.

[30] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. 2018. Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine* (2018), 1–17.

[31] Timothy NeCamp, Srijan Sen, Elena Frank, Maureen A Walton, Edward L Ionides, Yu Fang, Ambuj Tewari, and Zhenke Wu. 2020. Assessing real-time moderation for developing adaptive mobile health interventions for medical interns: micro-randomized trial. *Journal of medical Internet research* 22, 3 (2020), e15033.

[32] Rolf A Peterson and Robert L Heilbronner. 1987. The anxiety sensitivity index:: Construct validity and factor analytic structure. *Journal of anxiety disorders* 1, 2 (1987), 117–121.

[33] Tianchen Qian, Ashley E Walton, Linda M Collins, Predrag Klasnja, Stephanie T Lanza, Inbal Nahum-Shani, Mashfiqui Rabbi, Michael A Russell, Maureen A Walton, Hyesun Yoo, et al. 2022. The microrandomized trial for developing digital interventions: Experimental design and data analysis considerations. *Psychological methods* 27, 5 (2022), 874.

[34] Jakub Rajcani, Petra Solarikova, Daniela Turonova, Igor Brezina, and J Rajcáni. 2016. Heart rate variability in psychosocial stress: Comparison between laboratory and real-life setting. *Activitas Nervosa Superior Rediviva* 58, 3 (2016), 77–82.

[35] Parashar Ramanuj, Erin Ferenchik, Mary Docherty, Brigitta Spaeth-Rublee, and Harold Alan Pincus. 2019. Evolving models of integrated behavioral health and primary care. *Current psychiatry reports* 21 (2019), 1–12.

[36] Amy Leigh Rathbone, Laura Clarry, and Julie Prescott. 2017. Assessing the efficacy of mobile health apps using the basic principles of cognitive behavioral therapy: systematic review. *Journal of medical Internet research* 19, 11 (2017), e399.

[37] Deborah M Scharf, Nicole K Eberhart, Nicole Schmidt Hackbarth, Marcela Horvitz-Lennon, Robin Beckman, Bing Han, Susan L Lovejoy, Harold Alan Pincus, and M Audrey Burnam. 2014. Evaluation of the SAMHSA primary and behavioral health care integration (PBHCI) grant program. *Rand health quarterly* 4, 3 (2014).

[38] Nicholas J Seewald, Shawna N Smith, Andy Jinseok Lee, Predrag Klasnja, and Susan A Murphy. 2019. Practical considerations for data collection and management in mobile health micro-randomized trials. *Statistics in biosciences* 11 (2019), 355–370.

[39] Fred Shaffer and Jay P Ginsberg. 2017. An overview of heart rate variability metrics and norms. *Frontiers in public health* 5 (2017), 290215.

[40] Kathryn E Speer, Stuart Semple, Nenad Naumovski, and Andrew J McKune. 2020. Measuring heart rate variability using commercially available devices in healthy children: A validity and reliability study. *European Journal of Investigation in Health, Psychology and Education* 10, 1 (2020), 390–404.

[41] Charles Donald Spielberger. 1970. Manual for the State-Trait Anxiety Inventory (self-evaluation questionnaire). *(No Title)* (1970).

[42] Robert L Spitzer, Kurt Kroenke, Janet BW Williams, and Bernd Löwe. 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of*

internal medicine 166, 10 (2006), 1092–1097.

[43] Shufang Sun, Danhua Lin, Simon Goldberg, Zijiao Shen, Pujing Chen, Shan Qiao, Judson Brewer, Eric Loucks, and Don Operario. 2022. A mindfulness-based mobile health (mHealth) intervention among psychologically distressed university students in quarantine during the COVID-19 pandemic: A randomized controlled trial. *Journal of counseling psychology* 69, 2 (2022), 157.

[44] Bianca Lee Thomas, Nicolaas Claassen, Piet Becker, and Margaretha Viljoen. 2019. Validity of commonly used heart rate variability markers of autonomic nervous system function. *Neuropsychobiology* 78, 1 (2019), 14–26.

[45] Natasha Yan Chi Tung, Yang Yap, Bei Bei, Linda J Luecken, and Joshua F Wiley. 2022. A 14-day ecological momentary assessment study on whether resilience and early family risk moderate daily stress and affect on cortisol diurnal slope. *Scientific reports* 12, 1 (2022), 1240.

[46] Nathan W Whitmore, Samantha Chan, Jingru Zhang, Patrick Chwalek, Sam Chin, and Pattie Maes. 2024. Improving attention using wearables via haptic and multimodal rhythmic stimuli. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–14.

[47] Cassandra Wright, Paul M Dietze, Paul A Agius, Emmanuel Kuntsche, Michael Livingston, Oliver C Black, Robin Room, Margaret Hellard, Megan SC Lim, et al. 2018. Mobile phone-based ecological momentary intervention to reduce young adults' alcohol use in the event: A three-armed randomized controlled trial. *JMIR mHealth and uHealth* 6, 7 (2018), e9324.

[48] Cornelia Wrzus and Andreas B Neubauer. 2023. Ecological momentary assessment: A meta-analysis on designs, samples, and compliance across research fields. *Assessment* 30, 3 (2023), 825–846.

[49] Haoshi Zhang, Mingxing Zhu, Yue Zheng, and Guanglin Li. 2015. Toward capturing momentary changes of heart rate variability by a dynamic analysis method. *PLoS One* 10, 7 (2015), e0133148.

[50] Yiran Zhao, Jatin Arora, Yujie Tao, David Miller, Alexander Adams, and Tanzeem Choudhury. [n. d.]. Translation Effectiveness of Offset Heart Rate Biofeedback as Passive Intervention for Alcohol Craving: Pilot Controlled Experiment. *JMIR Preprints* ([n. d.]). https://doi.org/10.2196/preprints.54438 54438.

[51] Yiran Zhao, Yujie Tao, Grace Le, Rui Maki, Alexander Adams, Pedro Lopes, and Tanzeem Choudhury. 2023. Affective Touch as Immediate and Passive Wearable Intervention. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (2023), 1–23.