Using Sensor-Captured Patient-Generated Data to Support Clinical Decision-making in PTSD Therapy

HAYLEY I. EVANS, Georgia Institute of Technology, USA
MYEONGHAN RYU, Georgia Institute of Technology, USA
THERESA HSIEH, Georgia Institute of Technology, USA
JIAWEI ZHOU, Georgia Institute of Technology, USA
KEFAN XU, Georgia Institute of Technology, USA
KENNETH W. AKERS, Georgia Institute of Technology, USA
ANDREW M. SHERRILL, Emory University, USA
ROSA I. ARRIAGA, Georgia Institute of Technology, USA

Today, clinicians have limited visibility into the quality of homework exercises that occur outside of the clinical context; however, understanding patient performance in these exercises is essential for guiding patient-centered care. To address this, we present the Clinician Homework Review (CHR), a unique measure and interface that displays similarity ratings calculated using sensor-captured patient-generated data (sPGD; i.e. heart rate, phone usage, ambient noise, and physical activity) for therapeutic exercises outside of the clinical setting within the post-traumatic stress disorder (PTSD) treatment context. Through concept testing sessions with 10 clinicians, we examine how sPGD can be leveraged to measure and investigate what contributes to patient performance in a therapeutic exercise. We also share in-depth information regarding clinician interpretation and planned use of data displayed by CHR in clinical sessions with patients. We frame our results in the context of situated objectivity and propose the notion of "perceived reference weight," which describes the significance attributed to contextualized data. In doing so, we support clinical decision-making in PTSD therapy.

CCS Concepts: • **Human-centered computing** → *Usability testing*; **User studies**; • **Social and professional topics** → *Cultural characteristics.*

Additional Key Words and Phrases: sensor-captured patient-generated data, clinical decision-making, mental health, post-traumatic stress disorder, veterans, exposure therapy

ACM Reference Format:
Hayley I. Evans, Myeonghan Ryu, Theresa Hsieh, Jiawei Zhou, Kefan Xu, Kenneth W. Akers, Andrew M. Sherrill, and Rosa I. Arriaga. 2024. Using Sensor-Captured Patient-Generated Data to Support Clinical Decision-making in PTSD Therapy. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 149 (April 2024), 28 pages. https://doi.org/10.1145/3637426

Authors' addresses: Hayley I. Evans, hayley.evans@gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, USA; Myeonghan Ryu, mhryu@gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, USA; Theresa Hsieh, thsieh37@gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, USA; Jiawei Zhou, j.zhou@gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, USA; Kefan Xu, kefanxu@gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, USA; Kenneth W. Akers, kakers@gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, USA; Andrew M. Sherrill, andrew.m.sherrill@emory.edu, Emory University, Atlanta, Georgia, USA; Rosa I. Arriaga, arriaga@cc.gatech.edu, Georgia Institute of Technology, Atlanta, Georgia, USA.

ACM 2573-0142/2024/4-ART149
https://doi.org/10.1145/3637426

1 INTRODUCTION

For the last century, the delivery of nearly all psychotherapies has been constrained by data collected from patient self-report and clinician intuition [26, 33]. This paradigm is especially problematic when the clinician is evaluating the patient's between-session homework assignments in real-world contexts. Clinical data collected with subjective and narrow methods functions as an ever-present obstacle in the practice, training, and delivery of cognitive behavioral psychotherapies. Clinicians who treat mental illness are in need of ecologically sensitive tools to collect objective and comprehensive data to efficiently track, assess, and respond to mental health needs throughout the treatment process [15, 31, 58]. Patients are in need of tools that provide feedback to optimize their therapeutic exercises based on the objective and comprehensive data [28, 90]. Sensor-captured patient-generated data (sPGD) can be used to address these issues.

Recent work has explored the utility of sensor-captured patient-generated data (sPGD) in the context of identifying and treating a variety of mental health conditions [6, 27, 40, 52, 89, 91, 92]. Studies have leveraged mobile phones or wearables to collect sPGD such as phone usage, sleep data, and physical activity throughout a patient's day-to-day activities. From this data, studies have been able to predict relapse [91], monitor behavioral patterns [89], and investigate how such data might begin to be incorporated into the clinical setting [63, 64].

Our study extends this research by investigating how sPGD might be utilized by clinicians who are treating veterans with PTSD in an intensive 9-day outpatient treatment program. A cornerstone of this program is daily patient therapeutic homework and the decision-making that drives the clinician's assessment of how well the patient engaged during the homework session. Presently, the clinician only has the patient's report of how the session went and their own intuition about the session report to drive the patient's treatment agenda. We present a concept test of the Clinician Homework Review (CHR), a unique measure and interface that supports clinical decision-making. It displays similarity ratings for therapeutic exercises performed under the guidance of a clinician and at home for homework. The CHR is calculated using sPGD of heart rate, phone usage, ambient noise, and physical activity. Using the CHR prototype as a probe we address the following research questions:

- How can we design an interface using sPGD to inform clinicians about patient engagement during therapeutic exercises?
- How will clinicians perceive the CHR's utility and usability?
- How can CHR be used to support clinical workflow?

To explore these questions, we host sessions with ten clinicians who specialize in providing therapy to veterans with PTSD. During these sessions, we gain an understanding of their practice, reveal the CHR concept through a conceptual video and interactive prototype, and gather feedback from clinicians via semi-structured interviews and surveys. Our research makes the following contributions. We provide one of the few shared interfaces that display sPGD to clinicians and patients in the context of mental health [76] for PTSD. We examine how sPGD easily collected through commodity smartphones and fitness trackers can be leveraged to measure and investigate which features (e.g., heart rate, physical activity, ambient noise, and phone usage) directly contribute to patient performance in a therapeutic exercise. We also share in-depth information regarding clinician interpretation and planned use of data displayed by CHR in clinical sessions with patients.

Our paper is organized as follows. First, we provide necessary background information on PTSD and the setting (i.e., therapy point of care) in which our research was conducted. We then review related work regarding developments surrounding technology created for a specific cognitive behavioral therapy (i.e., Prolonged Exposure [74]) and how sPGD has been used to infer mental health and inform measurement-based care. We provide a detailed description of our methods, including a description of the CHR prototype. We organize our findings to reveal the current state of evaluating homework exercises in therapy, how clinicians interpret sPGD as presented in the CHR prototype, and how they would integrate this data into their clinical practice. We then share design implications, discuss the notion of "situated objectivity" [63, 69] as it applies to sPGD in CHR, and propose the notion of "perceived reference weight," describing the significance attributed to contextualized data. We conclude by identifying the limitations of our research and providing ideas for future work.

2 BACKGROUND

2.1 Post-traumatic Stress Disorder and Treatment in Veterans

Post-traumatic stress disorder (PTSD) is a trauma- and stressor-related disorder characterized by trauma re-experiencing (e.g. intrusive memories and nightmares); avoidance of trauma-related situations, thoughts, and feelings; negative alterations in thoughts and mood; and hyper-arousal [23]. PTSD is a chronic condition marked by considerable distress and dysfunction [23] with many comorbidities, including substance abuse [66], domestic violence [34], suicidal ideation, and medical illnesses [41, 77, 85]. US veterans are disproportionately affected by PTSD as compared to the US population, 16 percent to 6 percent respectively [65]. Approximately 20 percent of veterans of the conflicts in Iraq and Afghanistan meet diagnostic criteria for PTSD [19, 97].

Recovery from PTSD is possible, but treatment is an intensive process [74]. Several pharmacological and psycho-therapeutic approaches have been used to treat PTSD [5, 74], yet, only 50 percent of veterans with PTSD seek care [67] and the treatment dropout rate can be as high as 68% [29]. The treatment with the best evidence for therapeutic efficacy is prolonged exposure (PE) therapy [8, 74]. PE therapy is delivered by clinicians through outpatient, intensive outpatient, and inpatient programs and requires guided work in the clinical setting as well as homework assignments in real-world contexts. Unfortunately, access and utilization of PE is low [73, 82]. One aspect of high-quality PE implementation is the significant volume of data exchanged between the clinician and the patient about therapeutic exercise engagement. These exercises include: 1) in-vivo exposure to real-world stimuli and situations that the patient usually avoids; and 2) imaginal exposure via the patient's narrative of their distressing trauma memories.

Our study focuses on utilizing sPGD for imaginal exposure exercises in an intensive outpatient program (IOP) setting. In order to understand the design of our system, it is important to understand the structure and expectations of the imaginal exposure exercises in this setting.

2.2 Imaginal Exposures in the IOP Setting

Our study is situated within a specific point-of-care setting at a nationally renowned mental healthcare facility that offers a two-week intensive outpatient program (IOP) for veterans with PTSD in the United States. The veterans receive PE therapy, the gold-standard treatment for PTSD [20, 48, 95]. On the first day, veterans are introduced to the PE therapy process including key tenants of therapy such as "subjective units of distress" (SUDs), a quantified, self-report measure to indicate their level of arousal. SUD is a 0-100 measure, where "0" means no negative affect, "50" means manageable stress, and "100" the maximum level of distress the patient has ever experienced, which is typically the level of distress the individual experienced when the traumatic event was experienced or witnessed. On their first day, veterans are also assigned to a cohort of veterans who are also undergoing therapy and to a specific clinician that they stay with for the duration of the IOP. During the next 9 days, veterans are on a fixed daily schedule that includes one-on-one PE therapy with the clinician, group therapy with their cohort, and individual PE practice. Each morning consists of an individual therapeutic session during which the patient revisits the traumatic event through imaginal exposure with the clinician. During these individual sessions, the clinician is responsible for ensuring the patient is optimally engaged with imaginal exposure through the use of instruction, modeling, and immediate and direct feedback. Accordingly, we constrain our definition of patient engagement in this context to refer to the level of attention the patient is giving to the therapeutic exercise they are performing. Between individual sessions and outside of the clinical setting, veterans are expected to listen to the recording of the morning's imaginal exposure. They are instructed to record their pre-SUDs before the exposure and peak- and post-SUDs during and after each imaginal practice on paper worksheets [24] or via a mobile application such as PE Coach [44].

Best practices of imaginal exposure homework include sitting still in a quiet environment without distractions, closing one's eyes, and imagining the narrative as described in the recording and for its entire duration. Clinicians will typically review the homework with patients in session the following day and ask them to report their experience, SUDs levels, and how well they were able to stay focused, still, and engaged. They may also periodically review progress over time by visualizing progress in an Excel Sheet, having patients chart SUDs by hand in a graph, or by pulling up homework sheets from various points in treatment together [24]. Patients are also given self-report assessments throughout therapy such as the PHQ-9 [42], PCL-5 [9], and other depression and PTSD measures, which are stored in the patient's electronic medical record and used by the clinician to assess progress. The clinician is tasked with evaluating the various types of patient self-report to make decisions about the quality of the patient's imaginal exposure homework as an indicator of the patient's ability to engage with their traumatic memories (i.e., therapeutic progress).

3 RELATED WORK

In this section, we review related work in three relevant areas. First, we describe existing technologies for trauma-based care, highlighting patient performance. Then, we discuss how sPGD has been used to infer mental health status. Finally, we share how sPGD has been leveraged for measurement-based care.

3.1 Existing Technology to Support Patient-Performance in Trauma-Based Care

A variety of technologies have been researched and developed for trauma-based care. They work either as standalone solutions that are self-contained programs with no therapist support, or supportive solutions that are complementary support for conventional therapy. Due to the uniqueness of psychotherapy, supportive solutions are optimal as they ensure therapeutic effectiveness by working with clinicians to optimize and customize care. In PTSD assessment and treatment, there are three main types of technological solutions: Internet-based treatments, virtual reality, and mobile health (mHealth) applications. They share a promise to reduce barriers to care such as PTSD-related avoidance symptoms, stigma, lack of proximal feedback and support, ambulation and transportation difficulties [55, 84], and overall costs [39].

Internet-based or telehealth treatments are either standalone or supported. While standalone telehealth is highly scalable with small to moderate effect in reducing PTSD symptoms [43, 83], the focus is usually limited to education and skills-training modules [12, 94] and attrition rates are higher than in-person [12, 94]. At the same time, standalone telehealth cannot allow clinicians to monitor and address inadequate patient performance. On the other hand, supported telehealth is found to be more cost-effective than traditional in-person treatment [57] while not diminishing its effectiveness [1, 2, 51, 56, 99] or increasing attrition [86].

Virtual reality has been used in trauma-focused treatments to help patients engage in therapeutic exposures with their traumatic memories, which are objectively safe yet avoided as if dangerous [81]. The efficacy of virtual reality exposure therapy has been found effective with veterans and civilians [22, 30]. However, the costs associated with obtaining the appropriate hardware, software, and clinical training are currently prohibitive to most clinicians, thus limiting the scalability of virtual reality as a means to enhance patient engagement [80].

mHealth applications have recently emerged as a more accessible means to enhance engagement in trauma-focused therapies. The most downloaded standalone app is PTSD Coach, which is presented as a tool to assist in PTSD management with information and learning coping tools [45]. The empirical evidence regarding its clinical utility is mixed, and the observed effect on PTSD symptoms reduction is small compared to conventional in-person treatment [32, 45, 53, 71]. Adjunctive apps are designed to help the patient engage in a conventional treatment where the therapist is tasked with monitoring and addressing patient engagement. The most downloaded adjunctive app for PTSD therapy is PE Coach [72], which has no clinician-facing version. It consolidates PE resources digitally in the patient's own smartphone with functionalities of therapy session recording, conducting in-app assessments, progress trackers, appointment reminders, and psychoeducation. Adjunctive apps remain an area ripe for innovation that goes beyond digitizing paper forms that can improve the patient's and therapist's ability to monitor engagement (e.g., multi-sensor systems) [24, 76].

Despite efforts to create technologies to support trauma-focused treatments, few of them are designed to integrate with clinical treatments [72]. With our design of CHR for PE therapy, we aim to begin to address this gap by creating an interface for both clinicians and patients with therapeutic effectiveness and patient accessibility in mind.

3.2 Inferring Mental Health Through Sensor-Captured Patient-Generated Data

Smartphone applications and wearable activity tracking devices have been used to collect a range of sensor-captured patient-generated data (sPGD) in the context of mental health conditions including schizophrenia [91], bipolar disorder [6, 11, 27, 52, 89], depression [11, 40, 92], binge eating [38, 88], and PTSD [63, 64, 75]. Different forms and combinations of sPGD have been collected during non-clinical activities to characterize current mental health status, predict future mental states, and support care management. We review recent studies and their usage of sPGD in the paragraphs below.

Mobile phone sensing systems have been used to accurately predict psychotic relapse in patients with schizophrenia [91, 93]. Wang et al., [91] 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 clinical assessments.

Sensing systems for bipolar disorder such as MONARCA [6, 27] and MoodRhythm [52, 89] aim to support patient self-care and mental illness management while providing relevant information to clinicians. To do this, they collect sPGD, such as accelerometer readings, location, phone usage, social activity, or physical activity. They combine this passive sPGD with self-report assessments of mood [6, 27, 52, 89] or other factors, such as sleep, medication adherence, activity, warning signs, cognitive problem, stress, and alcohol consumption [6, 27]. To track the depression dynamics in college students, Wang et al. [92] use the symptom features that proxy the symptoms defined in DSM-5 to predict whether they are depressed on a weekly basis. In an observational study [11], Braund et al. used mobile GPS for extracting circadian rhythm and found this was associated with social support and predicted changes in anxiety levels of patients with mood disorders including bipolar and depression.

sPGD have also been used in eating disorders research [38, 88]. Juarascio et al. [38] collected heart rate variability data using wrist sensors and emotional eating behavior reports to predict emotional eating episodes. In another study, a mobile app collected accelerometer readings, app usage, calls, and ambient light to detect contextual triggers of binge eating episodes [88].

Researchers have used sPGD for both exploratory and real patient PTSD care management with clinicians and patients. Ng et al. presented clinicians with wearable activity tracker data (e.g., sleep quality, calories burned) and asked them to imagine how this could be incorporated during PTSD therapy sessions [63]. Patients with PTSD participating in an intensive treatment setting were provided with Fitbits. Results indicate that veterans had distinct motivations to use the Fitbit including enhanced self-awareness and social interaction support. In contrast, clinicians thought that Fitbit data could help identify a relationship between PTSD symptom measures [63].

Our study builds upon prior research which has investigated the utility of sPGD for mental health conditions, particularly in an intensive treatment setting for veterans with PTSD [63, 64]. In contrast with previous studies, we combine both mobile app sensor data and FitBit data from at-home therapeutic homework exercises. To our knowledge, this is the first attempt to leverage sPGD to inform clinicians' sense-making of patient performance of therapeutic homework.

3.3 Utilizing Sensor-Captured Patient-Generated Data for Measurement-Based Care

In measurement-based care, clinicians utilize systematic tracking of symptoms and responses to treatment to guide clinical decision-making [50]. In the case of trauma-based treatments such as PE therapy, patients regularly provide self-report measures through the PCL-5, PHQ-9, and SUDs to capture metrics regarding patient response to therapy [25, 47, 61]. Yet, the validity of self-report is limited as it is inherently biased [26, 33]. Furthermore, in a recent study, Das Swain et al. [21] demonstrated that in predicting mental wellness with passive sensing data, psychological signals, such as language, better predicted self-reported anxiety and stress, whereas behavioral signals, such as sleep, better predicted stressful arousal. As past research has noted, sPGD demonstrates promise in contributing additional data from different perspectives to measurement-based care [63].

sPGD can be utilized to inform clinical decision-making from a perspective of situated objectivity [63, 69] and empower patients to take an active role in their own care [17]. The term, "situated objectivity," concerns the fact that subjective data can be affected by the context in which it is collected and interpreted by people with different backgrounds [69, 70, 87]. Prior research argues that patient-generated data should be tracked and interpreted within its internal and external contexts [63] to avoid potential bias. In this sense, clinicians can use sPGD to help their patients recall the circumstances under which the data was collected and facilitate their conversations with patients. This can drive clinical decision-making and patient-centered care.

There are a variety of potential benefits in incorporating sPGD into measurement-based treatment for mental health. sPGD provides the foundation for scalable behavioral pattern recognition for clinicians and patients [60]. It can enhance the continuity of care by providing transparency into the patient's life outside the clinic between clinical sessions with the patient's data [62, 96]. It can make health data more accessible [40, 63, 64] and identify trends in treatment [18, 63]. Furthermore, patient-generated data has been found useful in establishing trust between patients and providers and assisting patient-clinician communication at the point-of-care [54, 78] to build stronger patient-clinician relationships [17]. In past studies, researchers also reported cases where both clinicians and patients are aware of the benefits of sharing patient-generated data and willing to use it in the treatment [68, 75, 88].

Yet, there are also a variety of perceived challenges in utilizing sPGD in the point-of-care setting. First and foremost, interpretation [59, 68, 96] and utility [63, 64] of data by key stakeholders is a concern. The volume and types of data may not seem directly relevant for clinical purposes and may conflict with existing workflows [3]. For clinicians, patient-generated data in unfamiliar and non-standard clinical formats can be challenging to use within the clinical setting [96]. Patients may misinterpret data or have unrealistic expectations which can lead to fixation, discouragement [63], emotional triggering [59], or lead them to conclude they are not progressing adequately [17, 40]. For example, Ng et al. found that some veterans using Fitbit data in PTSD treatment felt the purpose of collecting the data was unclear and did not trust that the data was meaningful [64]. Second, the addition of supplementary health data can be time-consuming [18, 63, 79] and create new barriers in the patient-clinician relationship [79]. Finally, research has highlighted issues of interoperability with other key systems in the ecosystem (e.g. EHR) and concerns for data privacy [24, 60]. Patients tend to withhold information that they perceive as personal or sensitive [78], and this may be exacerbated in the context of mental health, particularly for insular groups such as veterans with PTSD who have heightened privacy concerns [24].

To date, studies have examined perceptions, benefits, and challenges of the use of sPGD for measurement-based care. However, none have yet introduced a prototype to elicit reactions from clinicians on therapy-specific contexts (e.g., therapeutic homework). Our study bridges this gap by conducting an in-depth investigation of how clinicians interpret and plan to utilize sPGD presented by the CHR prototype in reviewing imaginal exposure sessions with patients with PTSD.

4 METHODS

In this section, we discuss the CHR prototype design, study design, recruitment and participants, and data collection and analysis. The Institutional Review Board at the Georgia Institute of Technology approved this research.

4.1 Design of the Clinician Homework Review

The design of Clinician Homework Review (CHR) dashboard is meant to support imaginal exposure activities. The purpose of imaginal exposure is to activate and modify trauma-related fear structures in order to break the cycle of fear activation and avoidance. Through imaginal exposure, patients revisit the memories of the traumatic event by recounting them aloud in vivid detail with the clinician during the clinical session. The patient's homework is to listen to the recording of the most recent session's trauma narrative while sitting in a quiet place with no disruptions. The CHR is a novel measure that compares and contrasts patient sensor data during clinician-led imaginal exposure exercises and homework sessions (See Figure 1). The sensor data during the clinical therapy sets the baseline for "patient engagement" since the clinician is actively working with the patient to stay focused during this time.

In order to design the CHR, the first, sixth, and last authors hosted weekly workshops over the course of 12 weeks to gather a variety of perspectives on what were the optimal sensor streams to be gathered during imaginal session activities. The workshop was attended by members of the research team that represent both the computing and clinical perspectives as well as other engineering and design professionals at the Georgia Institute of Technology and Emory University. Exposure therapy exercises were explained in detail by the sixth author, who is a practicing exposure therapy clinician. The goal was to consider sensor data that could be gathered to correlate with 1) fear activation and 2) avoidant behavior during imaginal therapy. Since the goal was also to use commodity devices that could be readily available. The Fitbit activity tracker was chosen to gather heart-rate data which is correlated with physiological response during fear activation. The smartphone sensors were used to capture avoidant behavior. Each week the sixth author would provide an anonymized scenario of a patient's struggles (e.g., avoidant behavior) to complete their imaginal homework. Example challenges included walking or driving around, or listening to the trauma narrative while doing housework (e.g., vacuuming the house) instead of sitting in a quiet place. The group would utilize these scenarios to discuss how mobile phone sensors might be able to help record deviations from best imaginal exposure practices. They did so to identify data that would suggest that the optimal homework practice was not occurring. For example, steady heart-rate data (indicating no response to the trauma narrative) recording ambient noise (e.g., TV or vacuum cleaner) or utilizing the GPS tracker (that would show the path they took during the homework) during homework sessions. These ideas were debated and refined over time until the final version of the CHR sPDGs were determined. The anonymized scenarios were also leveraged to create the "dummy data" to ensure that the CHR visualizations were realistic. For example, it would be expected that the heart rate data would be high at the beginning of a session and would level off by the end of the session to indicate that the patient was habituating to the trauma narrative.

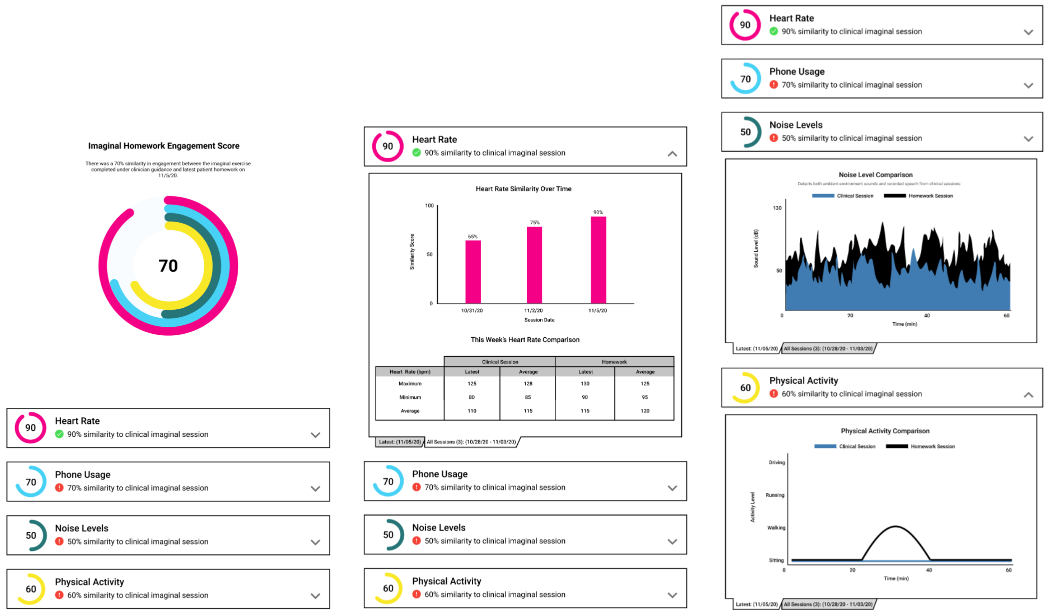
The final CHR concept was meant to fit into the IOP workflow. The clinician would ask the patient to use a mobile application and FitBit to record their imaginal exposure. When the app is turned on, various data streams including heart rate, physical activity, phone usage, and noise level in the environment are collected to establish a baseline level of engagement for optimal therapeutic progress Later, when the patient is ready to do their imaginal homework exercise, they turn on the app, and it collects the same data streams that were collected during the in-person clinical session. The app then compares the patient sensor data detected during the homework session to the sensor measures recorded earlier in the clinical session. These measures will be displayed in the CHR dashboard. Either before or during the clinical session the clinician can view the SPGD (i.e., comprehensive CHR score) and can use the data along with the patient self-report. This facilitates clinical decision-making about the current status of the patient's fear activation and avoidance as to drive an improved patient-centered care plan. Measures include:

- A comprehensive CHR score to give a high-level similarity rating that compares engagement during the clinical session to engagement during homework. This is an average of the scores of all data streams collected.
- Scores and visualizations for each data stream collected.
- Averages from all sessions within each data stream.

To reiterate, CHR collects and displays comparative data regarding patient engagement in imaginal exposure exercises throughout exposure therapy. Clinicians can use this data to inform their clinical decision-making about the delivery of PE therapy and guide the patient to optimal performance. The patient gets feedback on their performance that allows them to optimize the next homework exercise. Patients then get additional information that confirms what they did well and where they can improve. In short, the goal of CHR is to support clinician delivery of patient-centered care.

The mid-visual and functional fidelity prototype of the CHR interface was designed using Figma based on how patient imaginal exposure exercises are designed, assigned, and evaluated within the prolonged exposure therapy [74]. Clinicians were able to view the same "dummy data" that was built into the prototype and click through the prototype in the think-aloud exercise. We also reflected the findings from a previous study [24]. The prototype design was iteratively improved reflecting the feedback from our clinician co-author. The prototype is available at: [Link].

The overview page (see Figure 1a) of the prototype utilizes sensor data, as suggested in prior work [24], to help clinicians assess patient's practices outside of clinics. Based on that data, the system calculates a score to indicate patients' engagement during their homework, compared to the scenarios under clinician guidance. By clicking into each data category, clinicians can further review the data in detail and see their changes over time (e.g., Figure 1b shows how patients' heart rates compared to clinical practice throughout days), as well as a high-level summary of the data (e.g., Figure 1b shows the statistics comparing patients' heart rate data during homework and in clinic). To look at specific data points, clinicians can click on the "latest" tag. This will present the comparison of how data changes alongside the time on a minute scale (Figure 1c).



(a) Overview page: similarity ratings of collected data streams.

(b) Comparison of all sections: heart rate.

(c) Latest session: noise level and physical activity.

Fig. 1. CHR prototype dashboard.

4.2 Recruitment and Participants

We presented the CHR concept prototype to 10 clinicians and received their feedback. All clinicians were employed by the same mental healthcare facility and elected to participate after receiving an internal email sent on our behalf. Participants were compensated with a $20 Amazon gift card for their time. Table 1 provides a summary of participant information. 70% of our participants were female, which aligns with 69% of clinical psychologists workforce being female as of 2021 [4].

4.3 Data Collection

All concept test sessions were conducted by the first and third authors via video call lasting between 40 - 60 minutes. First, clinicians shared their experience collecting and evaluating patient self-reports for imaginal exposure exercises throughout exposure therapy in a semi-structured interview. Then, they were shown a short video explaining the CHR concept. Afterward, they participated in a think-aloud session [35] with the interactive CHR prototype in which they reviewed the CHR sensor-captured patient-generated data for an imaginary patient. This prototype did not contain real patient data, and mock data was used to create tables and visualizations to ascertain if it was "realistic." They were asked to share initial impressions of the prototype as well as to determine what CHR told them about the imaginal exposure exercise, which area(s) of the homework (if any) required clinician attention and how, and the time frame of the available data. After concluding the think-aloud tasks, clinicians were then asked about their experience interacting with the prototype in a semi-structured interview. Clinicians were asked to explain the CHR concept in their own words, how (if at all) they saw themselves using CHR as a part of their exposure therapy practice, what data (if any) they believed patients should have access to, and when. Clinicians were also

Table 1. Participant information from 10 clinicians. Expertise using technology (software and hardware) is self-rated on a scale of 1 to 5, with 1 being the lowest and 5 being the highest. System Usability Scale (SUS) score [13] is used to measure usability.

ID	Clinical psychology practice	Exposure therapy practice	Rated expertise with technology	Technology used in clinical practice	SUS Score
C1 (F)	7 (years)	3 (years)	4	e-Sense, PE Coach, Computer for telehealth, Zoom, MS Office Programs, audio recordings	87.5
C2 (F)	35	35	3	Computer, VR, Zoom, Mobile Apps, EMR, Online Portal	77.5
C3 (M)	5	3	4	Laptop, smartphone, video calling, Youtube videos, file sharing, audio recordings, Box for file sharing, Quicktime for audio editing, Zoom	97.5
C4 (F)	5	2	3	CBT Apps like PE Coach, OneDrive, Zoom, EMR	82.5
C5 (F)	13	8	3	Zoom, Email	60
C6 (F)	13	11	3	EMR, e-Sense, VR, mobile apps, Youtube goggles, Excel	62.5
C7 (M)	6	3	3	Video and audio recordings, Zoom, email, EMR, Salesforce, OneDrive, e-Sense	50
C8 (F)	6	0.75	4	Mobile apps, electronic documents, video recording, e-Sense	77.5
C9 (M)	8	8	4	e-Sense, mobile apps, video recordings	85
C10 (F)	20	20	4	Mobile apps, electronic forms, tablets, teleconferencing software	100

given an opportunity to provide open-ended feedback. Finally, they completed a system usability survey to evaluate their experience with the prototype.

4.4 Data Analysis

All sessions were transcribed by a third-party service and analyzed line-by-line by the first author utilizing thematic analysis [10]. An initial set of 10 deductive codes were generated utilizing key findings from Evans et al. [24] and Ng et al. [63] (e.g. "Clinician and patient relationship," "Self-Report," etc.) and another 10 inductive codes emerged from reviewing the transcripts (e.g. "Clinician Knowledge of Outside World," "Data Access," etc.). After the initial codebook was developed, the authors met as a group to review the codes and various samples presented by the first author. These were then refined as the group reviewed portions of the transcripts. In this first round of coding, the 20 codes were applied to the 10 transcripts 447 times by three authors, and two deductive codes were removed for lack of relevance to the data via group discussion and deliberation. The second and third rounds of coding were performed to further refine and apply the code book. A total of 16 parent codes (e.g. "Self-report") and 24 child codes (e.g. "Self-report quality," "SUDs calibration") were applied 869 times. In the final step of data analysis, the first author reviewed the findings with two of the clinician participants who volunteered to verify that the interpretations were representative of their experience. From this data analysis, the group found themes related to current clinical practices in evaluating imaginal exposure exercises, interpretation and imagined utilization of CHR, and possible barriers to the use of CHR.

5 FINDINGS

We organize our findings into three sections. First, to establish a baseline for understanding, we describe current practices in collecting and assessing self-reports, clinician workflow, and clinical decision-making for imaginal exposure exercises in PE therapy. Next, we share findings regarding clinician interpretation of the utility and usability of sPGD presented in the CHR prototype. Finally, we describe how clinicians report they would utilize this information in the therapeutic setting.

5.1 The Reality of Conducting Imaginal Exposures in PE Therapy

Conducting imaginal exposures in PE therapy is a specific procedure that includes many small clinical decisions informed by theory-driven case conceptualization [74]. However, we found that in delivering patient-centered care, the clinician practice extends beyond the PE guideline. Accordingly, it is essential that we describe how clinicians guide the imaginal exposure exercise in session, what expectations they have for at-home practice, and how this practice is evaluated in the following in-person clinical session. These insights can be especially valuable in guiding future design iterations of the CHR.

5.1.1 Guiding Imaginal Exposure Exercise In Session. In each session, clinicians guide patients in conducting an imaginal exposure exercise. These generally last 60 - 90 minutes, and the clinician offers guidance throughout to help the patient effectively articulate a narrative of the traumatic event with their eyes closed. They recommend that patients use a smartphone application (e.g. PE Coach) to record the session while their phone is in airplane mode. Clinicians have patients share their pre, peak, and post-exercise subjective units of distress (SUDs).

In early sessions, clinicians spend time instructing patients on how to calibrate their SUDs ratings. Clinicians help patients set anchor points from 0 to 100 that patients and clinicians will use to assess SUDs scores throughout therapy. Patients express what experiences would produce a 0 (no negative affect), 50 (manageable stress), and 100 (the most distressed you've ever been). C1, C4, and C10 explained that this is a highly individualized process and, subsequently, there is a lot of variability in how patients use SUDs. Clinicians must use their intuition to interpret this type of reporting and guide patients on how to best use SUD scoring. C1 explained:

> "[In speaking with a patient I'd say], I noticed that you said you were about a 70 or 80 in the imaginal exercise] but you didn't seem like you were having a lot of physiological response. Can you tell me a little about that? [Or maybe I have someone who under reports] and they tell me they are at a 30 but I'm noticing they are sweating. I normalize that this is something really different than they are used to doing and its not just about noting SUDs at any given time but allowing themselves to experience those emotions." (C1)

Clinicians look for emotional engagement with the imaginal exercise rather than avoidance of the memory. This includes the cognitive part (e.g., "this is dangerous"), the behavioral part (e.g., an urge to leave the situation or look away), and the physiological part (e.g., an increase of HR). They look for physiological reactions to the exercise such as sweating, flushed cheeks, breathing more heavily, speaking more quickly, and changes in body language. Three clinicians also utilized a device called eSense to measure galvanic skin response (GSR). Clinicians also seek to identify avoidance or safety behaviors during the imaginal exposure which might appear as any number of reactions as they are highly individualized but might include actions such as fidgeting, re-enacting, wringing hands, or hiding their face. C8 shared:

> "We're doing the imaginal exposure and I notice that [the patient] hides their face. If that's the first time I see it, then usually once we are done with the exercise, I name

> that behavior and we talk about it as a safety behavior. I ask them why they did it. I also would call it out during the exercise and ask them not to hide their face. And then I know it's a safety behavior they might engage in. I would follow up on it later [in session and homework.]" (C8)

Clinicians rely on their intuition and ability to customize the PE therapy protocol to meet unique patient needs. Clinicians shared several instances in which bending the protocol may occur. Specifically, they discussed over-engagement, which is characterized by experiencing an acute level of distress that impacts the patient's orientation (person, place, and time) and ability to identify their immediate environment is indeed safe. For example, clinicians may have patients open their eyes during the imaginal to reduce engagement if they are over-aroused, allow patients to act out memories, or even move around, if appropriate. C7 explained that clinicians who are well-versed in PE therapy are able to do this effectively. They said:

> "Sometimes you bend a little bit, especially a [clinician] who knows the protocol. They might bend the protocol to achieve what's necessary. So for example, the person is very over-engaged in their session, and during the imaginal, you say, okay, just do this with your eyes open and then we'll get to the appropriate amount of engagement as we go through the course of treatment." (C7)

Lastly, clinicians explained that the COVID-19 pandemic, which shifted their practice to fully virtual sessions, meant that they would have less control over the in-session imaginal exposure and potentially limited visibility into avoidance behaviors that could occur outside of the camera screen. For example, they might not notice hands wringing or a leg shaking. However, C9 shared that despite the limitations of telehealth, there is still more control if your therapist is involved and guiding a patient than if they are not present at all.

Clinicians provide guidance in session for imaginal exposures for optimal practices which are individualized on a patient-by-patient basis for that moment in time in their treatment whether in-person or via telehealth.

5.1.2 Expectations for At-Home Imaginal Exposure Practice. Following each in-session imaginal exposure exercise under clinician guidance, patients are expected to listen to the recording from that day by themselves for homework. In the context of the intensive outpatient program (IOP), the patients may complete this exercise while at the facility during downtime or later from their hotel room. However, during the time of this study, veterans enrolled in the IOP were participating virtually due to the COVID-19 pandemic. As such, they were completing their homework in their normal home environment.

First and foremost, clinicians expressed that they expected veterans to listen to the imaginal exposure for the full duration of the recording. Clinicians instruct patients to complete this exercise in a private space with the phone in airplane mode. However, clinicians said this was more easily controlled in an IOP setting and took a more pragmatic approach to homework expectations during the pandemic. They noted that some exceptions might need to be made based on the patient's real-world situation. C7 explained that some patients might need to use headphones so that others in the household did not hear the recording. C1, C3, and C4 said that finding a quiet, private space at home might be challenging or impossible. C3 said:

> "I pick my battles. Ideally, no, [the patient shouldn't walk] but if that's the only time they can get away - they've got a busy house, quarantine, kids, working from home, they don't have a space that's quiet to sit down and do this, I'd be okay with that. It's not ideal, right? I'd tell them, "lock yourself in the car" first, but if [walking] gets them

> to do [the homework] and assuming that it wasn't avoidance...getting them to engage is the priority." (C3)

Clinicians also instruct patients to note their pre-, peak, and post-SUDs just as they would in session. Patients may use a paper worksheet, a fillable PDF that can be sent virtually to the clinician, or PE Coach to note their SUDs. C2 said that compliance on filling out the SUDs measures for homework was probably about 50/50 and only because they emphasized the importance and actively used them in session. If a patient forgot, C2 allowed the patient to fill them out retroactively at the beginning of the session. In contrast, C6 explained that they didn't place much emphasis on completing SUDs measures for homework and rarely saw worksheets as they had in the clinic before the pandemic.

Clinicians aim to have patients complete their imaginal exposure at home as similarly as possible to the imaginal exposure completed in a clinical session. However, they have a realistic understanding of what this might look like and adapt their expectations accordingly so long as the patient is engaging with the memory.

5.1.3 Evaluating Imaginal Exposure Homework. Clinicians expressed a sense of underlying trust in the clinician-patient relationship. They predominantly believed their patients regarding whether or not they had completed the homework exercise and to what degree. Generally, clinicians felt homework compliance was high, and as C4 and C5 pointed out, patients are forthcoming if they have not completed it. On occasion, a patient may not be honest about homework compliance but this becomes obvious quickly. C3 explained that patients are not able to easily lie about engaging in the exercise. C3 said:

> "[If they haven't done the homework and say they have,] there's a lack of depth. They can't really talk about it, right? It's like when someone's lying about something, it's either very shallow or overly detailed in sort of the wrong way." (C3)

Clinicians engage patients in in-depth conversations around the homework to better understand and tailor the patient experience. Often, clinicians use SUDs as a way to initiate conversation around homework. They probe for additional information and seek descriptive language to explain when, where, and for how long they engaged in the imaginal homework exercise. They ask the patient about any avoidance or safety behaviors that may have taken place, particularly those that were identified in previous sessions. C1 noted that other avoidance behaviors may exist, but it can be difficult for a patient to identify these alone as they may be habitual actions. Clinicians also prompt patients to answer questions regarding the effort it took to engage in homework and what has changed since the last time they performed the exercise. These questions help the clinician understand the patient's progress and how to continuously shape optimal practice. C10 explained that engagement in imaginal homework early on may look different than later in the program; they may allow a patient to act out a memory (e.g. holding up an imaginary gun) during the imaginal early on to increase engagement.

> "In early sessions, sometimes people act out parts of the memory and that can be a sign of engagement, but as a therapist, I would also be talking with them about doing that less over the sessions...staying with the memory and re-experiencing but not re-enacting. It depends where we are in the sessions whether that's a good or bad thing." (C10)

Adherence to a strict homework regimen is not how the interviewed clinicians practiced PE therapy for imaginal homework exercises. Instead, the clinicians seek to better understand the experience of the patient by gathering self-reports, tailoring the treatment accordingly, and continually evaluating the patient's habituation to and progress in therapy.

5.2 Interpreting sPGD in the CHR Interface

Next, we share findings regarding clinician interpretation of the utility and usability of CHR prototype with sPGD data streams and visualizations. We report general feedback clinicians provided on the prototype's utility and usability. This is followed by an in-depth analysis of each data stream collected including heart rate, phone usage, physical activity, and noise level.

5.2.1 General Utility and Usability. Overall, clinicians found the CHR prototype to be useful and usable for informing their PE therapy practice for imaginal homework exercises. After interacting with the prototype, all ten clinicians filled out a survey to provide a system usability rating (see Table 1). On average, the CHR prototype received a SUS score of 78, with seven clinicians providing a score of 77.5 to 100. Three clinicians provided a score between 50 - 62.5. All three individuals who rated the CHR at this lower level also rated their expertise with technology at a 3 (out of 5). Those who rated their expertise at a 4 (out of 5) tended to provide higher scores, suggesting CHR may be a more attractive concept to clinicians with more technological expertise.

Clinicians unanimously understood the CHR concept correctly and were able to articulate its purpose and functionality at a high level in their own words. Before clicking into the individual data streams, they understood that the rings represented a similarity rating and that the number in the center of the rings provided an average of all data streams. Clinicians identified two areas that presented a challenge in interpreting the data. First, they wondered what an ideal and/or realistic similarity rating was. Second, they were unsure of what the individual data streams might entail but were able to bridge this gap in knowledge once they interacted with the prototype over time.

About half of the clinicians indicated that visualizing patient habituation (e.g. reductions in SUDs from session to session) to therapy over time might be more useful than general averages.

5.2.2 Heart Rate. Clinicians shared that heart rate provided useful information regarding patient performance in homework. Six of ten clinicians felt they would address heart rate with the patient while one stated they would use it for their own knowledge. They felt collecting comparative heart rate data over the course of all sessions could potentially demonstrate the body's physiological response to PE therapy and demonstrate progress over time. In reading the interface, most clinicians focused on the heart rate activity comparison graph showing heart rate (bpm) against the time of the imaginal exposure exercise. C3, C6, and C7 articulated that when reading the graph, they were looking for similarities in heart rate peaks around the same points in time to determine patient performance. C6 shared that heart rate would be a useful indicator for her as a clinician to determine habituation and would inform how she shaped therapy, but that she would not discuss this with a patient from session to session. In contrast, C9 felt this was the only piece of the CHR prototype that they would continually visit with the patient as it was singularly telling in terms of demonstrating habituation. C4 felt similarly and said that heart rate data could be particularly useful visualization for patients who are struggling to engage in treatment.

> "Heart rate can be one indicator of how emotionally engaged they are, especially if this is a patient that we're seeing a lot of avoidance and that they're having trouble emotionally engaging with the memory after we've done a few repetitions." (C4)

Clinicians did question what heart rate might look like in real patients and how this might change their view on the utility of the data stream. C1, C2, and C7 questioned what heart rate might be able to show in terms of patient performance. C1 questioned:

> "[For homework the heart rate data is] a little bit lower, so that looks theoretical but it's a good theory [of what could happen]." (C2)

C3 questioned whether the Fitbit would provide high enough quality data to conduct heart rate variability analysis, though, they noted this was likely outside of the concern of an average clinician.

C8, C10, and C7 acknowledged that heart rate had some utility but would prefer a galvanic skin response (GSR) data stream.

5.2.3 Phone Usage. Clinicians were able to easily interpret the amount of calls, number of contacts, and time spent calling, texting, and utilizing apps. In our prototype, we displayed this data to demonstrate the patient was engaging in all of these behaviors during homework. Perhaps unsurprisingly, nine of ten clinicians stated they would use this data stream as part of their practice. C4 and C9 explained that they imagined this data would be particularly helpful early on in treatment as they could catch and correct it early. C1 believed collecting and visualizing this data could open doors to behaviors not previously detected by the patient. He said:

> "Sometimes I think a call or a text comes in and they don't think about it; it's just kind of natural to [respond]. And so that would be kind of a learned habit that we can talk more about and work on not engaging [with calls or texts] with when they're doing the homework. I think some of this data is information that maybe they wouldn't talk to you about maybe when you're asking about how their homework went." (C1)

However, six clinicians were confused about the functionality of the phone usage data stream. All six assumed the phone would be in airplane mode as is standard in their current recommended instructions for imaginal homework exercises in-session and for homework (even when using an application like PE Coach). As a result, they supposed that the patient had not complied with that request. C1 and C7 also questioned where the indicators for a paused recording were on the graph given that phone calls had taken place. C7 thought that perhaps it also picked up apps running in the background of the phone rather than those in active use. C3, C7, and C10 questioned whether or not phone usage included usage of the CHR mobile application itself. C3 shared a statement that demonstrates this confusion in interpreting the similarity score and data stream details:

> Phone usage would be a big one [for my attention] if [it] was high. If similarity, I guess, was low, that would be a big concern for me. (C3)

While phone usage received a strong reaction from clinicians, interpretation was variable and clinicians struggled to understand what phone usage during the imaginal *really* meant.

5.2.4 Physical Activity. Seven of ten clinicians indicated that they found physical activity to be a useful source of information for patient performance. C1, C2, C6, C7, C8, and C10 indicated that this could provide some additional context to homework performance. It might show that a patient was engaging in avoidance behaviors. C1 said:

> "Walking for a few minutes might mean that they were also engaging in other things like chores or things around the house." (C1)

However, C3, C6, and C7 explained that while physical activity was not ideal, it might be the only way that a patient was able to complete the homework. C3 and C6 said that walking might be okay, especially early on in therapy, so long as it was not avoidance behavior. Clinicians stated that walking was preferable to driving as they should be devoting their attention to the homework to properly engage. C7 said:

> "I have had patients say, 'I went on a walk and I listened to my imaginal.' And I'm like, 'I would prefer that you did not do that.' But I do also feel like sometimes that may be a thing that comes up, especially with COVID and people working from home, [sometimes] there no other place." (C7)

After reviewing this data stream, five clinicians asked if it was possible to see smaller movements they typically address (e.g. fidgeting, wringing hands, shaking legs). C4 and C6 initially mistook the physical activity graph to depict data representing these smaller movements. C4 said:

> Within sessions [I] will talk with patients about how much they're fidgeting,if I notice a leg shaking, if I notice they're covering their face or moving their hands a lot, fixing their hair a lot. So this is really nice that this tracks it. And even within session, patients don't always know in session how much they're moving around. They're so in the imaginal that unless I point it out to them, they don't know, so this would be a really nice way to have some objective data saying, 'look, you're moving around quite a bit. What's going on?' We really want your activity level to be pretty low. We want you sitting still, hands on your lap. I think that could facilitate a conversation about what's happening in session, but then also having a peek into what's going on during their homework session, too." (C4)

Clinicians did find some utility in physical activity but noted that movement may be okay for some patients, especially early on. Several clinicians desired the ability to see smaller movements.

5.2.5 Noise Level. Noise level in the environment received mixed reviews from clinicians. Six stated that they found the data stream useful but noted it had less utility than other data streams collected. C1, C3, C4, and C10 explained that they would assume the noise levels in the homework environment would almost certainly be higher than in the clinical environment, which is abnormally quiet. C10 said that they expected a higher noise level, particularly in earlier sessions as the patient was acclimating to best practices in PE therapy. C1 and C9 pointed out that noise levels could be especially relevant for telehealth as it might help manage the clinicians' expectations around what was realistic for both the imaginal exercise completed under their guidance and for homework. C9 said:

> "Noise levels is a cool one because you would expect not to have any noise problems inside a therapist's office. This is interesting to think about in terms of telehealth. So you know, now that we do so much over telehealth, I guess we'll see a lot less controlled in-session things [like noises in the background]. But I'd still expect more control if your therapist is actually watching you [than for homework]." (C9)

Seven of ten clinicians had some confusion about the utility and functionality of noise levels. C2, C6, C7, C8, and C9 struggled to understand what noises the sensor stream was detecting; was it the recording or just ambient noise? How did it work if the patient was wearing headphones? What happens if there is a loud noise on the street while the data is being collected? C1 raised a concern of particular relevance to veterans who have a heightened sense of privacy concerns. He expressed confusion over how the app collected noise, in what format, if it was associated with the patient name, and how it was stored. C6 and C7 questioned what an optimal noise level really was inside and outside of the clinical setting. C5 and C6 weren't sure how to interpret a similarity rating between the imaginal exposures. C5 said:

> "I'm thinking about what information I don't have that could be the attribution for the higher level of noise during homework versus in the PE session. But either way, so what? I really don't know what to draw from that." (C5)

While noise levels could offer some insights into patient performance, clinicians stated that they might already expect a difference in noise levels. Furthermore, they had several questions regarding the functionality of noise levels that required clarification.

5.3 Integrating CHR sPGD into the Review of Imaginal Exposure Homework

Here we describe how clinicians would address the sPGD data presented by the CHR prototype in therapy. We also discuss clinician perspectives on patient data access and barriers to use.

5.3.1 Providing Feedback to Patients. Clinicians indicated that they would utilize the data presented by CHR in conjunction with their already existing processes for evaluation such as SUDs, PCL-5, verbal reports, and clinician intuition. C2 said:

> "If their SUDs aren't going down, if they don't seem to be making improvement, I'm going to look at this and I'm going to say, I think this is why. I think you're not really engaged in your imaginal exposure, you're doing other things, so that's why it's not getting better. So just use it as another bit of data." (C2)

The majority of the clinicians said they would integrate CHR data into their already existing flow for evaluating homework exercises. They would ask general questions about how the patient felt the homework went, what changes they might have noticed, and if there were any distractions. By asking these open-ended questions, they would allow the patient to respond with or without the data or go through the data together and discuss the best way to move forward. C10 shared:

> "I would probably start with a more open-ended question about how did they feel like their homework went. And if they didn't bring up that they were distracted and doing other things on their phone while they were supposed to be listening, then I would say, oh, well, in looking through CHR, it looks like there were some interruptions during your exposure. Let's talk about that. What's going on? Why were you getting calls, and then, getting more specific? And from looking at all that, I wouldn't want to nitpick 20 different things with the patient. I would kind of try to put it under, if it's appropriate to put it under one umbrella of, let's find a quiet place next time where you can be by yourself and not interrupted, you know, turn off the getting calls and getting notifications while you're doing this. And you know, set in giving some more specific instruction in how to make a quiet place where they can focus just on doing their imaginal exposure." (C10)

However, there was a sense among four clinicians that using this data could introduce the risk of a patient feeling criticized. While they felt the data could be useful, discussing it with a patient has to be a balance between providing useful feedback and ensuring engagement. This would have to be determined on a patient-by-patient basis. C9 explained:

> "I do think that one downside to this could be that we don't want to be like 'Hey, you were walking around during your imaginal!' We don't want to wag our fingers at patients. But we do want to make sure they're engaged. So we need to balance those two things." (C9)

Clinicians indicated that many of the sensor streams would be best addressed early on in treatment, particularly when the patient was more likely to engage in some of these distractions and when they were still in the learning phase. Afterward, clinicians envisioned that CHR, or parts of it (e.g. heart rate) could be visited periodically.

5.3.2 Data Access. There was variability in opinions surrounding data access. Clinicians had distinct opinions on what data they would allow patients to see and when. Three clinicians (C1, C2, C3) believed patients should gain access to all of their CHR data immediately after finishing the exercise at home. That way, they could review it and bring it into the session ready to discuss it. Half of the clinicians (C4, C6, C7, C9, and C10) believe that patients should gain access to all of their CHR but only once they are in-session with the clinician. They worried that patients might not understand how to interpret the data or might fixate on it without clinician guidance. C8 believed that patients should have access to all of their data but when should be determined by the needs of the patient. They said:

> "If someone is very prone to worry or rumination that might not be someone that I would give access to right away. I don't think it would be harmful; I just think that it could be something that they look at and then really think about a lot, and it might take away from the main work, which is just doing the imaginal exposure." (C8)

The only clinician who did not believe patients should have access to all of their data was C5. She shared that they would share data only if it was noteworthy. She said:

> "If I'm reviewing the data and there is nothing in particular that jumps out at me, I'm not sure I would find a clinical reason to share it with the patient. But I might do that if there's something that I want to talk about with the patient." (C5)

In general, clinicians believed patients should have access to all of their CHR data, but *when* they ought to receive access varied.

5.3.3 Barriers to Use. A variety of barriers arose in our conversations with clinicians. First, and perhaps most obvious, is the time burden of sPGD integration. It is one more demand on top of an already demanding therapeutic process. C7 said:

> "I don't know is [this] a discussion that takes two minutes? Or is [it] a discussion that's now all of a sudden taking 30 minutes and pulling away [from the rest of the session]?" (C7)

Furthermore, as insinuated in the data access section, clinicians are not confident that patients will be able to interpret this data on their own. C3 expressed:

> "Patients might not know what to do with all this information." (C3)

Concerns about data privacy, particularly for the veteran population, were mentioned by three clinicians. They worried that this technology might feel overly intrusive and that veterans might have concerns over what was collected, how it was stored, and for how long. C1 said:

> "[The veteran] population is even more sensitive to privacy concerns, so they would probably want to know a little bit more about, you know, what is recorded as far as sound [for example]." (C1)

There was also concern from one clinician regarding how data collection through the phone and FitBit might be cost-prohibitive for some patients. C1 questioned how the volume of data collected might not align with what their smartphone plans allow.

6 DISCUSSION

Through our investigation of the CHR prototype dashboard, for supporting clinical decision-making, we contribute to the growing body of literature that examines the use of sensor-captured patient-generated data (sPGD) for clinical mental health [7, 27, 38, 88, 89, 91, 93]. Previous research has investigated the usefulness of incorporating passive mobile sensing systems [7, 11, 27, 89, 91, 93] and wearable devices [64, 92] to measure physical activity, phone usage, sleep quality, and more to indirectly assess mental health conditions. In measurement-based care, these practices have been shown to improve patient-clinician communication and establish greater levels of trust [54, 78]. Ng et al.'s application of FitBit tracking data to intensive trauma-based care for veterans with PTSD suggests that passive sensing systems can improve patient self-monitoring habits [64]; however, when examining reasons for non-use, veterans shared that they did not see a clear connection between the FitBit tracking and its relevance to therapy goals. Similarly, West et al. [96] pointed out the misalignment of objectives as one of the barriers for clinicians in using patient-generated data in clinical settings. In the design of CHR, we extend the research of the application of sPGD in clinical mental health by directly measuring imaginal exposure exercises. Thus, we make explicit connections between sPGD and therapy practice. In presenting our CHR concept through an interactive prototype, we also gather specific feedback from clinicians about how they might use such data. This contributes to a growing literature on clinical decision-making in situ. It reveals how visualizing such data can help give clinicians insights into treatment trajectory [76] and patient outcomes [96]

The use of sPGD is one of many efforts to include additional perspectives in creating a more holistic understanding of mental health status for veterans with PTSD in clinical therapy [7, 24, 64, 75]. This effort echoes previous findings that consider PGD or self-tracking data as context boundary negotiating artifacts [46] by which patients' life expertise and clinicians' expertise in clinical practice are mediated [18]. The automated nature of sPGD [16] requires minimal patient engagement and has thus been considered to be more objective than patient self-report. However, as Ng et al. point out while sPGD is free from some of the biases of self-report, it is not fully objective [63]. Instead, they suggest that it should be viewed as "situated objectivity" [69] which combines the notion of 1) "mechanical objectivity" in which sensor data creates a frame of reference for interpreting self-tracking information and 2) "trained judgment" in which experts are relied upon for contextualization of knowledge formation. They highlight how clinicians play a critical role in reframing and refocusing data interpretation for patients [63]. The current findings highlight this interplay. It shows that a patient's imaginal exposure and SUD ratings are highly individualized and require a clinician's interpretation within the temporal resolution of the PE intensive outpatient program (e.g. day 1 versus day 9). Clinicians adapt the PE therapy protocol to the patient's progress or setbacks during the nine-day stay. Thus, clinicians expressed an interest in leveraging multiple streams of sPGD (e.g., physical activity, noise level) as contextual information to better understand patients' self-reported imaginal exposure homework practice. This study demonstrates a use case for "situated objectivity" by showing how the CHR's multiple streams of sPGD help the clinician to contextualize the patient's self-report and to guide the next steps in the therapy. Still, depending on different scenarios (e.g., making sense of sensor-detected noise data), more contextual information (e.g., location, presence of others) might be needed to assist the clinician in the interpretation of the data.

Since figuring out a patient's avoidance behaviors during imaginal exposure exercises is a crucial part of Prolonged Exposure (PE) therapy [74], we speculate that CHR can lead to a more holistic understanding of patient performance. This may help facilitate discussions between clinicians and patients and drive patient engagement. Furthermore, adding a patient-facing homework review dashboard might improve patient accountability [17, 18] and sense-making about the PE homework since they would now have objective contextual, behavioral, and physiological data at their disposal. This suggests that future research should investigate the impact of collecting and communicating about sPGD in this context of patient-provider collaboration in PE with studies involving both providers and patients.

In considering the interpretation of sPGD for clinical decision-making, it is crucial to consider the experience level of the clinician. Our study interviewed PE practitioners whose extensive experience enables them to anticipate the needs of patients and make clinical decisions that individualize the practice of PE therapy to meet those needs. Individual clinician practices may then be reflected in the sPGD collected for each patient. While our study focused on the individual patient, aggregate forms of sPGD from multiple patients will require additional investigation to uncover meaningful patterns. Conversely, novice clinicians may require more support than what is currently presented when interpreting CHR data. As pointed out by one of our expert clinicians, it is possible that a patient's at-home practice may outperform the clinician-led session when the clinician is a novice. While this is an edge case, the current design of CHR assumes that the patient performs their therapeutic activity better under the guidance of a clinician. Considering the skill of the individual interpreting the data leads us to an extension of situated objectivity. Specifically, the latter takes into account the non-systematic way in which experts contextualize self-tracking data. Thus, "perceived reference weight" describes the significance attributed to contextualized data for making clinical decisions based on sPGD. For clinicians, when expertise is high and self-report is consistent, the reference weight of self-tracking data is low, acting as a supplemental source to inform decision-making. When expertise is low and self-report is unreliable, the reference weight of self-tracking data is high, leading the data to be a more influential source for discussion.

6.1 Design Implications

We have identified several design implications useful for future designs that utilize sPGD to measure patient performance in shared interfaces to support clinical decision-making and patient engagement.

6.1.1 CHR Validation and Extension. Current practices of conducting and reviewing Imaginal Exposure therapy point to ways in which the CHR concept is valid, but also to new features. As clinicians already utilize mHealth technology (e.g., PE Coach) to assist PE therapies, CHR pushes this practice forward by making it a tool that supports telehealth sessions, allowing clinicians to easily access patient behavior. CHR also facilitates comparing data streams between patients' homework and in-session engagement (i.e., adhering to PE best practices such as sitting quietly while listening to their trauma narrative recording).

Additionally, the concept of CHR facilitates clinicians' interpretation of patients' data (i.e., sPGD and SUDs) in a situated context. In this sense, future tools should first support forming a mutual understanding and calibrating patients' data. CHR may be extended to provide the charts visualizing both sPGD and self-reported pre, peak, and post-SUDs for this purpose. To allow both clinicians and patients to review the data, a tool might incorporate a Patient Homework Review (PHR) interface. PHR would be able to support patients by providing them with access to their sPGD, following the personalized decision about when and what data stream would be made available to the patients. As reported in the findings, clinicians spend time instructing patients on how to calibrate their SUDs ratings. Integrating sPGD in PHR interface might support this practice.

By adding rich contextual information to patients' data (e.g., duration of homework, location), the design of such tools could enable clinicians to identify a patient's avoidance behaviors. This may require future designs to include different features (e.g., visualizations, filtering, comparisons) that allow clinicians and patients to make sense of their data. With the help of these new features, patients may be empowered to bring aspects of their practice that they haven't mastered and clinicians can better tailor treatment to each patient's specific needs.

6.1.2 Facilitate Customization. Clinicians demonstrated distinct preferences for specific data streams, data access, and how they would approach the use of this data in session with patients. These distinct preferences are closely related to the workflow of clinicians and must be considered in designing clinician interfaces that support reviewing clinical homework and making clinical decisions. This relates to previous studies [14, 37, 49, 98] that highlight the lack of user-centered design approaches as one of the main obstacles in the adoption of clinical decision support systems in the field. Clinicians did not adopt the systems when they felt that the systems fit poorly into the clinical workflow and hamper their expertise. Therefore, future interfaces should be designed to accommodate clinicians' distinct preferences in their practices. We suggest future interfaces to allow clinicians to customize their own dashboards (e.g., choose preferred data streams) as well as patient access to data. The clinicians were very clear that patient access would be based on individual patient profiles; the implication was that uncurated data streams could have a negative implication on some patients' treatment performance. Likewise, some clinicians might be given the option between multiple physiological data streams (e.g., heart rate or GSR).

6.1.3 Develop Contextualized Training Material. Regardless of the expertise of clinicians, there are barriers to incorporating sPGD into their workflow or clinical practices. Echoing what has been discussed in previous studies on the use of PGD in clinical practice [18, 68, 78, 96], our clinicians also had concerns about having limited time to incorporate sPGD into their workflow. In addition, clinicians can face challenges while trying to integrate a novel and potentially useful datastream that they are not already familiar with [68, 96]. For example, seven out of ten participants were confused about the functionality of ambient noise level. Designing technology to support highly manualized psychotherapy like PE [25], it is recommended to work closely with clinicians to develop instructional materials with concrete use cases with detailed settings and clear guidelines on when and how to use each data stream during clinical conversations.

6.1.4 Identify What Data is Collected and Stored. Activities like imaginal exposures (i.e., trauma narratives) are deeply personal. Clinicians expressed concern over the storage of this data and the need to protect patient privacy. Past studies have also highlighted legal liability from the storage of such data [24]. Future designs should avoid storing personal health information when possible and abstract to sPGD data points only. For example, rather than storing the audio content, sPGD stores only the decibel levels. Furthermore, clear guidelines and policies should be displayed for both clinicians and patients. Patients should also have the opportunity to decline transmission of the data from any homework session recorded to protect their privacy.

6.1.5 Leverage Current Work Practices. Existing technologies affected clinician interpretation of sPGD data. First, due to the functionality of PE Coach, clinicians expected that patients would be able to utilize CHR in airplane mode. This is not the case but did impact the interpretation of data and the instruction clinicians said they would provide to patients. Second, three clinicians used eSense to measure GSR. As such, clinicians felt heart rate was less precise than the data they were accustomed to using. Future designs could consider these existing workflows and incorporate clinician preferences.

7 LIMITATIONS AND FUTURE WORK

Our CHR concept was designed with clinicians who receive state-of-the-art Prolonged Exposure therapy training. It also assumes that patients are attending therapy at the intensive outpatient program where they also are provided housing at a nearby hotel. This also assumes that patients have access to a smartphone and have free wifi. This may not be the case for patients who are doing therapy in their communities. Future research might consider how to reduce this barrier to access. It would also be interesting to determine how novice clinicians use sPGD. How might their interpretations of the data and perceived use differ from the expert clinicians enrolled in this study?

We concocted "dummy data" using the approach used in previous works [36, 37]. While we did so in co-designed sessions with an expert clinician it still may not represent real patient data. Real patients' data might reveal different patterns from what we derived (e.g., patients' data during a trip might be significantly different from normal life), which might pose challenges for the current visualizations. Future researchers may want to experiment with different design directions (e.g., allowing patients to annotate their data) to facilitate patient sense-making and patient-centered discussions. Future work should also determine how clinician usage of sPGD streams varies from patient to patient and if we can learn meaningful patterns from cohorts of patients. For example, how might sPGD collected in-session vary between patients that have combat trauma versus military sexual trauma? Is it possible to predict when habituation will occur among sub-populations of veterans? And how will patients interpret and react to the use of this data in therapy? Finally, future research should explore how CHR might be utilized in the context of patient engagement and empowerment in therapy.

8 CONCLUSION

Clinician intuition and self-report are fundamental to the delivery of psychotherapies. Our results show that technologies such as the CHR can be useful in situ. They highlight how sPGD from commodity devices can be integrated within a manualized PTSD therapy. Also, that clinicians can use this data to compare patient performance between in-clinic and homework sessions; and to gauge adherence to therapeutic best practices. They reveal the usefulness and challenges of using CHR during homework review with the patient. In sum, our findings suggest that clinicians can leverage sPGD interfaces to support their workflow, and decision-making and enhance patient-centered mental health progress.

ACKNOWLEDGMENTS

This work is supported by NSF research grant (1915504).

REFERENCES

[1] Ron Acierno, Daniel F Gros, Kenneth J Ruggiero, Melba A Hernandez-Tejada, Rebecca G Knapp, Carl W Lejuez, Wendy Muzzy, Christopher B Frueh, Leonard E Egede, and Peter W Tuerk. 2016. Behavioral activation and therapeutic exposure for posttraumatic stress disorder: A noninferiority trial of treatment delivered in person versus home-based telehealth. *Depression and Anxiety* 33, 5 (2016), 415–423.

[2] Ron Acierno, Anna E Jaffe, Amanda K Gilmore, Anna Birks, Carol Denier, Wendy Muzzy, Cristina M Lopez, Peter Tuerk, and Anouk L Grubaugh. 2021. A randomized clinical trial of in-person vs. home-based telemedicine delivery of Prolonged Exposure for PTSD in military sexual trauma survivors. *Journal of Anxiety Disorders* (2021), 102461.

[3] Patricia A Areán, Emily C Friedman, Abhishek Pratap, Ryan Allred, Jaden Duffy, Sara Gille, Shelley Reetz, Erin Keast, and Gregory Clarke. 2021. Using Real-world Data for Decision Support: Recommendations from a Primary Care Provider Survey. *The Permanente Journal* 25 (2021).

[4] American Psychological Association. 2022. Demographics of U.S. Psychology Workforce [Interactive data tool]. Retrieved 2023-07-14, from https://www.apa.org/workforce/data-tools/demographics.

[5] American Psychological Association et al. 2017. Clinical practice guideline for the treatment of posttraumatic stress disorder (PTSD) in adults.

[6] Jakob E Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. 2013. Designing mobile health technology for bipolar disorder: a field trial of the monarca system. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2627–2636.

[7] Greg Barish, Hilary Aralis, Eric Elbogen, and Patricia Lester. 2019. A mobile app for patients and those who care about them: A case study for veterans with PTSD+ Anger. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 1–10.

[8] Jonathan I Bisson, Neil P Roberts, Martin Andrew, Rosalind Cooper, and Catrin Lewis. 2013. Psychological therapies for chronic post-traumatic stress disorder (PTSD) in adults. *Cochrane database of systematic reviews* 12 (2013).

[9] Christy A Blevins, Frank W Weathers, Margaret T Davis, Tracy K Witte, and Jessica L Domino. 2015. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *Journal of traumatic stress* 28, 6 (2015), 489–498.

[10] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[11] Taylor A Braund, May The Zin, Tjeerd W Boonstra, Quincy J J Wong, Mark E Larsen, Helen Christensen, Gabriel Tillman, and Bridianne O'Dea. 2022. Smartphone Sensor Data for Identifying and Monitoring Symptoms of Mood Disorders: A Longitudinal Observational Study. *JMIR Mental Health* 9, 5 (May 2022), e35549. https://doi.org/10.2196/35549

[12] Deborah J Brief, Amy Rubin, Terence M Keane, Justin L Enggasser, Monica Roy, Eric Helmuth, John Hermos, Mark Lachowicz, Denis Rybin, and David Rosenbloom. 2013. Web intervention for OEF/OIF veterans with problem drinking and PTSD symptoms: A randomized clinical trial. *Journal of consulting and clinical psychology* 81, 5 (2013), 890.

[13] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[14] Carrie J. Cai, Samantha Winter, David Steiner, Lauren Wilcox, and Michael Terry. 2019. "Hello AI": Uncovering the Onboarding Needs of Medical Practitioners for Human-AI Collaborative Decision-Making. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–24. https://doi.org/10.1145/3359206

[15] Emil Chiauzzi and Paul Wicks. 2021. Beyond the Therapist's Office: Merging Measurement-Based Care and Digital Medicine in the Real World. *Digital Biomarkers* 5, 2 (2021), 176–182.

[16] Eun Kyoung Choe, Saeed Abdullah, Mashfiqui Rabbi, Edison Thomaz, Daniel A. Epstein, Felicia Cordeiro, Matthew Kay, Gregory D. Abowd, Tanzeem Choudhury, James Fogarty, Bongshin Lee, Mark Matthews, and Julie A. Kientz. 2017. Semi-Automated Tracking: A Balanced Approach for Self-Monitoring Applications. *IEEE Pervasive Computing* 16, 1 (Jan. 2017), 74–84. https://doi.org/10.1109/MPRV.2017.18

[17] Chia-Fang Chung, Jonathan Cook, Elizabeth Bales, Jasmine Zia, and Sean A Munson. 2015. More than telemonitoring: health provider use and nonuse of life-log data in irritable bowel syndrome and weight management. *Journal of medical Internet research* 17, 8 (2015), e203.

[18] 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*. 770–786.

[19] Beth E Cohen, Kris Gima, Daniel Bertenthal, Sue Kim, Charles R Marmar, and Karen H Seal. 2010. Mental health diagnoses and utilization of VA non-mental health medical services among returning Iraq and Afghanistan veterans. *Journal of general internal medicine* 25, 1 (2010), 18–24.

[20] Karen Cusack, Daniel E Jonas, Catherine A Forneris, Candi Wines, Jeffrey Sonis, Jennifer Cook Middleton, Cynthia Feltner, Kimberly A Brownley, Kristine Rae Olmsted, Amy Greenblatt, et al. 2016. Psychological treatments for adults with posttraumatic stress disorder: A systematic review and meta-analysis. *Clinical psychology review* 43 (2016), 128–141.

[21] Vedant Das Swain, Victor Chen, Shrija Mishra, Stephen M. Mattingly, Gregory D. Abowd, and Munmun De Choudhury. 2022. Semantic Gap in Predicting Mental Wellbeing through Passive Sensing. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–16. https://doi.org/10.1145/3491102.3502037

[22] Wenrui Deng, Die Hu, Sheng Xu, Xiaoyu Liu, Jingwen Zhao, Qian Chen, Jiayuan Liu, Zheng Zhang, Wenxiu Jiang, Lijun Ma, et al. 2019. The efficacy of virtual reality exposure therapy for PTSD symptoms: A systematic review and meta-analysis. *Journal of affective disorders* 257 (2019), 698–709.

[23] Fifth Edition. 2013. Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc* 21 (2013).

[24] 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*. 1–15.

[25] E.B. Foa, E. A. Hembree, B. O. Rothbaum, and S. A. M. Rauch. 2019. Prolonged Exposure Therapy for PTSD. In *Emotional Processing of Traumatic Experiences - Therapist Guide (2nd ed.)*. Oxford University Press.

[26] Sharon L Foster and John D Cone. 1995. Validity issues in clinical assessment. *Psychological assessment* 7, 3 (1995), 248.

[27] Mads Frost, Afsaneh Doryab, Maria Faurholt-Jepsen, Lars Vedel Kessing, and Jakob E Bardram. 2013. Supporting disease insight through data analysis: refinements of the monarca self-assessment system. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. 133–142.

[28] Madelyn R Frumkin, Marilyn L Piccirillo, Emorie D Beck, Jason T Grossman, and Thomas L Rodebaugh. 2021. Feasibility and utility of idiographic models in the clinic: a pilot study. *Psychotherapy Research* 31, 4 (2021), 520–534.

[29] Hector A Garcia, Lance P Kelley, Timothy O Rentz, and Shuko Lee. 2011. Pretreatment predictors of dropout from cognitive behavioral therapy for PTSD in Iraq and Afghanistan war veterans. *Psychological Services* 8, 1 (2011), 1.

[30] Raquel Gonçalves, Ana Lúcia Pedrozo, Evandro Silva Freire Coutinho, Ivan Figueira, and Paula Ventura. 2012. Efficacy of virtual reality exposure therapy in the treatment of PTSD: a systematic review. *PloS one* 7, 12 (2012), e48469.

[31] Molly Gromatsky, Sarah R Sullivan, Angela Page Spears, Emily Mitchell, Samantha Walsh, Nathan A Kimbrel, and Marianne Goodman. 2020. Ecological momentary assessment (EMA) of mental health outcomes in veterans and servicemembers: a scoping review. *Psychiatry research* (2020), 113359.

[32] Haijing Wu Hallenbeck, Beth K Jaworski, Joseph Wielgosz, Eric Kuhn, Kelly M Ramsey, Katherine Taylor, Katherine Juhasz, Pearl McGee-Vincent, Margaret-Anne Mackintosh, and Jason E Owen. 2022. PTSD Coach Version 3.1: A Closer Look at the Reach, Use, and Potential Impact of This Updated Mobile Health App in the General Public. *JMIR mental health* 9, 3 (2022), e34744.

[33] John Hunsley and Eric J Mash. 2007. Evidence-based assessment. *Annu. Rev. Clin. Psychol.* 3 (2007), 29–51.

[34] Washington University in St. Louis. 2008. Growing Problem For Veterans: Domestic Violence. *Science Daily* (2008).

[35] Riitta Jääskeläinen. 2010. Think-aloud protocol. *Handbook of translation studies* 1 (2010), 371–374.

[36] Maia Jacobs, Jeffrey He, Melanie F. Pradier, Barbara Lam, Andrew C. Ahn, Thomas H. McCoy, Roy H. Perlis, Finale Doshi-Velez, and Krzysztof Z. Gajos. 2021. Designing AI for Trust and Collaboration in Time-Constrained Medical Decisions: A Sociotechnical Lens. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*.

1–14. https://doi.org/10.1145/3411764.3445385 arXiv:2102.00593 [cs].

[37] Eunkyung Jo, Myeonghan Ryu, Georgia Kenderova, Samuel So, Bryan Shapiro, Alexandra Papoutsaki, and Daniel A. Epstein. 2022. Designing Flexible Longitudinal Regimens: Supporting Clinician Planning for Discontinuation of Psychiatric Drugs. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–17. https://doi.org/10.1145/3491102.3502206

[38] Adrienne S. Juarascio, Rebecca J. Crochiere, Tinashe M. Tapera, Madeline Palermo, and Fengqing Zhang. 2020. Momentary changes in heart rate variability can detect risk for emotional eating episodes. *Appetite* 152 (Sept. 2020), 104698. https://doi.org/10.1016/j.appet.2020.104698

[39] Alan E Kazdin. 2017. Addressing the treatment gap: A key challenge for extending evidence-based psychosocial interventions. *Behaviour research and therapy* 88 (2017), 7–18.

[40] Christina Kelley, Bongshin Lee, and Lauren Wilcox. 2017. Self-tracking for mental wellness: understanding expert perspectives and student experiences. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 629–641.

[41] Ronald C Kessler, Wai Tat Chiu, Olga Demler, and Ellen E Walters. 2005. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of general psychiatry* 62, 6 (2005), 617–627.

[42] Kurt Kroenke, Robert L Spitzer, and Janet BW Williams. 2001. The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine* 16, 9 (2001), 606–613.

[43] Annika Kuester, Helen Niemeyer, and Christine Knaevelsrud. 2016. Internet-based interventions for posttraumatic stress: a meta-analysis of randomized controlled trials. *Clinical Psychology Review* 43 (2016), 1–16.

[44] Eric Kuhn, Jill J Crowley, Julia E Hoffman, Afsoon Eftekhari, Kelly M Ramsey, Jason E Owen, Greg M Reger, and Josef I Ruzek. 2015. Clinician characteristics and perceptions related to use of the PE (prolonged exposure) coach mobile app. *Professional Psychology: Research and Practice* 46, 6 (2015), 437.

[45] Eric Kuhn, Nitya Kanuri, Julia E Hoffman, Donn W Garvert, Josef I Ruzek, and C Barr Taylor. 2017. A randomized controlled trial of a smartphone app for posttraumatic stress disorder symptoms. *Journal of consulting and clinical psychology* 85, 3 (2017), 267.

[46] Charlotte P. Lee. 2007. Boundary Negotiating Artifacts: Unbinding the Routine of Boundary Objects and Embracing Chaos in Collaborative Work. *Computer Supported Cooperative Work (CSCW)* 16, 3 (June 2007), 307–339. https://doi.org/10.1007/s10606-007-9044-5

[47] D. Lee, S. Kleiman, and F. Weathers. 2019. Assessment: Trauma- and stressor-related disorders. In *Handbook of Anxiety and Related Disorders*, B. O. Olatunji (Ed.). Cambridge Press. https://doi.org/10.1017/9781108140416.014

[48] Daniel J Lee, Carla W Schnitzlein, Jonathan P Wolf, Meena Vythilingam, Ann M Rasmusson, and Charles W Hoge. 2016. Psychotherapy versus pharmacotherapy for posttraumatic stress disorder: Systemic review and meta-analyses to determine first-line treatments. *Depression and anxiety* 33, 9 (2016), 792–806.

[49] Min Hun Lee, Daniel P. Siewiorek, Asim Smailagic, Alexandre Bernardino, and Sergi Bermúdez i Badia. 2020. Co-Design and Evaluation of an Intelligent Decision Support System for Stroke Rehabilitation Assessment. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (Oct. 2020), 1–27. https://doi.org/10.1145/3415227

[50] Cara C Lewis, Meredith Boyd, Ajeng Puspitasari, Elena Navarro, Jacqueline Howard, Hannah Kassab, Mira Hoffman, Kelli Scott, Aaron Lyon, Susan Douglas, et al. 2019. Implementing measurement-based care in behavioral health: a review. *JAMA psychiatry* 76, 3 (2019), 324–335.

[51] Kelly P Maieritsch, Tracey L Smith, Jonathan D Hessinger, Eileen P Ahearn, Jens C Eickhoff, and Qianqian Zhao. 2016. Randomized controlled equivalence trial comparing videoconference and in person delivery of cognitive processing therapy for PTSD. *Journal of Telemedicine and Telecare* 22, 4 (2016), 238–243.

[52] Mark Matthews, Saeed Abdullah, Elizabeth Murnane, Stephen Voida, Tanzeem Choudhury, Geri Gay, and Ellen Frank. 2016. Development and evaluation of a smartphone-based measure of social rhythms for bipolar disorder. *Assessment* 23, 4 (2016), 472–483.

[53] Adam Miner, Eric Kuhn, Julia E Hoffman, Jason E Owen, Josef I Ruzek, and C Barr Taylor. 2016. Feasibility, acceptability, and potential efficacy of the PTSD Coach app: A pilot randomized controlled trial with community trauma survivors. *Psychological Trauma: Theory, Research, Practice, and Policy* 8, 3 (2016), 384.

[54] Sonali R Mishra, Andrew D Miller, Shefali Haldar, Maher Khelifi, Jordan Eschler, Rashmi G Elera, Ari H Pollack, and Wanda Pratt. 2018. Supporting collaborative health tracking in the hospital: patients' perspectives. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–14.

[55] Ramin Mojtabai, Mark Olfson, Nancy A Sampson, Robert Jin, Benjamin Druss, Philip S Wang, Kenneth B Wells, Harold A Pincus, and Ronald C Kessler. 2011. Barriers to mental health treatment: results from the National Comorbidity Survey Replication. *Psychological medicine* 41, 8 (2011), 1751–1761.

[56] Leslie A Morland, Margaret-Anne Mackintosh, Craig S Rosen, Emy Willis, Patricia Resick, Kathleen Chard, and B Christopher Frueh. 2015. Telemedicine versus in-person delivery of cognitive processing therapy for women with

posttraumatic stress disorder: A randomized noninferiority trial. *Depression and Anxiety* 32, 11 (2015), 811–820.
[57] Leslie A Morland, Michelle Raab, Margaret-Anne Mackintosh, Craig S Rosen, Clara E Dismuke, Carolyn J Greene, and B Christopher Frueh. 2013. Telemedicine: A cost-reducing means of delivering psychotherapy to rural combat veterans with PTSD. *Telemedicine and e-Health* 19, 10 (2013), 754–759.
[58] Debbie S Moskowitz and Simon N Young. 2006. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. *Journal of Psychiatry and Neuroscience* 31, 1 (2006), 13.
[59] Elizabeth L Murnane, Dan Cosley, Pamara Chang, Shion Guha, Ellen Frank, Geri Gay, and Mark Matthews. 2016. Self-monitoring practices, attitudes, and needs of individuals with bipolar disorder: implications for the design of technologies to manage mental health. *Journal of the American Medical Informatics Association* 23, 3 (2016), 477–484.
[60] Elizabeth L Murnane, Tara G Walker, Beck Tench, Stephen Voida, and Jaime Snyder. 2018. Personal informatics in interpersonal contexts: towards the design of technology that supports the social ecologies of long-term mental health management. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–27.
[61] M. W. Miller T. M. Keane N. A. Livingston, D. J. Brief. 2021. Assessment of PTSD and its comorbidities in adults. In *Handbook of PTSD: Science and practice (3rd Edition)*, P. P. Schnurr M. J. Friedman and T. M. Keane (Eds.). Guilford Press.
[62] Gina Neff and Dawn Nafus. 2016. *Self-tracking.* MIT Press.
[63] Ada Ng, Rachel Kornfield, Stephen M Schueller, Alyson K Zalta, Michael Brennan, and Madhu Reddy. 2019. Provider perspectives on integrating sensor-captured patient-generated data in mental health care. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–25.
[64] Ada Ng, Madhu Reddy, Alyson K Zalta, Stephen M Schueller, et al. 2018. Veterans' perspectives on fitbit use in treatment for post-traumatic stress disorder: an interview study. *JMIR mental health* 5, 2 (2018), e10415.
[65] U.S. Department of Veterans Affairs. 2018. How Common is PTSD in Veterans? - PTSD: National Center for PTSD. https://www.ptsd.va.gov/understand/common/common_veterans.asp
[66] U.S. Department of Veterans Affairs. 2020. PTSD and Substance Abuse in Veterans - PTSD: National Center for PTSD. https://www.ptsd.va.gov/professional/continuing_ed/ptsd_sud_comorbidity.asp
[67] U.S. Department of Veterans Affairs. April 24, 2014. Study explores reasons why Veterans seek—or don't seek—PTSD care. https://www.research.va.gov/currents/spring2014/spring2014-25.cfm
[68] Chi Young Oh, Yuhan Luo, Beth St. Jean, and Eun Kyoung Choe. 2022. Patients Waiting for Cues: Information Asymmetries and Challenges in Sharing Patient-Generated Data in the Clinic. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (March 2022), 1–23. https://doi.org/10.1145/3512954
[69] Mika Pantzar and Minna Ruckenstein. 2017. Living the metrics: Self-tracking and situated objectivity. *Digital health* 3 (2017), 2055207617712590.
[70] Sarah Pink and Vaike Fors. 2017. Being in a mediated world: self-tracking and the mind–body–environment. *cultural geographies* 24, 3 (2017), 375–388.
[71] Kyle Possemato, Eric Kuhn, Emily Johnson, Julia E Hoffman, Jason E Owen, Nitya Kanuri, Leigha De Stefano, and Emily Brooks. 2016. Using PTSD Coach in primary care with and without clinician support: a pilot randomized controlled trial. *General Hospital Psychiatry* 38 (2016), 94–98.
[72] Carolina Rodriguez-Paras, Kathryn Tippey, Elaine Brown, Farzan Sasangohar, Suzannah Creech, Hye-Chung Kum, Mark Lawley, and Justin K Benzer. 2017. Posttraumatic stress disorder and mobile health: app investigation and scoping literature review. *JMIR mHealth and uHealth* 5, 10 (2017), e156.
[73] CS Rosen, MM Matthieu, S Wiltsey Stirman, JM Cook, S Landes, NC Bernardy, KM Chard, J Crowley, A Eftekhari, EP Finley, et al. 2016. A review of studies on the system-wide implementation of evidence-based psychotherapies for posttraumatic stress disorder in the Veterans Health Administration. *Administration and Policy in Mental Health and Mental Health Services Research* 43, 6 (2016), 957–977.
[74] Barbara Olasov Rothbaum and Ann C Schwartz. 2002. Exposure therapy for posttraumatic stress disorder. *American journal of psychotherapy* 56, 1 (2002), 59–75.
[75] Jason J. Saleem, Nancy R. Wilck, John J. Murphy, and Jennifer Herout. 2022. Veteran and Staff Experience from a Pilot Program of Health Care System–Distributed Wearable Devices and Data Sharing. *Applied Clinical Informatics* 13, 03 (May 2022), 532–540. https://doi.org/10.1055/s-0042-1748857
[76] Elaine Schertz, Hue Watson, Ashok Krishna, Andrew Sherrill, Hayley Evans, and Rosa I Arriaga. 2019. Bridging the Gap: Creating a Clinician-Facing Dashboard for PTSD. In *IFIP Conference on Human-Computer Interaction*. Springer, 224–233.
[77] Paula P Schnurr and Bonnie L Green. 2004. *Trauma and health: Physical health consequences of exposure to extreme stress.* American Psychological Association.
[78] Jessica Schroeder, Jane Hoffswell, Chia-Fang Chung, James Fogarty, Sean Munson, and Jasmine Zia. 2017. Supporting patient-provider collaboration to identify individual triggers using food and symptom journals. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1726–1739.

[79] Aviv Shachak and Shmuel Reis. 2009. The impact of electronic medical records on patient–doctor communication during consultation: a narrative literature review. *Journal of evaluation in clinical practice* 15, 4 (2009), 641–649.

[80] Andrew M Sherrill, Jessica R Goodnight, Mark S Burton, and Barbara O Rothbaum. 2020. Use of Virtual Reality Exposure Therapy for Anxiety-and Trauma-Related Disorders. In *Technology and Mental Health*. Routledge, 75–91.

[81] Andrew M Sherrill, Alex O Rothbaum, Lauren B McSweeney, and Barbara O Rothbaum. 2019. Virtual reality exposure therapy for posttraumatic stress disorder. *Psychiatric Annals* 49, 8 (2019), 343–347.

[82] 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. *Administration and Policy in Mental Health and Mental Health Services Research* 40, 4 (2013), 311–318.

[83] Marit Sijbrandij, Ivo Kunovski, and Pim Cuijpers. 2016. Effectiveness of internet-delivered cognitive behavioral therapy for posttraumatic stress disorder: A systematic review and meta-analysis. *Depression and anxiety* 33, 9 (2016), 783–791.

[84] Tracy Stecker, Brian Shiner, Bradley V Watts, Meissa Jones, and Kenneth R Conner. 2013. Treatment-seeking barriers for veterans of the Iraq and Afghanistan conflicts who screen positive for PTSD. *Psychiatric Services* 64, 3 (2013), 280–283.

[85] Terri L Tanielian, Terri Tanielian, and Lisa Jaycox. 2008. *Invisible wounds of war: Psychological and cognitive injuries, their consequences, and services to assist recovery*. Vol. 1. Rand Corporation.

[86] David Turgoose, Rachel Ashwick, and Dominic Murphy. 2018. Systematic review of lessons learned from delivering tele-therapy to veterans with post-traumatic stress disorder. *Journal of telemedicine and telecare* 24, 9 (2018), 575–585.

[87] Yoni Van Den Eede. 2015. Tracing the tracker: A postphenomenological inquiry into self-tracking technologies. In *Postphenomenological investigations: Essays on human–technology relations*. Lexington Books, 143–158.

[88] Julio Vega, Beth T Bell, Caitlin Taylor, Jue Xie, Heidi Ng, Mahsa Honary, and Roisin McNaney. 2022. Detecting Mental Health Behaviors Using Mobile Interactions: Exploratory Study Focusing on Binge Eating. *JMIR Mental Health* 9, 4 (April 2022), e32146. https://doi.org/10.2196/32146

[89] Stephen Voida, Mark Matthews, Saeed Abdullah, Mengxi Xi, Matthew Green, Won Jun Jang, Donald Hu, John Weinrich, Prashama Patil, Mashfiqui Rabbi, et al. 2013. Moodrhythm: tracking and supporting daily rhythms. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. 67–70.

[90] Sophie Walsh, Eoin Golden, and Stefan Priebe. 2016. Systematic review of patients' participation in and experiences of technology-based monitoring of mental health symptoms in the community. *BMJ open* 6, 6 (2016), e008362.

[91] Rui Wang, Min SH Aung, Saeed Abdullah, Rachel Brian, Andrew T Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Michael Merrill, Emily A Scherer, et al. 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*. 886–897.

[92] 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. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–26.

[93] Weichen Wang, Shayan Mirjafari, Gabriella Harari, Dror Ben-Zeev, Rachel Brian, Tanzeem Choudhury, Marta Hauser, John Kane, Kizito Masaba, Subigya Nepal, et al. 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*. 1–15.

[94] Zhiyun Wang, Jianping Wang, and Andreas Maercker. 2013. Chinese My Trauma Recovery, a Web-based intervention for traumatized persons in two parallel samples: randomized controlled trial. *Journal of medical Internet research* 15, 9 (2013), e213.

[95] Bradley V Watts, Paula P Schnurr, Lorna Mayo, Yinong Young-Xu, William B Weeks, and Matthew J Friedman. 2013. Meta-analysis of the efficacy of treatments for posttraumatic stress disorder. *The Journal of clinical psychiatry* 74, 6 (2013), 0–0.

[96] Peter West, Max Van Kleek, Richard Giordano, Mark J. Weal, and Nigel Shadbolt. 2018. Common Barriers to the Use of Patient-Generated Data Across Clinical Settings. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC Canada, 1–13. https://doi.org/10.1145/3173574.3174058

[97] Julia M Whealin, Emily C Jenchura, Ava C Wong, and Donna M Zulman. 2016. How veterans with post-traumatic stress disorder and comorbid health conditions utilize ehealth to manage their health care needs: a mixed-methods analysis. *Journal of medical Internet research* 18, 10 (2016), e280.

[98] Qian Yang, John Zimmerman, Aaron Steinfeld, Lisa Carey, and James F. Antaki. 2016. Investigating the Heart Pump Implant Decision Process: Opportunities for Decision Support Tools to Help. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 4477–4488. https://doi.org/10.1145/2858036.2858373

[99] Erica K Yuen, Daniel F Gros, Matthew Price, Stephanie Zeigler, Peter W Tuerk, Edna B Foa, and Ron Acierno. 2015. Randomized controlled trial of home-based telehealth versus in-person prolonged exposure for combat-related PTSD

in veterans: Preliminary results. *Journal of Clinical Psychology* 71, 6 (2015), 500–512.

A APPENDICES

A.1 Interview / think aloud protocol

A.1.1 Introduction.

Hi [PARTICIPANT NAME], thank you very much for joining us and for participating in our study! My name is [NAME]. I am joined by my colleague, [NAME] who will be taking notes during this session.

Today, I will be taking you through an exercise with a dashboard that displays a new technology to assist clinicians as they guide patients through exposure therapy. In our session today we will ask a few questions, show you a short video about a new technology, and then have you use a prototype of the technology. We will finish the session by having you answer some questions about your experience with the prototype.

Before we jump into things, I want to remind you of your rights as a participant. Your participation is anonymous. You do not need to answer any questions you do not want to or participate in any tasks you do not want to. If at any time you decide you no longer want to participate, please let me know and we will discontinue the session immediately.

Do you have any questions about the session before we get started?

Do you mind if we record this session?

A.1.2 Baseline questions.

- First, I'd like to ask you about your experience with collecting self-report from patients about their homework activity during exposure therapy.
- How do you evaluate patient self-report for imaginal exposure homework activity?
- Can you give me an example of a time when you evaluated self-report? NOTE: Tailor this question to how the previous question (e.g. if they say patients sometimes aren't truthful, ask for an example of that). You can dig in deeper depending on what they share.

A.1.3 Showing the video about the prototype (script below).

- Many mental health interventions, such as Prolonged Exposure Therapy, have a homework component. Imagine if you could be a fly on the wall while your patient completed their homework. You would be able to understand how the context impacted their performance and provide necessary feedback for improvement.
- The Clinician Homework Review, or CHR, is a novel concept that aims to do just that. It is a measure that compares and contrasts patient engagement during clinician led imaginal exposure exercises and for homework sessions throughout exposure therapy.
- During the clinical session a patient uses a mobile app and FitBit, to record their imaginal exposure exercises. When the app is turned on, various data streams including heart rate, physical activity, location, phone usage, and noise level in the environment are collected to establish a baseline level of engagement for optimal learning.
- Later, when the patient is ready to do their homework, they turn on the app and it collects the same data streams that were collected during the in-person clinical session.
- It then compares and contrasts the patient engagement measures detected during the homework session to the engagement measures recorded earlier in the clinical session.
- These measures will be displayed as a Clinician Homework Review on both a clinician dashboard and patient-facing mobile application. Both clinicians and patients will be able to view patient engagement performance from the clinical and homework sessions.

- Measures include: a comprehensive CHR score to give a high-level similarity rating that compares engagement during the clinical session to engagement during homework. This is an average of the scores of all data streams collected; scores and visualizations for each data stream collected; averages from all sessions within each data stream.
- CHR collects and displays comparative data regarding patient engagement in imaginal exposure exercises throughout exposure therapy. Clinicians can use this data to inform their delivery of exposure therapy and guide the patient to optimal performance. The patient gets feedback on their performance that allows them to optimize the next homework exercise. They then get additional information that confirms what they did well and where they can improve.
- In short, CHR helps clinicians to individualize feedback to each patient and helps patients know how well they did during their homework.

A.1.4 Main Interview Questions. Now we are going to ask you a few questions about your experience with the prototype as well as your background.

Prototype-related questions

- In your own words, what does the dashboard do?
- How, if at all, do you see yourself using the dashboard as a part of your exposure therapy practice?
- What data, if any, do you believe patients should have access to? When?
- Is there anything else you would like to share?

Experience-related questions

- How long have you been practicing clinical psychology?
- How long have you been practicing exposure therapy?
- On a scale of 1 to 5, with one being the lowest and 5 being the highest, how would you rate your expertise using technology?
- What technology (software, hardware) do you typically use as part of your clinical practice?

B SURVEY QUESTIONS

Responses were given on a Likert scale with 5 items, from 1 (strongly disagree) to 5 (strongly agree).

- I think that I would like to use the dashboard frequently.
- I found the dashboard unnecessarily complex.
- I thought the dashboard was easy to use.
- I think that I would need the support of a technical person to be able to use the dashboard.
- I found that the various functions in the dashboard were well integrated .
- I thought that there was too much inconsistency in the dashboard.
- I would imagine most people would learn to use the dashboard very quickly.
- I found the dashboard very cumbersome to use.
- I felt very confident using the dashboard.
- I needed to learn a lot of things before I could get going with the dashboard.

Received January 2023; revised July 2023; accepted November 2023