Incorporating wearable activity tracking into prolonged exposure treatment for PTSD

A Thesis
Presented to
The Academic Faculty

by

Kenneth William Akers

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Science in the
College of Computing

Georgia Institute of Technology

Approved by:

Rosa I. Arriaga, Ph.D. Advisor
School of Interactive Computing
Georgia Institute of Technology

DocuSigned by:
Rosa Arriaga
ABA8C26128D24BD...

Peter Presti
Interactive Media Technology Center
Georgia Institute of Technology

DocuSigned by:
Peter Presti
238BF2740ACE4DB...

Date Approved: 12/17/2021 | 7:46 AM EST

TABLE OF CONTENTS

INTRODUCTION

The rising prevalence of wearable technologies in society has catalyzed their use in clinical settings. This has run the gamut from treatment programs for weight management [3,12] to mental health. Within the last decade, an abundance of research has investigated incorporating wearables into clinical therapy programs for patients with schizophrenia, depression, and other chronic mental health conditions [37,38]. Traditionally, clinical evaluations only provide a short glimpse into patients' progress. They require highly-trained professionals, can be time-costly, and still provide sparse data points throughout a multi-week therapy program [36]. The rapidly evolving nature of this area of human-computer interaction is thus no wonder: these relatively unobtrusive tools grant clinicians and care teams an unprecedented level of insight into their patients' behavior when outside of their office.

The improved insight provided by wearable technology increases continuity of care and provides the clinicians with higher resolution, ubiquitous data of their patients' treatment trajectory. Crucially, such passive sensors - sensors that collect data with minimal human interaction - do so with little burden to the patient. This is supported by previous research, which has shown a general willingness among patients to use wearables in mental health therapy [11,26,29]. Notably, a major challenge facing this application of wearables has been patient understanding, privacy, and usability [26]. As such, a truly effective implementation of clinical wearables for mental health therapy must prioritize patient agency over their data, as well as education of how physiological tracking benefits treatment [9,29].

This research project will extend from the ongoing work conducted on the Prolonged Exposure Collective Sensing System for PTSD (PECSS) team at Georgia Tech. Existing mobile health (mHealth) and ubiquitous computing systems for mental health treatment have incorporated a patchwork of sensors to give clinicians insight into the treatment trajectory and status of their patients during intensive outpatient programs (IOPs) for post-traumatic stress disorder (PTSD). Specifically, these IOPs involve patients undergoing prolonged exposure (PE) therapy, in which clinicians guide patients through exercises in the clinical settings, as well as real-world environments [16]. PECSS builds on this by creating a comprehensive system of sensors and apps collecting patient-generated data (PGD), including metrics like location data, subjective units of distress (SUDs), app usage, SMS activity, and audio data. This project aims to introduce to this digital ecosystem the use of physiological data, namely heart rate collected by provided Fitbits, to address the broader HCI question: how can clinicians and researchers effectively incorporate wearables into PE therapy for PTSD? Exploring this question involves answering, is the PECSS system usable and learnable? Encouraged by prior research [9,29], this study will investigate how users interact with the mobile app and wearable components of the system. It will design a protocol for educating patients regarding the use, applicability, and privacy of their data. Data collected from participants will then be analyzed to evaluate the capabilities of the system to provide insight into participation in mock exposures.

LITERATURE REVIEW

Post-traumatic stress disorder (PTSD) is a trauma- and stressor-related disorder characterized by re-experiencing of trauma (e.g., intrusive memories and nightmares); avoidance of trauma-related situations, thoughts, and feelings; negative alterations in thoughts and mood; and hyperarousal [39]. PTSD burdens 6% of U.S. citizens and disproportionately 16% of U.S. veterans [35]. Over the past few decades, PE therapy has become one of the leading treatments for the condition. The strategy is a specific type of cognitive behavioral therapy that, through *imaginal* and *in-vivo* sessions, teaches individuals to gradually approach these intrusive memories, feelings, and situations [17]. Despite the prevalence of research supporting its efficacy [32,33], patients frequently drop out of treatment [18,19,21]. Thus, there has been a recent push in the areas of human-computer interaction and ubiquitous computing for mHealth solutions to augment IOPs and improve patient engagement.

Significant research involving mHealth technologies has been conducted within other vulnerable populations, namely schizophrenia patients. As early as 2016, a study by Wang et al. [36] proposed a mobile sensing system that collected data regarding physical activity, location, environmental setting, digital communication, and phone usage. Using machine learning models to predict the risk of psychotic relapse for patients in an intensive schizophrenia clinic, the study demonstrated significant predictive power from passively sensed metrics as compared to traditional assessments. Though foundational for later research on predictive mHealth technology in a clinical setting, this study's findings are hardly extrapolatable to other mental health conditions like PTSD. To yield more relevant insight regarding the social implications of PTSD, a

subsequent study [38] applied the same sensing system to explore the viability of assessing social functioning in schizophrenic patients. Critically, the study's results suggested that quantitative sensing from mobile phones alone could offer insight into behavioral and social health, and thus apply to PTSD treatment. However, in both of these studies, physiological data was not collected.

While the tracked metrics of mobile phones have been shown to provide insights into patient treatment trajectories for schizophrenia, the predictive power of physiological metrics for PTSD is yet to be empirically explored in a clinical setting. Results from Ernala et al. [15] suggest that physiological data, such as heart rate, may serve as proxy diagnostic signals for mental health conditions typically evaluated using PHQ-8, PHQ-4, and PCL-5. Specifically for PTSD, heart rate variability has been shown [34] to be an indicator of autonomic nervous system dysregulation, a key marker of PTSD. Still, despite these promising results for the clinical insight offered by these metrics, the incorporation of both mobile sensing and passively collected physiological data is yet to be studied.

In recent years, wearable technology has grown in popularity among researchers for unobtrusive, scalable applications in mobile health. Furthermore, Fitbits appear to be sufficiently accurate for heart rate (MAE = 5.96% [27]), and especially suitable for applications where scalability and low participant burden are central to the population or research question [27,28]. Research into veteran perspectives on the use of wearable activity trackers in IOPs for PTSD further motivates the introduction of Fitbits into therapy. A study by Ng et al. [29] interviewed veterans with PTSD immediately following a 3-week intensive IOP for PTSD, observing that veterans cite three principal reasons supporting wearable activity tracker use in treatment: they increase

self-awareness, support social interactions, and facilitate giving back to other veterans. However, unaddressed areas in this study's methodology constitute an opportunity for improvements in Fitbit use in IOP treatment. Chief among these include structured training and education on the function and rationale behind Fitbit use in mental health treatment, the absence of which is a major motivation for nonadherence to continued Fitbit use. As per Powers et al.'s [31] results demonstrating the positive impact of exercise augmentation in PE therapy, it may follow that educating patients on the link between physical and mental health for PTSD treatment could increase adherence.

With this body of research suggesting the multitude of benefits of activity tracking and physiological augmentation of mHealth systems for PTSD, it is crucial that such implementations achieve these benefits while preserving patient agency and privacy. As such, following closely behind this rise in prevalence of mHealth solutions have been growing ethical and privacy concerns. Research by Paul et al. [30] suggests that the use of these relatively novel wearable devices in clinical settings introduces seldom-addressed HIPAA concerns. As subsequent research [9,23] has identified, visualizing PGD is paramount to ensuring patients understand their metrics as it pertains to their treatment. Such research indicates a need for transparency and security in patient data collection, storage, and ownership.

The current study will combine the findings of previous research on the utility of heart rate using activity tracking wearables to assess the usability of this sensor suite in simulated PE therapy. To do this, the study will explore user perceptions and attitudes about the incorporation of Fitbit activity tracking into the PECSS sensor suite. Participants will complete a collection of simulated exposures while using the PECSS system. Data collected from the PECSS mobile app and the Fitbit will be used to

evaluate the ability of these data streams to provide insight into participant performance during such exposures. Participant feedback on the usability of the system will be combined with feedback on data privacy concerns to maximize integration adherence and promote data privacy for real PTSD patients in PECSS's subsequent deployment. In doing so, this study aims to demonstrate how wearable PGD can most effectively be introduced to PE therapy such that patient engagement is maximized and clinicians are better equipped to design more effective treatment, ultimately increasing the efficacy of PE therapy for PTSD.

METHODS

OVERVIEW

The aforementioned usability testing has three principle aims:

1. Collect heart rate data from participants in a variety of physically or mentally stimulating scenarios using Fitbits.
2. Using this data and ground truth data from the PECSS mobile app, evaluate the efficacy of Fitbit integration during the simulated "PE therapy" (i.e., in-vivo and imaginal sessions).
3. Ensure the use of Fitbit in the simulated PECSS meets users' needs.

The following sections detail the experiment's participant pool, materials, trial design, and procedure.

PARTICIPANTS

We recruited a convenience sample of nine participants (male = 6, female = 3; median age = 21) for this study. The sample consisted of Georgia Tech undergraduate students working on the PECSS research project. Each recruited participant was randomly assigned to one of four scenario groups, resulting in two participants per group except for group B, which contained three participants. Participants were provided with a Fitbit Inspire HR to facilitate passive heart rate tracking. These groups and their trial sequence are described in *Table 1* below.

Table 1: Trial Timeline

Group	Trial Timeline									
	n		*In-vivo* Activity	Clip (Valence)		Clip (Valence)		*In-vivo* Activity	Clip (Valence)	Clip (Valence)
A	2	Pre-survey	Jog	Serenity (=)	Imaginal instructions	Enjoyment (+)	Break	Walk	Serenity (=)	Fear (-)
B	3		Walk	Serenity (=)		Enjoyment (+)		Jog	Serenity (=)	Fear (-)
C	2		Jog	Serenity (=)		Fear (-)		Walk	Serenity (=)	Enjoyment (+)
D	2		Walk	Serenity (=)		Fear (-)		Jog	Serenity (=)	Enjoyment (+)

MATERIALS

Like many wearable devices, the Fitbit Inspire HR is equipped with a multitude of sensors including an accelerometer, gyroscope, and heart rate sensor [40]. To measure heart rate, the Fitbit uses photoplethysmography (PPG), a popular low-cost and noninvasive technique for optical measurement of volumetric changes in blood vessels [4]. LEDs mounted on the bottom of the device flash multiple times per second, then light-sensitive diodes detect these volumetric changes in the wrist capillaries [2]. Such volumetric changes have been shown to act as a surrogate for heart rate [27].

To simulate both *in-vivo* and *imaginal* aspects of PE treatment (see *Appendix C - Testing Protocol*), a collection of videos was chosen based on those compiled by Chen et al. [10] and Di Crosta et al [13]. The first category, “serenity,” was chosen to act as a control category, meant to calm the participant and establish a baseline of arousal. Hence, the “serenity” video clip was shown first for each participant, followed by the positive/negative valence clip. This clip was selected based on the work of Bednarski [5], which found that clips in this category most reliably elicited a feeling of serenity across a

wide range of participants. Serenity, or calmness, is a mental state free from anxiety, tension, negativity, and disturbance. Often linked with the feeling of "doing nothing" and the emotion of contentment [14,20], serenity encourages individuals to slow down and savor their current experiences. The video categories of "enjoyment" and "fear" were chosen to elicit two of the five basic emotional triggers defined by the Paul Ekman Group [10]. Enjoyment was chosen as the positive valence, while fear was selected for negative valence since fear commonly arises in PTSD through the re-experiencing of a traumatic event.

Table 2: Videos selected from the CAAV dataset

Title: Fear Videos	Mean Valence Score	Mean Arousal Score
Slapping someone (male)	1.74	7.11
Being threatened with a knife (male)	1.79	7.04
Kicking something	1.81	6.13
Being kidnapped	1.85	6.41
Snorting cocaine	1.98	5.15
Stealing from someone	2.03	5.67
Being threatened with a knife (female)	2.05	6.96
Drawing a swastika	2.07	5.53
Poisoning a person	2.08	6.22
Losing hair	2.17	6.2
Title: Enjoyment Videos	Mean Valence Score	Mean Arousal Score
Eating pizza (female)	7.76	6.05
Eating pizza (male)	7.73	5.54
Hugging someone	7.49	4.75
Finding cash (male)	7.48	5.84
Finding cash (female)	7.41	6.66
Receiving flowers from someone	7.38	5.64
Hugging someone	7.35	5.27
Celebrating with someone	7.35	5.21

Smiling at someone	7.31	5.77
Kissing someone	7.3	4.44

Table 3: Videos selected from Chen et. al (2020)

Video	Description	Intensity Score (1-5)
Enjoyment 1	Two pandas were playing together, and a sudden sneeze made people laugh.	5
Enjoyment 2	A little girl's father, who is in the military, surprises her at her birthday party.	5
Enjoyment 3	A little girl pretends to be a veterinarian and bandages a dog.	4
Fear 1	A person tries to capture a large snake that is hiding under a bed.	5
Fear 2	A person hanging off a tall building with one hand; a person sliding off a tall, slanted building.	4
Fear 3	A spider crawls over a young child's body.	4

Tables 2 and 3 summarize the video clips used in this study. Within each emotional category, amateur video clips studied by Chen et al. [10] were combined with a series of short clips from the Chieti Affective Action Videos (CAAV) dataset [13] to act as emotional elicitors for participants. For both sources, clips with the highest emotional intensities were selected to maximize emotive - and thus physiological - responses. Videos in the CAAV dataset used a 9-point Self-Assessment Manikin (SAM) scale to measure emotional valence and arousal, with 1 indicating a negative valence/low arousal and 9 corresponding to a positive valence/high arousal [7,13]. Videos chosen from Chen et al. [10] each had high (i.e. 4 to 5) emotional intensity scores as rated by the researchers on a 5-point Likert scale [24]. Likewise, out of the 361 clips in the CAAV database, the top 10 clips with the highest mean arousal score were selected for each category. Videos eliciting negative emotions (here, the fear category) had high valence scores, whereas more positive videos had lower valence ratings.

PROCEDURE

A usability testing protocol was established to ensure continuity and uniformity across all trials. The full protocol can be found in *Appendix C - Testing Protocol*. Before each trial, participants completed a pre-trial survey that gauged how frequently they perform cardiovascular exercise each week, their susceptibility to various phobias, as well as their hand dominance. These responses are summarized in *Table 4* along with participant demographic information. Data on activity levels were collected to control for different heart rate variations across participants during the *in-vivo*'s (e.g. sensed heart rates among highly active participants would be lower compared to less-active participants). Previous work has indicated that increased movement in the dominant hand leads to less accurate heart rate readings in wearables [27]. As such, data on participants' hand dominance was polled to determine the non-dominant hand on which to wear the Fitbit, thereby maximizing data accuracy across all participants. Participants were also polled on their propensity for Acrophobia (fear of heights), Arachnophobia (fear of spiders), and Ophidiophobia (fear of snakes). Since several of the most emotive videos selected for the "fear" category of videos depicted such phobias, controlling for this allowed for standardizing heart rate responses across the *imaginal* trials. The survey can be found in *Appendix A: Pre-Trial Survey*.

During the trial, both participants in the group were instructed to perform the given *in-vivo* activity (jogging or walking) while wearing the provided Fitbit on their non-dominant wrist and holding a provided Android smartphone running the PECSS mobile app. The watch had already been configured in the PECSS system, and each participant was instructed to keep the smartphone on their person for the duration of the trial. In each trial, participants first started a pre-assigned *in-vivo* assignment in the

app: either “Jog” or “Walk” corresponding to their group. The participants began by performing the *in-vivo* activity for 5 minutes. At the end of the period, they ended the assignment in the app and entered their pre-, intra-, and post-activity levels of exertion on a 1-100 scale. They then sat and watched a neutral clip for 5 minutes meant to elicit a state of serenity. This intermediate step allowed participants to calm down and return to a baseline level of physical arousal before proceeding. Continuing then to the pre-assigned *imaginal* task, the participants started the *imaginal* assignment in the PECSS app before being verbally informed about the contents and purpose of this section of the trial. Resembling the patient-clinician imaginal conversations in PE therapy, audio features were collected by the mobile app during this time. They then watched a 5-minute compilation of clips intended to elicit a specific emotional response (i.e. either enjoyment or fear). Participants were then allowed to take a 5-minute break before repeating this process for the other *in-vivo* activity and *imaginal* valence. This design allows all permutations of *in-vivo*/*imaginal* pairs to be tested, thereby ensuring the order of performed activity and emotional valence can be accounted for when analyzing the sensed heart rate. After both *in-vivo* and *imaginal* exposures were completed, participants were instructed to upload their data collected in the app. Such data were uploaded to a remote database for later retrieval and analysis.

Following the trial, participants completed a post-trial system usability scale (SUS) that asked participants for feedback on the usability of the Fitbit and mobile app in these mock exposures (see *Appendix B: Post-Trial Usability Survey*). The SUS [8] is an industry-standard, 10-item questionnaire that reliably scores systems (out of 100) on their usability and learnability [22]. In addition to the standard SUS questions, the survey also asked to what degree it increased their awareness of their physical responses

to stimuli, what concerns they have in the collection and ownership of their heart rate data, and what frustrations they had while completing the assignments. The results of these surveys, as well as the data from the rest of the trial, are displayed in the following section.

RESULTS

In this section, we present our findings and summarize the results of the trials. This includes responses from the aforementioned surveys and data collected by the PECSS mobile app and Fitbit watch.

SURVEYS

Table 4: Pre-Trial Participant Information

Group	Participant	Sex	Dominant Hand	Frequency of cardiovascular exercise	Fear of Heights (1-5)	Fear of Spiders (1-5)	Fear of Snakes (1-5)	Age	Used a Fitbit before?
A	1	Female	Right	Never	2	3	1	21	Yes
	2	Male	Right	A few times per month	2	2	2	19	No
B	3	Male	No preference	A few times per month	1	5	5	22	No
	4	Male	Right	Every day	2	4	4	21	No
	5	Male	Right	Once per week	4	2	2	18	No
C	6	Male	Right	A few times per month	2	3	1	21	No
	7	Female	Right	Once per week	1	3	4	20	No
D	8	Female	Right	Once per week	5	5	5	20	Yes
	9	Male	Right	A few times per month	4	2	5	28	Yes
				Averages:	2.56	3.22	3.22	21.1	
				Medians:	2	3	4	21	
				Std. Dev:	1.42	1.20	1.72	2.85	

The sample of nine participants reported a variety of activity levels and degrees of fear for the phobias shown during the trials. Most individuals reported engaging in rigorous cardiovascular exercise a few times per month or once a week, with one person each exercising every day and not at all. All of the participants wore the Fitbit on their left wrist given the near-universal right-hand dominance. Participants also reported a wide range of perceived fearfulness of heights, spiders, and snakes. Acrophobia (fear of

heights) saw the lowest median score of 2, whereas Arachnophobia (fear of spiders), and Ophidiophobia (fear of snakes) had median scores of 3 and 4 respectively.

Table 5 summarizes responses for the post-trial survey. The first 10 questions consisted of questions from the system usability survey (SUS), and the final question asked participants to assess to what degree wearing the Fitbit increased their awareness of their physiological reactions to the videos. All questions were scored on the standard SUS scale of 1-5, with 1 indicating a strong disagreement with the statement and 5 indicating a strong agreement [8,22]. Median scores for each question are also shown. Written feedback collected from this survey is included in *Appendix B - Participant Feedback*.

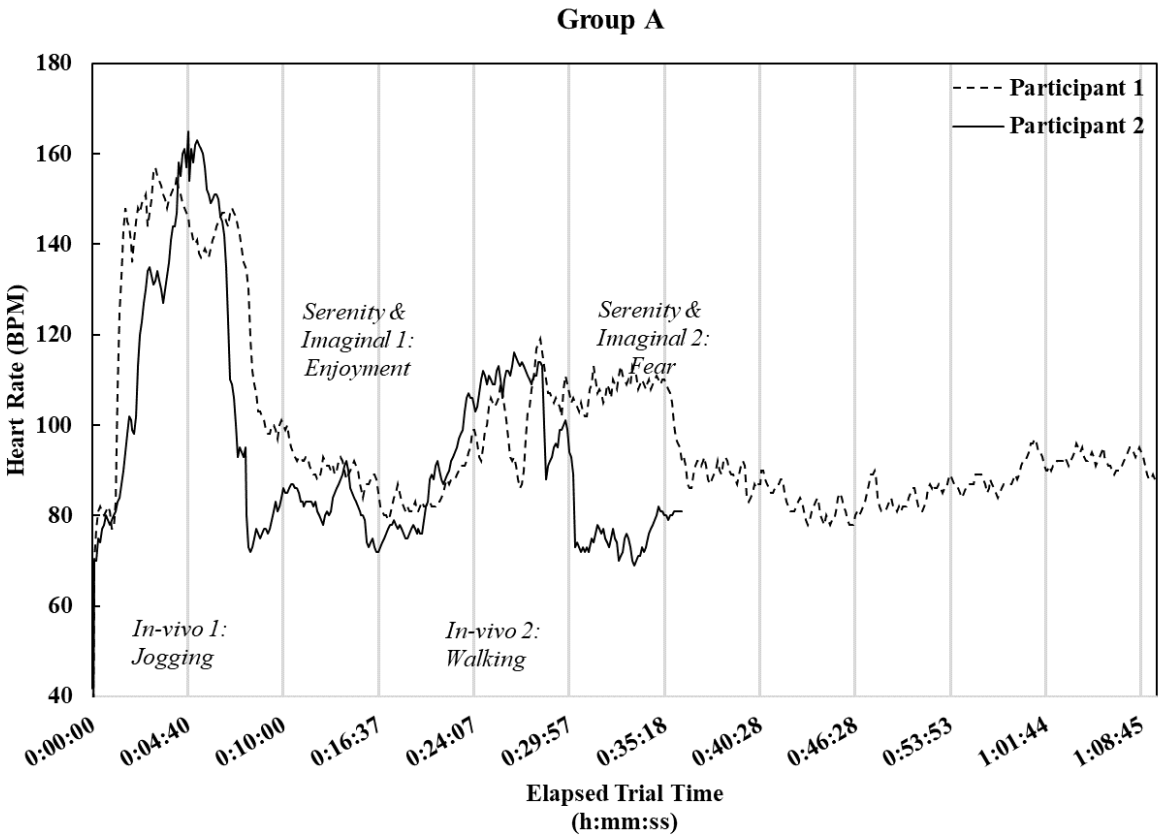
Table 5: Post-Trial Survey Results

		System Usability Scale (SUS) Questions											
Group	Participant	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	SUS Score	The Fitbit made me more aware of my physical responses to the videos shown.
A	1	3	1	3	1	4	1	5	1	3	2	80	5
	2	3	2	4	1	5	4	5	4	2	1	67.5	3
B	3	4	2	4	1	4	2	5	3	4	1	80	4
	4	4	1	4	2	4	1	4	1	5	2	85	2
	5	3	1	5	4	3	3	5	1	4	1	75	3
C	6	3	1	5	1	4	2	5	1	5	1	90	2
	7	4	1	5	1	3	1	5	1	5	1	92.5	3
D	8	4	3	3	4	4	3	4	4	3	3	52.5	4
	9	1	1	5	1	5	2	5	1	5	1	87.5	2
	Medians:	3	1	4	1	4	2	5	1	4	1		3

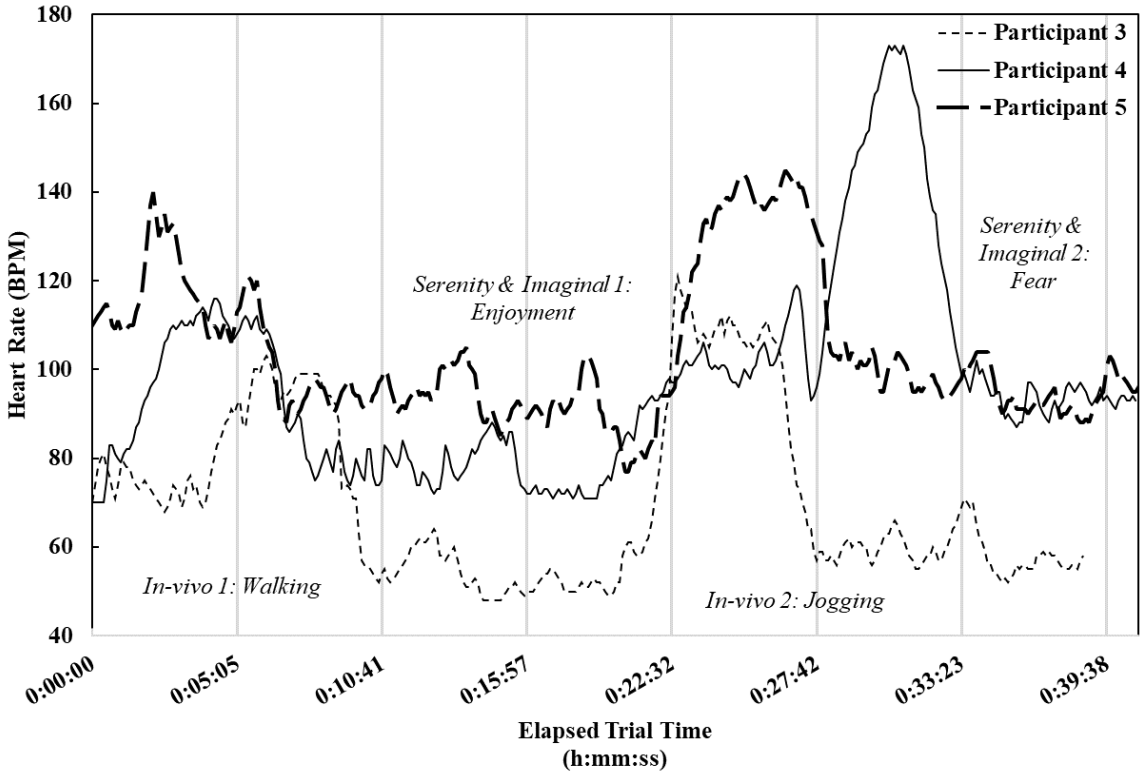
HEART RATE DATA

The graphs in *Figure 1* show participants' heart rate throughout the exposures in each trial group. Time segments are labeled according to the exposure (i.e. jogging or walking *in-vivo*; enjoyment, serenity, or fear *imaginal*). Note that since the Fitbit reports heart rate at a maximum granularity of 5-second intervals, the precision of these labels is inconsequential. *Table 6* summarizes the average, highest, lowest, and standard deviation in heart rate for the 5 exposures across all participants.

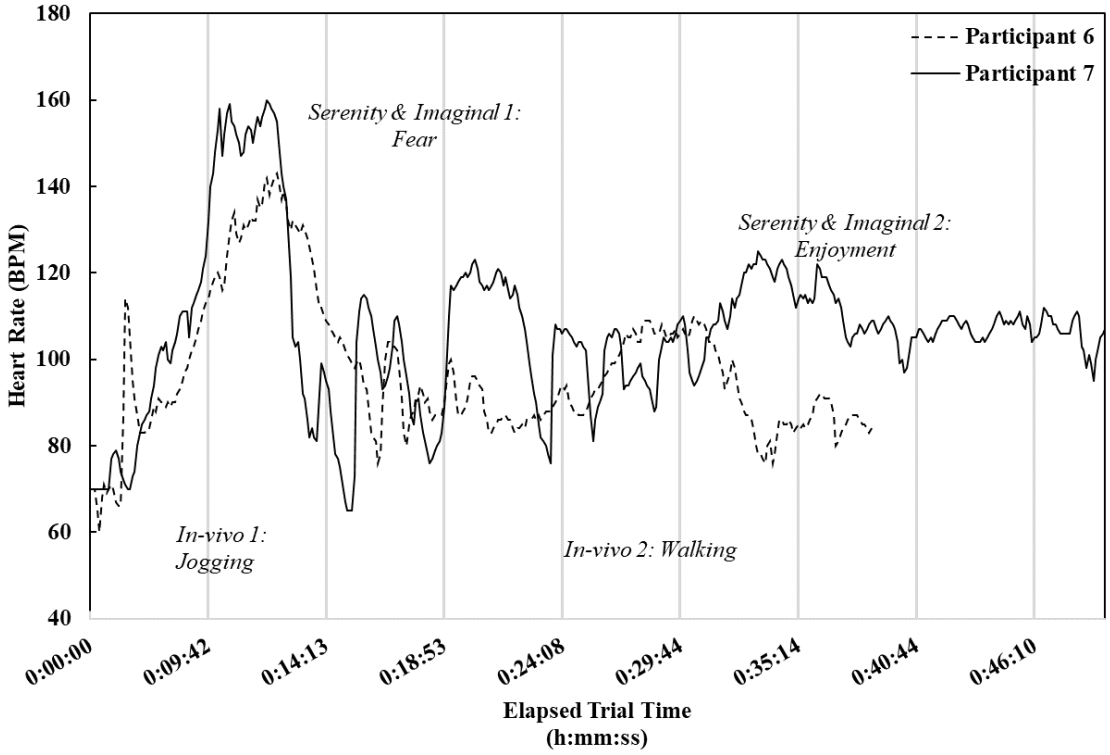
Figure 1: Heart Rate Graphs Across Trial Groups



Group B



Group C



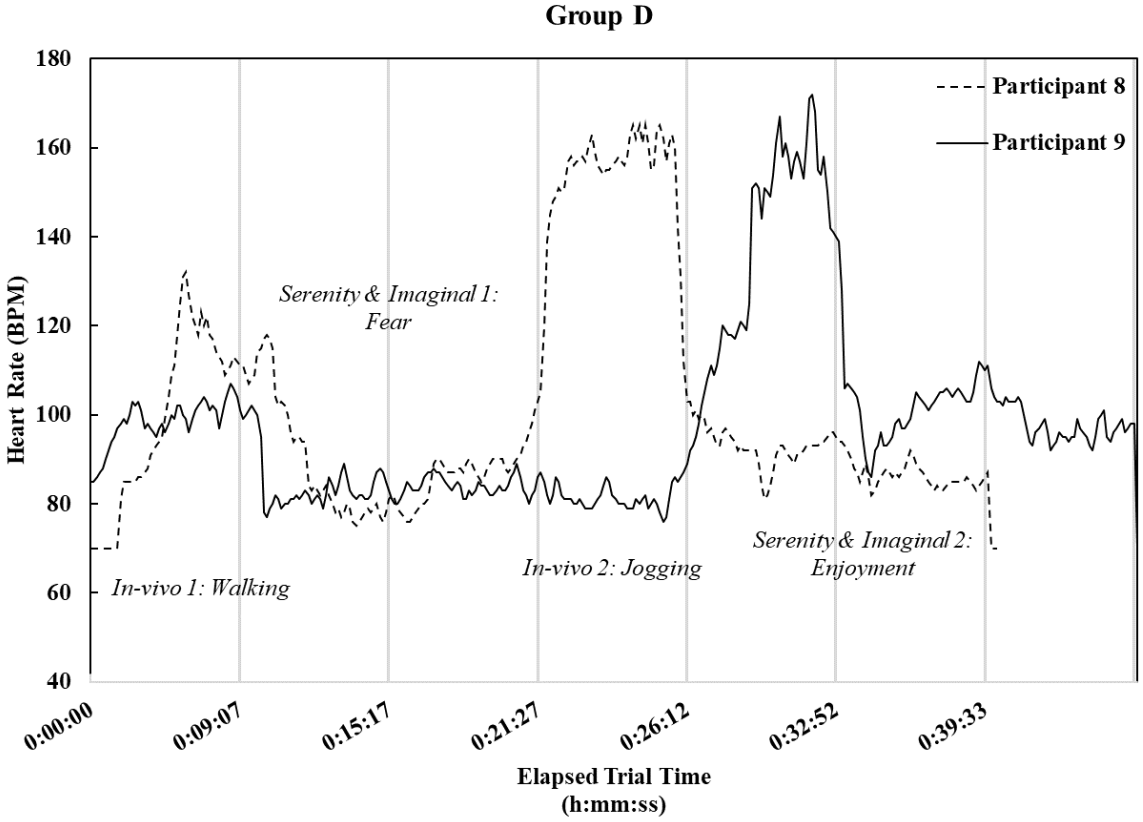


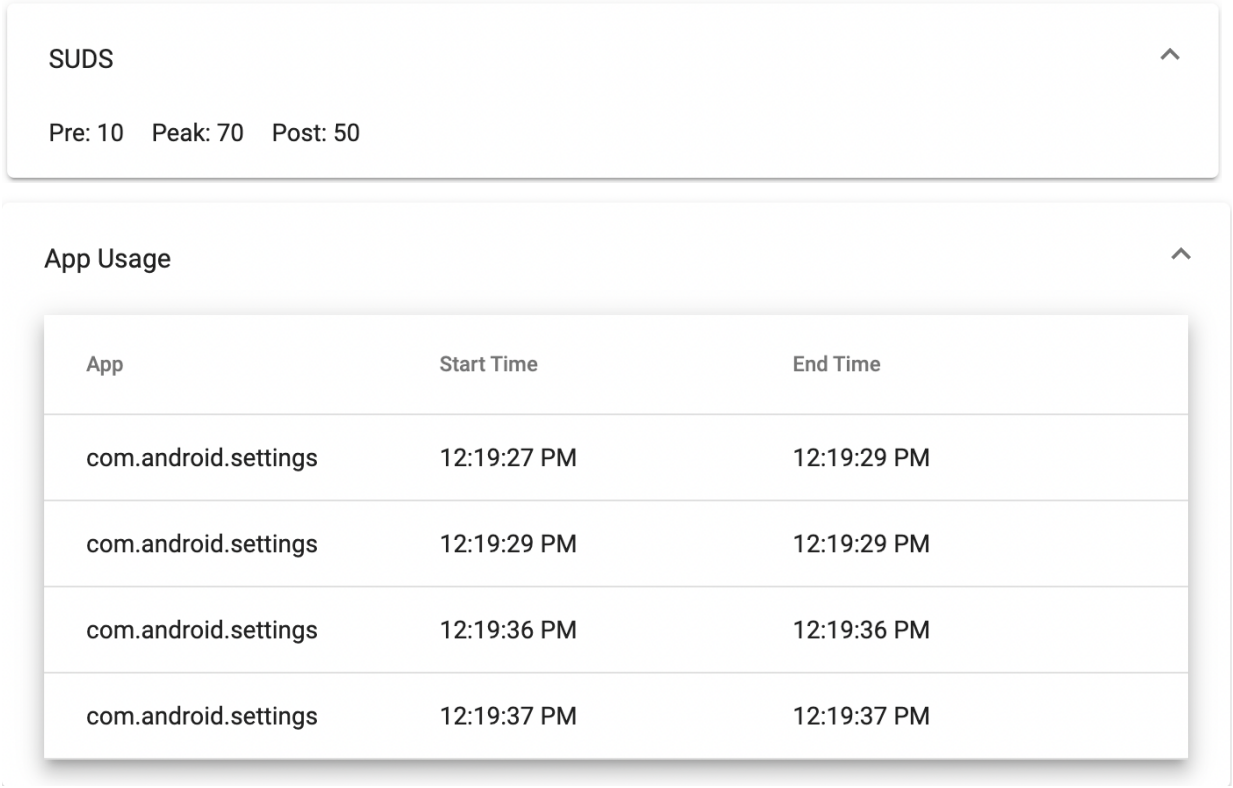
Table 6: Cross-Participant Exposure Statistics

Exposure	Mean Heart Rate (BPM) +/- 5.96%	Peak Heart Rate (BPM) +/- 5.96%	Lowest Heart Rate (BPM) +/- 5.96%	Standard Deviation
Walking	105.1	140	68	12.95
Jogging	132.5	173	70	23.58
Enjoyment	85.2	112	48	15.08
Fear	84.4	128	52	13.16
Serenity	86.7	115	52	11.18

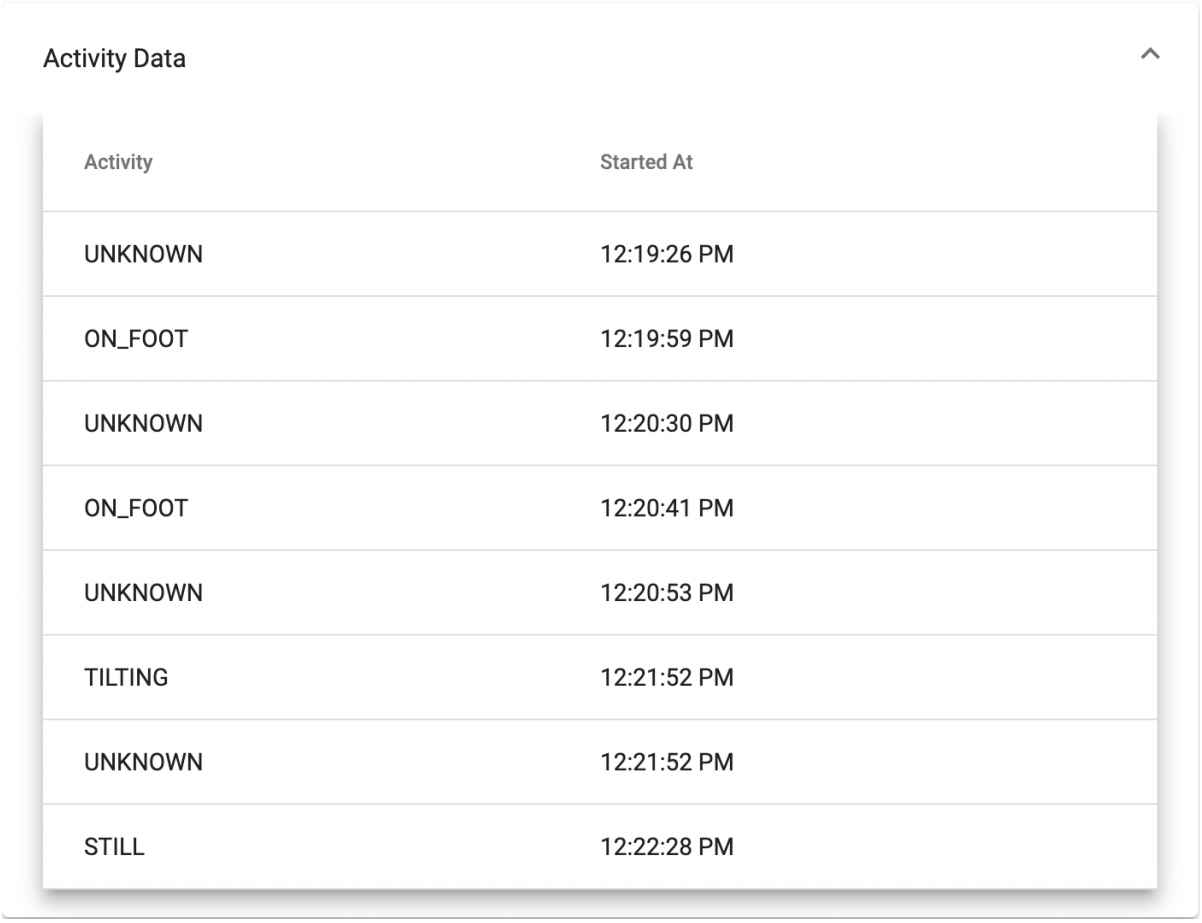
MOBILE APP DATA

Data collected from the mobile app supplemented the Fitbit data. Location data, app usage, activity data, and self-reported ratings of exertion were successfully collected for the *in-vivo* exposures. The phone generally classified the *in-vivo* activity correctly, detecting the participant was on foot while walking or jogging (see *Figure 2(b)*). Periodically, however, the phone could not classify the movement pattern, and thus reported "unknown." Participants did not exit the PECSS app to use other apps, except for Participant 1 who accidentally entered the settings app. The app still recorded this event, which can be found with the rest of the mobile app data in *Figure 2* below.

Figure 2: Mobile App Data



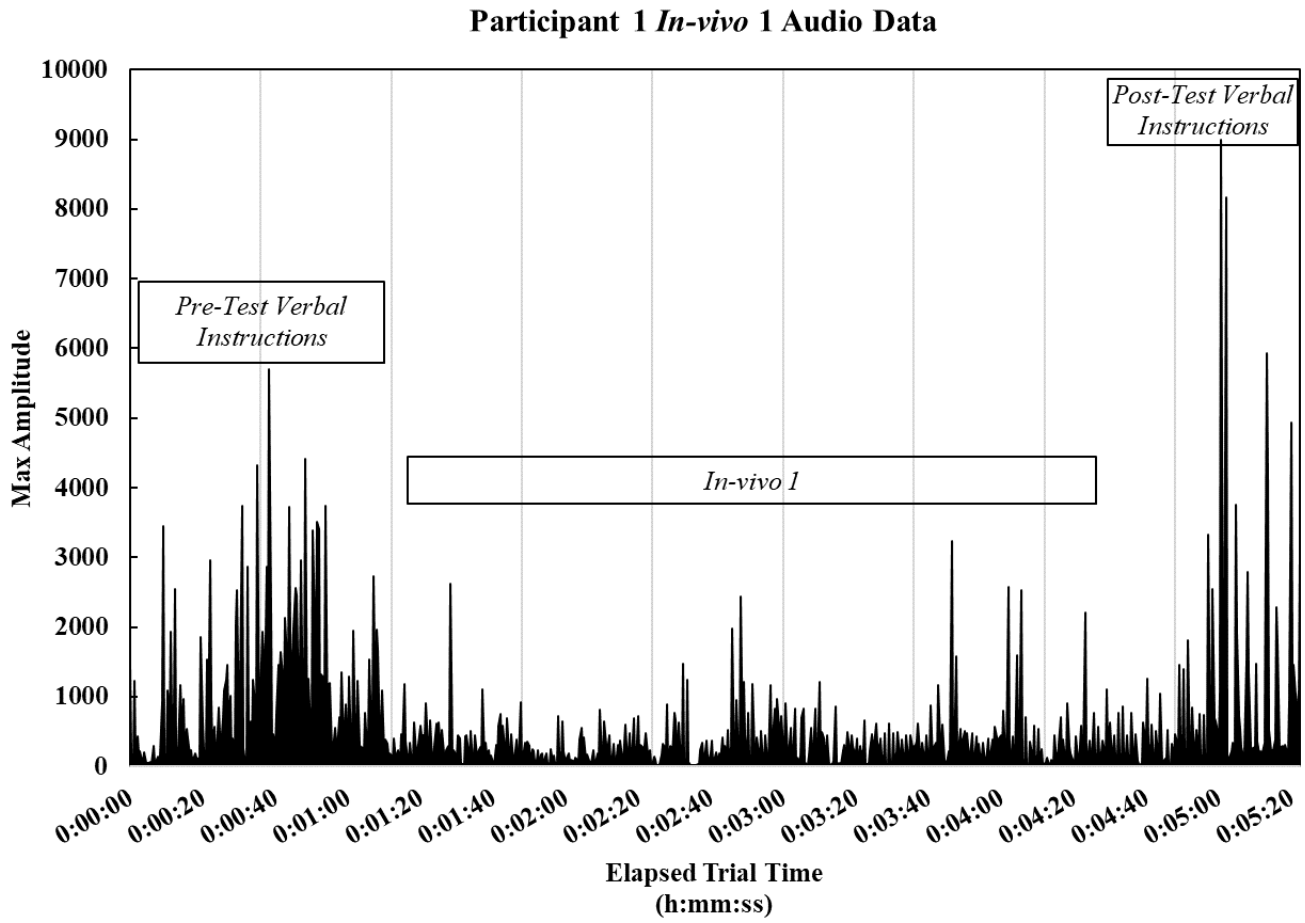
(a): Self-reported ratings of exertion and recorded app usage



(b): Activity classifications during *in-vivo* 1: jogging

Audio metrics (not raw recordings) were collected as well. *Figure 3* shows the maximum amplitude readings from the phone's internal microphones shortly before, during, and briefly after Participant 1's first *in-vivo* (jogging).

Figure 3: Audio Data



DISCUSSION

The Prolonged Exposure Collective Sensing System (PECSS) is intended to be used by patients undergoing Prolonged Exposure (PE) therapy. During *imaginal* sessions, patients use the app while verbalizing their trauma under clinician guidance, allowing them to listen back to the recording at a later date. In out-of-clinic *in-vivo* exposures, patients complete homework assignments in the app while wearing the Fitbit. We conducted a usability study with three aims. First, to collect heart rate data from participants in a variety of physically and mentally stimulating scenarios using Fitbits. Second, to use this data and ground truth data from the PECSS mobile app, to evaluate the efficacy of Fitbit integration in "PE therapy" (i.e., *in-vivo* and *imaginal* sessions). Finally, we wanted to ensure that using the Fitbits during the simulated PECSS deployment met the users' needs. The results of the study support each of these aims; however, there are some limitations. We will discuss the major findings related to aims one and two in this section and those related to the third aim in the design implications section. We will also discuss the results of this work relating to the existing literature.

Regarding the first and second aims, heart rate data collected from the Fitbits during the trial provided insights into the participants' responses to various physical and mental stimuli. Across all four groups, distinct peaks in heart rate appear during the jogging and, to a lesser extent, walking *in-vivo*. Furthermore, the inclusion of the serenity video appeared to accomplish its intended purpose: to allow for a period of relaxation to restore the participant's heart rate to pre- *in-vivo* levels before performing the *imaginal* exposure. In support of existing work [15,27] regarding wearable heart rate

accuracy during a variety of physical and non-physical activities, the Fitbit was indeed able to capture differences in heart rate between the various exposures, particularly among the physical *in-vivo's* (see *Table 6*). The average heart rate for jogging (132.5 BPM) was 27.4 BPM higher than that of walking (105.1 BPM), which itself was 18.4 BPM higher than the average heart rate while watching the serenity video (86.7 BPM). Therefore, the Fitbit can not only discern whether the user is engaged in physical activity but also the degree of intensity of that activity. This contributes an additional dimension of feedback for clinicians, who can use the mobile app's activity classification as a coarse indication of what homework segment was being performed, then refer to the Fitbit data for information on how stimulating the activity was. Furthermore, participants were fairly neutral regarding whether the Fitbit affected their awareness of their physical reactions to the exposures. Combined with the lack of negative feedback regarding the Fitbit inhibiting their ability to complete the exposures properly, this may substantiate previous findings [27,28] that such wearables are suitable where low patient burden is imperative.

However, heart rate responses while watching the videos intended to elicit emotions of fear or joy were less uniform. Participants 1, 6, 7, and 9 showed small, or in the case of Participant 7 extreme, jumps in heart rate while watching the fearful videos. In contrast, the other participants saw normal heart rate responses, even if they had indicated having one or more strong phobias (e.g., Participant 8). The enjoyment videos yielded a somewhat stronger pattern. Participants 2, 4, 5, 6, 7, and 8 showed marked spikes in heart rate while watching these clips. Notably, despite self-scoring a low fear of all three phobias (rating of 2 out of 5 for each), Participant 2 had one of the most responsive heart rates while watching the fearful videos. This highlights a crucial

affordance of the Fitbit: it may provide more accurate insight into user responses to exposures when participants may diminish or exaggerate such responses.

The audio metrics also yielded promising results. When the participant was instructed how to perform the exposure, both before and after the *in-vivo* and just before the serenity video portion, the maximum amplitude increased accordingly. Similarly, we observed a period of decreased amplitude while the participant was jogging/walking (see *Figure 3*). It is important to note that since little speaking was involved beyond short instructions and *imaginal* briefings (see *Appendix C - Testing Protocol*), such data served simply as proof of concept that the phone could effectively measure periods of speaking. Still, the phone's apparent ability to capture audio with reasonable fidelity complements Wang et al.'s [38] assertion of the value of other smartphone metrics in mHealth therapy.

DESIGN IMPLICATIONS

In addressing the third aim of this study, the trials revealed several design implications for the Prolonged Exposure Collective Sensing System (PECSS) mobile app and its pairing with the Fitbit. In this section, we discuss these implications and their effect on the usability of the PECSS system, then make recommendations for updating the system. Overall, participants gave the system an average SUS score of 78.89, indicating a positive usability experience compared to the average SUS score of 68 [22]. However, participants still reported problems while performing the mock activities. One participant reported difficulty in attaching the Fitbit to their wrist, as the default band size was too small for their wrist circumference. Moving forward, it will be good practice to have both regular and extended bands on hand to best suit every patient. The trials

quickly revealed bugs in the mobile app as well. While Participant 1 completed the *in-vivo* activity, they found that screen rotation, which could occur while holding the phone and walking/jogging, would cancel the attempt in the app. This was quickly remedied by disabling screen rotation in the phone, but going forward, starting an attempt should disable screen rotation for the duration of that attempt. Similarly, two participants reported concerns over the navigation buttons present at the bottom of the Android's screen. Accidentally pressing the phone's back button cancels the ongoing attempt in the app's current version, thus subsequent iterations should disable these buttons while the attempt is in progress. Participants also reported confusion in the UI while starting the imaginal exposure recordings. The app required users to hit "Start" then "Start Imaginal." Conversely, to end the exposure, participants were presented with both a "Stop Imaginal" button and a "Stop Complete Imaginal" button. The former intended to allow users to pause the recording and resume at a later time, but such labeling confused participants as to what button to press to end the exposure after the videos had concluded. Finally, it is important to note that none of the participants pressed the "Upload Data" button on the homepage of the app. Uploading assignment data is crucial for clinicians to view & analyze the patient's data. Assignment data was not synchronized automatically - for example, on completion of an attempt - to promote patient agency in deciding whether to share their data with their clinician or not. However, since the participants in this study appeared to be unaware that uploading the data was a deliberately manual process, this raises two implications for the PECSS system. First, patients should be informed that they possess the choice to share their exposure homework data with their care team or not. This should include discussing the consequences of sharing versus not sharing data on their experience in care. Second, to

ensure patients remember to upload their data (or not), the app should ask patients upon completion of an attempt whether they would like to share their data with their care team. Post-trial feedback also highlighted this concern for data privacy, agreeing with Paul et al. [30] that data privacy is a critical concern where wearables are employed.

CONCLUSION

The current study showed that the Fitbit was able to capture physiological manifestations of physical and emotional stimulation in the simulated PE activities via heart rate. In integrating the wearable into the PECSS sensor suite, this study also found the system to be usable and learnable while providing an additional dimension for the assessment of users' performance during such exposures. In this section, we summarize the importance of these findings and discuss the study's limitations and areas for future researchers to explore.

The Fitbit's ability to reliably capture changes in heart rate could equip clinicians with additional insight into the *in-vivo* and *imaginal* exposures. Upon aligning each attempt's timestamps with the heart rate data, it became immediately obvious when the participants were performing the exposure. This could allow clinicians to not only verify patient participation in out-of-clinic exercises, but also tune their intuition and decision-making to align with how viscerally their patients respond physiologically to the exposures.

Additionally, the Fitbit showed promising results in its capability to sense smaller changes in heart rate induced by joyful or scary imagery during the *imaginal* exposures. This is a crucial dimension of data for clinicians since the re-experiencing of traumatic events - which occurs in PE through narrated recounts during *imaginal* sessions - has been shown to manifest in heart rate variability in those with PTSD [32]. Specifically, our observation that heart rate variability decreased while viewing fearful videos as compared to the joyful clips aligns with previous findings that phobic responses to negative stimuli can manifest in decreased heart rate variability [32].

LIMITATIONS

Since this is a pilot usability study, there are limitations to the results presented. In this section, we discuss several notable limitations to consider when interpreting our data. First, this data was collected in a controlled setting via a convenience sample of PECSS researchers (age 19-28) who were at least somewhat familiar with the app and PE process, not patients with PTSD. Though for this study - that is, evaluating the usability and learnability of the PECSS system with Fitbits - this does not pose a significant threat to the data gathered, it is important to note that PTSD patients may find different aspects of the app or use of the Fitbit confusing. The participants in this study were at least somewhat familiar with the app and generally technologically savvy, whereas real-world patients may not be. Consequently, the scores collected from the SUS are likely skewed towards a more positive usability rating. Second, our inability to find a single, 5-minute video supported by prior research to elicit either a positive or negative emotional response caused us to stitch 10 shorter clips together. This resulted in a random compilation of videos previously shown to elicit a particular valence, but the lack of continuity between them confused some participants, possibly at the detriment of the videos' effectiveness. Third, it is important to note that real *in-vivo* and *imaginal* exposures deployed in PE therapy are typically much longer than 5 minutes, sometimes exceeding an hour in length [17]. As such, usability issues that were not raised in these short exposures may become visible over longer ones. Finally, heart rate data were analyzed observationally and not subjected to standard psychophysiological analysis methods such as intragroup t-tests, heart rate variability calculations, or electrocardiographic (ECG) R-wave time interval analyses [1,6].

FUTURE WORK

The findings presented in this pilot study serve as evidence of the usability of the system and support previous research into the affordances of wearable technology in clinical settings, specifically for heart rate sensing in physically and emotionally stimulating situations. In seeking to fulfill one of the aims of the study, the pilot data collected in this study and design implications discussed above will be employed as part of the greater PECSS project to improve usability in the app. The protocol developed for this study contributes several design elements for the deployment of Fitbits in the PECSS system (e.g., on which wrist to wear the watch, how users should be informed on the use of their heart rate data, etc.). As discussed previously, the major limitation of this study is the lack of PTSD patients evaluating the usability of the PECSS mobile app and Fitbit during PE treatment. Future work should explore how the mobile app's UI can be improved to increase usability for this specific population. Furthermore, previous research has indicated a correlation between the tonality of speech in PTSD patients and their response to trauma [25]. The PECSS system should be deployed in real *imaginal* exposures to explore whether the audio metrics can capture this phenomenon.

APPENDIX A

PRE-TRIAL SURVEY

Pre-Trial Survey

Thank you for participating in our pilot study. Please answer the questions below.

* Required

1. First Name *

2. Last Name *

3. Age *

4. Please indicate your sex *

Mark only one oval.

- ◯ Male
- ◯ Female
- ◯ Other: _____

5. Have you used a Fitbit fitness tracker before? *

Mark only one oval.

- ◯ Yes
- ◯ No

6. Please select your dominant hand, or "No preference" if you are ambidextrous. *

Mark only one oval.

- ◯ Right
- ◯ Left
- ◯ No preference

7. How often do you run or engage in rigorous cardio? *

Mark only one oval.

- ◯ Never
- ◯ A few times per month
- ◯ Once per week
- ◯ Several days per week
- ◯ Every day

8. How fearful are you of heights? *

Mark only one oval.

	1	2	3	4	5	
Not scared at all	◯	◯	◯	◯	◯	Terrified

9. How fearful are you of spiders? *

Mark only one oval.

	1	2	3	4	5	
Not scared at all	◯	◯	◯	◯	◯	Terrified

10. How fearful are you of snakes? *

Mark only one oval.

	1	2	3	4	5	
Not scared at all	◯	◯	◯	◯	◯	Terrified

APPENDIX B

POST-TRIAL USABILITY SURVEY

Post-Trial Survey

Thank you for participating in our pilot study. Please answer the questions below.

* Required

1. First Name *

2. Last Name *

System Usability Survey (SUS)

3. I think that I would like to use this system frequently. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

4. I found the system unnecessarily complex. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

5. I thought the system was easy to use. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

6. I think that I would need the support of a technical person to be able to use this system. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

7. I found the various functions in this system were well integrated. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

8. I thought there was too much inconsistency in this system. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

9. I would imagine that most people would learn to use this system very quickly. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

10. I found the system very cumbersome to use. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

11. I felt very confident using the system. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

12. I needed to learn a lot of things before I could get going with this system. *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

13. The Fitbit made me more aware of my physical responses to the videos shown *

Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

Written feedback

14. What frustrations did you have with using the app or Fitbit during the trial?

15. What concerns (e.g. privacy) do you have regarding the collection and use of your heart rate data by third-parties?

PARTICIPANT FEEDBACK

Participant	What frustrations did you have with using the app or Fitbit during the trial?	What concerns (e.g. privacy) do you have regarding the collection and use of your heart rate data by third parties?
1	1. I cannot actually see current heart rate data being recorded (and there is no confirmation message like "your Fitbit data is being recorded now"), so I'm unsure if it actually tracks my heart rate; 2. Android system have a back button that can be accidentally clicked during the trial and I'm nervous about clicking on it (so I have to redo the trial again); 3. I don't like holding my phone while walking (I naturally have a desire to play with my phone) or jogging (hard to hold and hard to avoid the accidental click on the screen). 4. It is weird that the timer is already started in the imaginal exposure but I need to click on start imaginal again (I feel I may forget to do this sometimes); it's confusing.	As long as they don't have my name on it, everything is fine.
2	I think that everything worked fine. I just think that the "taps" could be simplified. Like sometimes I need to hit start twice or stop twice or I need to pause and then stop. There should just be like three buttons. Play, pause and stop and then like a save button. Just very intuitive with icons and I think that would be very beneficial for people. Also, I am not sure how people set up the Fitbit so that could be a cause for a different answer on some of those questions.	I have no concerns
3	No frustrations	No concerns
4	None	None
5	Having to press the start button twice for imaginals felt a bit redundant	None
6	None	None
7	I did not have any frustrations	I do not have any concerns
8	I had no frustrations with the Fitbit during the trial but I do feel that the system might be confusing to those unaware or unfamiliar with parts of the system, i.e. what is the purpose of an imaginal or in vivo and how will it help me?	Concerns regarding privacy are that it is safely secured and not used in a way that would be harmful to the person.
9	None	I have no concern since it is just heart rate data.

APPENDIX C

TESTING PROTOCOL

Pre-trial Requirements

- ☐ Participant is provided with a Fitbit Inspire 2 that is connected to a Fitbit account to store/retrieve data.
- ☐ Participant is provided with an Android phone with PECSS app installed
    - ☐ The phone's screen timeout is set to 10 minutes. Screen rotation is disabled.
- ☐ The Fitbit has been entered into the PECSS database and linked to the participant's ID. This can be done through the dashboard if implemented, otherwise, it can be done manually through the API.
- ☐ An *in-vivo* homework assignment is assigned to them in the PECSS app corresponding to the type of physical activity to be performed.

Briefing

- ☐ Participant fills out pre-trial survey

  [Researcher 1] - "Before we proceed, please fill out this anonymous pre-trial survey."

- ☐ Participant is briefed on the study

  [Researcher 1] - "Thank you for agreeing to participate in this study. I will briefly explain the goals of this study and describe what we will be doing today. This pilot study has the following goals: (1) to collect heart rate data from participants in a variety of physically or mentally stimulating scenarios using Fitbits. (2) Using this data and ground truth data from the PECSS mobile app, evaluate the usability and learnability of Fitbits integrated into PE therapy. And (3) ensure the use of Fitbits in the deployment of PECSS will be based on the findings of this study to address patient needs."

  [Researcher 2] - "In today's trial, you will complete two mock "in-vivo" assignments and two mock "imaginal" assignments. You will hold a provided smartphone in your hand and wear a Fitbit watch throughout these assignments. In the first in-vivo, you will jog or walk around this area for 5 minutes. Once

those 5 minutes are up, you will sit and watch a series of videos. Upon completion, you will record your levels of emotional arousal in the app. After a short break, you will repeat this process with the other in-vivo and imaginal assignments."

In-vivo: participants jog/walk for 5 minutes

Pre-test:

- ☐ Before participant equips the Fitbit, Researcher 1 walks them around the route.
- ☐ Researcher 1 selects an *in-vivo* assignment and instructs participant to hit "Start New *In Vivo* exposure."
- ☐ Participant enters a "Subjective Unit of Exertion" 1-100 as the "Pre-SUD" and starts the assignment in the PECSS app.

  [Researcher 1] - "In the Pre Subjective Unit of Distress box, put your current level of exertion. Then press 'Start *In Vivo*'."

Intratest:

- ☐ Participant performs the activity - jogging or walking - for 5 minutes.
  - ☐ Participant is to wear the Fitbit on their non-dominant wrist
  - ☐ Participant is to hold the phone in their dominant hand during the activity
  - ☐ [Researcher 2] - counts the number of laps they complete in the 5 minutes
- ☐ After 5 minutes, participant is instructed to answer the questions on peak and post-SUD values in the app, then press "Complete *In Vivo*."

Post-test:

- ☐ Participant is instructed to watch the "serenity" video for 5 minutes.

Imaginal: watching video clips

Pre-test:

- ☐ Participant is provided with two videos: "enjoyment" and "fear"
- ☐ Participant is given headphones

Intratest:

- ☐ Participant presses “Record a New Imaginal Exposure” in the PECSS app and hit start imaginal (twice).
- ☐ Participant is instructed what to do next

  [Researcher 1] - “You will now view a series of clips. Please pay attention and wear headphones while doing so, and let me know when the video has finished. If at any time you feel too uncomfortable, please let me know and we will stop the exposure.”
- ☐ Participant watches the video
  - ☐ The phone is sitting on the table in front of them
- ☐ Once the video ends, participant presses “Stop imaginal” then “Complete imaginal” in the app.

Participant is given a 5-minute break to do whatever they want. Afterward, the above in-vivo and imaginal are repeated with the other activity & valence.

Post-trial

- ☐ Participant fills out post-trial survey

  [Researcher 1] - “Thank you for completing the trial and for participating in our pilot study. No identifying information will be included in the report of this trial. Please fill out this anonymous post-trial survey.”
- ☐ Researcher 1 presses the “Upload data” button in the PECSS app.

REFERENCES

[1] Rajendra Acharya U, Kannathal N, Ong Wai Sing, Luk Yi Ping, and TjiLeng Chua. 2004. Heart rate analysis in normal subjects of various age groups. *Biomed. Eng. OnLine* 3, 1 (July 2004), 24. DOI:https://doi.org/10.1186/1475-925X-3-24

[2] Daniel A. Adler, Vincent W.-S. Tseng, Gengmo Qi, Joseph Scarpa, Srijan Sen, and Tanzeem Choudhury. 2021. Identifying Mobile Sensing Indicators of Stress-Resilience. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 2 (June 2021), 51:1-51:32. DOI:https://doi.org/10.1145/3463528

[3] Rawan Alharbi, Angela Pfammatter, Bonnie Spring, and Nabil Alshurafa. 2017. WillSense: Adherence Barriers for Passive Sensing Systems That Track Eating Behavior. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '17), Association for Computing Machinery, New York, NY, USA, 2329–2336. DOI:https://doi.org/10.1145/3027063.3053271

[4] John Allen. 2007. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* 28, 3 (March 2007), R1-39. DOI:https://doi.org/10.1088/0967-3334/28/3/R01

[5] Juliana Doria Bednarski. 2012. Eliciting Seven Discrete Positive Emotions Using Film Stimuli. (March 2012), 69.

[6] Ravi Bhoja, Oren T. Guttman, Amanda A. Fox, Emily Melikman, Matthew Kosemund, and Kevin J. Gingrich. 2020. Psychophysiological Stress Indicators of Heart Rate Variability and Electrodermal Activity With Application in Healthcare Simulation Research. *Simul. Healthc.* 15, 1 (February 2020), 39–45. DOI:https://doi.org/10.1097/SIH.0000000000000402

[7] Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 1 (March 1994), 49–59. DOI:https://doi.org/10.1016/0005-7916(94)90063-9

[8] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval Ind* 189, (November 1995).

[9] Steven Chan, John Torous, Ladson Hinton, and Peter Yellowlees. 2015. Towards a Framework for Evaluating Mobile Mental Health Apps. *Telemed. E-Health* 21, 12 (December 2015), 1038–1041. DOI:https://doi.org/10.1089/tmj.2015.0002

[10] HongYi Chen, Kai Ling Chin, and Chrystalle B. Y. Tan. 2021. Selection and validation of emotional videos: Dataset of professional and amateur videos that elicit basic emotions. *Data Brief* 34, (February 2021), 106662. DOI:https://doi.org/10.1016/j.dib.2020.106662

[11] Jenny Chum, Min Suk Kim, Laura Zielinski, Meha Bhatt, Douglas Chung, Sharon Yeung, Kathryn Litke, Kathleen McCabe, Jeff Whattam, Laura Garrick, Laura O'Neill, Stefanie Goyert, Colleen Merrifield, Yogita Patel, and Zainab Samaan. 2017. Acceptability of the Fitbit in behavioural activation therapy for depression: a qualitative study. *Evid. Based Ment. Health* 20, 4 (November 2017), 128–133. DOI:https://doi.org/10.1136/eb-2017-102763

[12] Chia-Fang Chung, Kristin Dew, Allison Cole, Jasmine Zia, James Fogarty, Julie A. Kientz, and Sean A. Munson. 2016. Boundary Negotiating Artifacts in Personal Informatics: Patient-Provider Collaboration with Patient-Generated Data. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* (CSCW '16), Association for Computing Machinery, New York, NY, USA, 770–786. DOI:https://doi.org/10.1145/2818048.2819926

[13] Adolfo Di Crosta, Pasquale La Malva, Claudio Manna, Anna Marin, Rocco Palumbo, Maria Cristina Verrocchio, Michela Cortini, Nicola Mammarella, and Alberto Di Domenico. 2020. The Chieti Affective Action Videos database, a resource for the study of emotions in psychology. *Sci. Data* 7, 1 (January 2020), 32. DOI:https://doi.org/10.1038/s41597-020-0366-1

[14] Phoebe C. Ellsworth and Craig A. Smith. 1988. Shades of Joy: Patterns of Appraisal Differentiating Pleasant Emotions. *Cogn. Emot.* 2, 4 (October 1988), 301–331. DOI:https://doi.org/10.1080/02699938808412702

[15] Sindhu Kiranmai Ernala, Michael L. Birnbaum, Kristin A. Candan, Asra F. Rizvi, William A. Sterling, John M. Kane, and Munmun De Choudhury. 2019. Methodological Gaps in Predicting Mental Health States from Social Media: Triangulating Diagnostic Signals. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (CHI '19), Association for Computing Machinery, New York, NY, USA, 1–16. DOI:https://doi.org/10.1145/3290605.3300364

[16] Hayley Evans, Udaya Lakshmi, Hue Watson, Azra Ismail, Andrew M. Sherrill, Neha Kumar, and Rosa I. Arriaga. 2020. Understanding the Care Ecologies of Veterans with PTSD. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20), Association for Computing Machinery, New York, NY, USA, 1–15. DOI:https://doi.org/10.1145/3313831.3376170

[17] Edna B. Foa, Elizabeth Hembree, and Barbara Rothbaum. *Prolonged Exposure Therapy for PTSD: Emotional Processing of Traumatic Experiences, Therapist Guide*. Oxford University Press. Retrieved October 3, 2021 from https://www.oxfordclinicalpsych.com/view/10.1093/med:psych/9780195308501.001.0001/med-9780195308501

[18] Zac E. Imel, Kevin Laska, Matthew Jakupcak, and Tracy L. Simpson. 2013. Meta-analysis of dropout in treatments for posttraumatic stress disorder. *J. Consult. Clin. Psychol.* 81, 3 (June 2013), 394–404. DOI:https://doi.org/10.1037/a0031474

[19] Shannon M. Kehle-Forbes, Laura A. Meis, Michele R. Spoont, and Melissa A. Polusny. 2016. Treatment initiation and dropout from prolonged exposure and cognitive processing therapy in a VA outpatient clinic. *Psychol. Trauma Theory Res. Pract. Policy* 8, 1 (January 2016), 107–114. DOI:https://doi.org/10.1037/tra0000065

[20] Richard S. Lazarus. 1991. *Emotion and adaptation*. Oxford University Press, New York, NY, US.

[21] Catrin Lewis, Neil P. Roberts, Samuel Gibson, and Jonathan I. Bisson. 2020.

Dropout from psychological therapies for post-traumatic stress disorder (PTSD) in adults: systematic review and meta-analysis. *Eur. J. Psychotraumatology* 11, 1 (March 2020), 1709709. DOI:https://doi.org/10.1080/20008198.2019.1709709

[22] James R. Lewis and Jeff Sauro. 2009. The Factor Structure of the System Usability Scale. In *Human Centered Design* (Lecture Notes in Computer Science), Springer, Berlin, Heidelberg, 94–103. DOI:https://doi.org/10.1007/978-3-642-02806-9_12

[23] Ian Li, Anind K. Dey, and Jodi Forlizzi. 2011. Understanding my data, myself: supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th international conference on Ubiquitous computing* (UbiComp '11), Association for Computing Machinery, New York, NY, USA, 405–414. DOI:https://doi.org/10.1145/2030112.2030166

[24] R. Likert. 1932. A technique for the measurement of attitudes. *Arch. Psychol.* 22 140, (1932), 55–55.

[25] Charles R. Marmar, Adam D. Brown, Meng Qian, Eugene Laska, Carole Siegel, Meng Li, Duna Abu-Amara, Andreas Tsiartas, Colleen Richey, Jennifer Smith, Bruce Knoth, and Dimitra Vergyri. 2019. Speech-Based Markers for Posttraumatic Stress Disorder in U.S. Veterans. *Depress. Anxiety* 36, 7 (July 2019), 607–616. DOI:https://doi.org/10.1002/da.22890

[26] John A. Naslund, Kelly A. Aschbrenner, and Stephen J. Bartels. 2016. Wearable devices and smartphones for activity tracking among people with serious mental illness. *Ment. Health Phys. Act.* 10, (March 2016), 10–17. DOI:https://doi.org/10.1016/j.mhpa.2016.02.001

[27] Benjamin W. Nelson and Nicholas B. Allen. 2019. Accuracy of Consumer Wearable Heart Rate Measurement During an Ecologically Valid 24-Hour Period: Intraindividual Validation Study. *JMIR MHealth UHealth* 7, 3 (March 2019), e10828. DOI:https://doi.org/10.2196/10828

[28] Benjamin W. Nelson, Carissa A. Low, Nicholas Jacobson, Patricia Areán, John Torous, and Nicholas B. Allen. 2020. Guidelines for wrist-worn consumer wearable

assessment of heart rate in biobehavioral research. *Npj Digit. Med.* 3, 1 (June 2020), 1–9. DOI:https://doi.org/10.1038/s41746-020-0297-4

[29] Ada Ng, Madhu Reddy, Alyson K. Zalta, and Stephen M. Schueller. 2018. Veterans' Perspectives on Fitbit Use in Treatment for Post-Traumatic Stress Disorder: An Interview Study. *JMIR Ment. Health* 5, 2 (June 2018), e10415. DOI:https://doi.org/10.2196/10415

[30] Greig Paul and James Irvine. 2014. Privacy Implications of Wearable Health Devices. In *Proceedings of the 7th International Conference on Security of Information and Networks* (SIN '14), Association for Computing Machinery, New York, NY, USA, 117–121. DOI:https://doi.org/10.1145/2659651.2659683

[31] Mark B. Powers, Johnna L. Medina, Stephanie Burns, Brooke Y. Kauffman, Marie Monfils, Gordon J. G. Asmundson, Allison Diamond, Christa McIntyre, and Jasper A. J. Smits. 2015. Exercise Augmentation of Exposure Therapy for PTSD: Rationale and Pilot Efficacy Data. *Cogn. Behav. Ther.* 44, 4 (2015), 314–327. DOI:https://doi.org/10.1080/16506073.2015.1012740

[32] C. S. Rosen, M. M. Matthieu, S. Wiltsey Stirman, J. M. Cook, S. Landes, N. C. Bernardy, K. M. Chard, J. Crowley, A. Eftekhari, E. P. Finley, J. L. Hamblen, J. M. Harik, S. M. Kehle-Forbes, L. A. Meis, P. E. Osei-Bonsu, A. L. Rodriguez, K. J. Ruggiero, J. I. Ruzek, B. N. Smith, L. Trent, and B. V. Watts. 2016. A Review of Studies on the System-Wide Implementation of Evidence-Based Psychotherapies for Posttraumatic Stress Disorder in the Veterans Health Administration. *Adm. Policy Ment. Health* 43, 6 (November 2016), 957–977. DOI:https://doi.org/10.1007/s10488-016-0755-0

[33] Brian Shiner, Leonard W. D'Avolio, Thien M. Nguyen, Maha H. Zayed, Yinong Young-Xu, Rani A. Desai, Paula P. Schnurr, Louis D. Fiore, and Bradley V. Watts. 2013. Measuring use of evidence based psychotherapy for posttraumatic stress disorder. *Adm. Policy Ment. Health* 40, 4 (July 2013), 311–318. DOI:https://doi.org/10.1007/s10488-012-0421-0

[34] Gabriel Tan, Tam K. Dao, Lorie Farmer, Roy John Sutherland, and Richard Gevirtz.

2011. Heart rate variability (HRV) and posttraumatic stress disorder (PTSD): a pilot study. *Appl. Psychophysiol. Biofeedback* 36, 1 (March 2011), 27–35. DOI:https://doi.org/10.1007/s10484-010-9141-y

[35] U.S. Department of Veterans Affairs. How Common is PTSD in Veterans? - PTSD: National Center for PTSD. Retrieved October 3, 2021 from https://www.ptsd.va.gov/understand/common/common_veterans.asp

[36] Rui Wang, Min S. H. Aung, Saeed Abdullah, Rachel Brian, Andrew T. Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Michael Merrill, Emily A. Scherer, Vincent W. S. Tseng, and Dror Ben-Zeev. 2016. CrossCheck: toward passive sensing and detection of mental health changes in people with schizophrenia. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '16), Association for Computing Machinery, New York, NY, USA, 886–897. DOI:https://doi.org/10.1145/2971648.2971740

[37] Rui Wang, Weichen Wang, Alex daSilva, Jeremy F. Huckins, William M. Kelley, Todd F. Heatherton, and Andrew T. Campbell. 2018. Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1 (March 2018), 43:1-43:26. DOI:https://doi.org/10.1145/3191775

[38] Weichen Wang, Shayan Mirjafari, Gabriella Harari, Dror Ben-Zeev, Rachel Brian, Tanzeem Choudhury, Marta Hauser, John Kane, Kizito Masaba, Subigya Nepal, Akane Sano, Emily Scherer, Vincent Tseng, Rui Wang, Hongyi Wen, Jialing Wu, and Andrew Campbell. 2020. Social Sensing: Assessing Social Functioning of Patients Living with Schizophrenia using Mobile Phone Sensing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20), Association for Computing Machinery, New York, NY, USA, 1–15. DOI:https://doi.org/10.1145/3313831.3376855

[39] 2013. *Diagnostic and statistical manual of mental disorders* (5th ed.). American Psychiatric Association. Retrieved from https://doi.org/10.1176/appi.books.9780890425596

[40] Fitbit Development: Sensor Guides. Retrieved October 3, 2021 from https://dev.fitbit.com/build/guides/sensors/