New Ideas in Psychology 76 (2025) 101127

Contents lists available at ScienceDirect

New Ideas in Psychology

journal homepage: www.elsevier.com/locate/newideapsych

Virtual reality analytics map (VRAM): A conceptual framework for detecting mental disorders using virtual reality data

Vibhav Chitale[a,*], Julie D. Henry[b], Hai-Ning Liang[c], Ben Matthews[a], Nilufar Baghaei[a]

[a] School of Electrical Engineering and Computer Science, The University of Queensland, Australia
[b] School of Psychology, The University of Queensland, Brisbane, Australia
[c] Computational Media and Arts Thrust, Information Hub, The Hong Kong University of Science and Technology (Guangzhou), China

ARTICLE INFO

Keywords:
Virtual reality
Analytics
Mental disorders
Conceptual framework
Diagnosis
Digital biomarkers

ABSTRACT

Virtual reality (VR) is an emerging tool in mental health care yet its potential in diagnostic assessments remains underexplored. Recognizing the growing need of technological advancements that support traditional methods for mental health assessment, this paper introduces the Virtual Reality Analytics Map (VRAM), a novel conceptual framework designed to leverage VR analytics for the detection of symptoms of mental disorders. The VRAM framework integrates psychological constructs with VR technology, systematically mapping and quantifying behavioral domains through specific VR tasks. This approach potentially allows for the precise capture and identification of nuanced behavioral, cognitive, and affective digital biomarkers associated with symptoms of mental disorders. The benefits of the VRAM framework are demonstrated with its example application across various mental disorders ensuring the utility and versatility of the framework. By bridging the gap between psychology and technology, the VRAM framework aims to contribute to the early detection and assessment of mental disorders.

1. Introduction

According to the World Health Organization (WHO), one in every eight or about 970 million people globally, were living with a mental illness in 2019 (World Health Organization, 2022). With mental disorders continuing to rapidly increase worldwide, the WHO predicts that by 2030 mental disorders will be the leading cause of disease burden globally (World Health Organization, 2023). Clinical criteria defined in standardized manuals such as the Diagnostic and Statistics Manual (DSM - 5) (American Psychiatric Association, 2022) or the International Classification of Diseases (ICD - 11) (World Health Organization, 2019) are used by clinicians, practitioners, and researchers for diagnosing mental disorders. Such manuals classify or distinguish mental disorders based on the type, duration and intensity of observable symptoms which are defined as the "diagnostic criteria". Clinician-administered interview-based methods are widely regarded as the gold standard, yet resource shortages or societal stigma can limit its accessibility and availability (Kilbourne et al., 2018; Kumar & Phookun, 2016; Wainberg et al., 2017). Alternatively, self-report measures have the advantage of being quick and easy to administer yet their diagnostic sensitivity can be influenced by factors such as social desirability bias, lack of self-awareness or the respondents' willingness to disclose personal information (Althubaiti, 2016; Krumpal, 2013). Therefore, there has been a growing interest towards digital technologies that can potentially support the traditional methods by providing more objective and ecologically valid assessments of mental health symptoms that can assist in the early diagnosis, monitoring, and treatment of mental disorders (Balcombe & De Leo, 2021; D Alfonso et al., 2020).

1.1. *Digital technologies and virtual reality in mental health assessments*

Digital technologies have emerged as promising tools in mental health assessments because they offer data collection from numerous sources to develop computational models that help predict symptoms of mental illness. Several studies have now successfully applied machine learning, deep learning or natural language processing algorithms to data collected from a range of sources such as text-based data from social media or text messages (Corcoran et al., 2018; D Alfonso et al., 2020; Gkotsis et al., 2017; Kim, Lee, Park, & Han, 2020; Wongkoblap, Vadillo, & Curcin, 2021), audio, speech or facial expression data (Gavrilescu &

* Corresponding author. School of Electrical Engineering and Computer Science, The University of Queensland, Brisbane, Queensland, Australia.
E-mail addresses: v.chitale@uq.edu.au (V. Chitale), julie.henry@uq.edu.au (J.D. Henry), HaiNing.Liang@xjtlu.edu.cn (H.-N. Liang), matthews@uq.edu.au (B. Matthews), n.baghaei@uq.edu.au (N. Baghaei).

https://doi.org/10.1016/j.newideapsych.2024.101127
Received 7 January 2024; Received in revised form 10 September 2024; Accepted 26 September 2024
Available online 2 October 2024

Vizireanu, 2019; Hashim, Wilkes, Salomon, Meggs, & France, 2017; He & Cao, 2018; Ozkanca et al., 2019; Pampouchidou et al., 2020), or data from wearable devices and sensors (Meyer et al., 2018; Miranda, Calderón, & Favela, 2014, pp. 34–41; O'Brien et al., 2017; Wang et al., 2018).

Although many digital technologies are now available, Virtual Reality (VR) has been the focus of increasing interest as an innovative approach for use in mental health care (Baghaei et al., 2021; Freeman et al., 2017). Virtual Reality is the use of computer modelling and simulation that enables a person to interact within a virtual environment giving an immersive feel of the virtual world (Lowood, 2012). VR has several benefits over other digital technologies because it can offer an immersive and interactive experience that can improve user participation and engagement which are crucial to digital solutions (Voinescu et al., 2021). Indeed, VR not only allows for the creation of real-world simulations, but ones that can be tweaked and tailored according to the needs of the participants in a controlled manner simply not be possible with conventional methods (Parsons, 2015; A. S. Rizzo & Koenig, 2017). Additionally, users are often more willing to be exposed to triggering scenarios in virtual environments since they are just a simulation yet respond both physiologically and psychologically similar to how they would in the real-world (Freeman et al., 2017; Riva et al., 2007). VR also has the important strength of being more ecologically valid in helping to translate research findings to a real-world setting (Parsons, 2015) and can offer multimodal objective measurements by tracking physiological responses as well as user interaction within the virtual worlds (Bell, Nicholas, Alvarez-Jimenez, Thompson, & Valmaggia, 2020; Chitale, Baghaei, et al., 2022; Chitale, Playne, Liang, & Baghaei, 2022; Freeman et al., 2017).

Despite the many advantages of using VR in mental health care, its implementation beyond research settings in detecting or assessing symptoms of mental disorders has to date been relatively limited (Bell et al., 2020; Chitale, Baghaei, et al., 2022; Meyerbröker & Morina, 2021). This may be attributable to several factors, including the current lack of standardized VR frameworks, higher operational costs, availability of simpler solutions that are more accessible than VR systems. Furthermore, the lack of psychometric evaluations (e.g., reliability, validity) in VR studies combined with the need to compare VR measures with traditional diagnostic methods have slowed its wider translation to clinical settings (Freeman et al., 2017; A. S. Rizzo & Koenig, 2017). Additionally, the empirical evidence supporting VR is generally based on relatively small sample sizes, whereas studies concerning conventional assessment methods typically rely on much larger sample sizes for validation (Freeman et al., 2017).

While VR offers several advantages, it should be considered adjunct to, and not a replacement for, traditional assessment approaches such as self-reports measures that are well-established, validated and clinical interviews, that are widely considered to be the gold standard in mental health diagnosis. At present, VR-based assessments have not been subject to the same rigorous validation that has been applied to these traditional methods. Traditional diagnostic methods (and in particular, clinical interview), also offer deeper insights into a patient's overall psychological states, personal history and broader social and environmental factors. These factors often play a critical role in the development and progression of mental health conditions and seem unlikely to be fully captured through VR systems alone. Nonetheless, VR holds substantial potential to support the existing traditional assessment methods by offering immersive, ecologically valid and safe environments rather than act as their replacement.

1.2. *Virtual reality analytics*

When users interact within a virtual environment in VR, large volumes of interactional data are left behind reflective of an individual's thoughts, feelings, physiological responses, and behaviors pertinent to their mental states. Several studies have reported encouraging findings (Bell et al., 2020; Chitale, Playne, et al., 2022; Freeman et al., 2017) on using VR as an assessment tool despite their exploratory nature.

VR analytics (i.e., tracking, visualizing, modelling, analyzing and interpreting VR data) can help identify the data metrics or correlates (digital biomarkers) indicative of symptoms of mental disorders (Bell et al., 2020; Chitale, Playne, et al., 2022; Freeman et al., 2017). These digital biomarkers are crucial both from a clinical and a design standpoint, playing a pivotal role in the development of tailored VR applications aimed at detecting the symptoms of mental disorders. However, suitable design guidelines and standardized frameworks are first needed to fully realize the potential of VR analytics. The multidisciplinary nature of this field also complicates these issues, as each discipline relies on its own specific set of design principles and frameworks. Therefore, in this study, we propose a novel conceptual framework called the Virtual Reality Analytics Map (VRAM), which has been designed to guide researchers in leveraging VR analytics for mental health assessment. The VRAM framework presents a step-by-step roadmap to support the design and development of VR applications integrated with VR analytics for detecting symptoms of mental disorders. Specifically, the VRAM framework provides a standardized approach on how to effectively leverage data from VR headsets, sensors and participants' interactions within the virtual environments to identify key metrics that should theoretically be linked to symptoms of mental disorders.

A key issue in current research is lack of transparency in design decisions. Researchers often provide little to no explanation of the design decisions or the scientific basis for introducing specific VR tasks or why certain data metrics were measured. For VR applications to be used as reliable and valid assessment tools, researchers must first consider how the design of different elements and components may influence user behaviors. Each design element has the capacity to influence the generated data affecting the validity and reliability of the VR applications. For instance, introducing time limits to complete tasks in the virtual world can potentially affect decision making and physiological measures due to the added stressor thereby significantly influencing the collected data metrics. By identifying which VR tasks elicit meaningful data related to specific symptoms, researchers can design VR scenarios that are both engaging and scientifically grounded in their ability to predict these symptoms. As we argue below, there is a clear gap in the literature with no existing mapping models that provide guidance on how to determine which VR tasks are best suited for appropriately capturing the data metrics associated with symptoms of mental disorders.

2. Objectives

The primary aim of this study is to introduce a novel conceptual framework called Virtual Reality Analytics Map (VRAM) that can serve as a guide to understanding how VR data and analytics can be harnessed to detect symptoms of mental disorders. The VRAM conceptual framework provides a step-by-step structured approach to leverage VR analytics for mental health assessments, bridging the gap between psychology and technology.

3. Related conceptual frameworks

Due to the lack of existing frameworks within the scope of VR specifically, we turned our attention to the previously established frameworks in the fields of serious games, education, digital phenotyping, and health. This section provides an overview of frameworks, or models that supported the conception of the proposed VRAM conceptual framework.

A theory-driven evidence-based framework for Serious Games for Health (SGH) was proposed by Verschueren et al., that contains five stages comprising Scientific Foundations, Design Foundations, Development, Validation, and Implementation (Verschueren, Buffel, & Vander Stichele, 2019). The Scientific Foundation stage focuses on determining the goals or expected outcomes of the research, and the formulation of the hypotheses or underlying models required to achieve these goals. The choice of game mechanics, design, and technological features in the Design Foundations stage is then guided by these scientific foundations, prior to the practical Development, Validation, and Implementation stages. The SGH framework acts only as a general heuristic, as it does not identify the potential game mechanics or game tasks, nor does it specify how to appropriately select them to achieve the intended goals. Although the framework only pertains to SGH, the same principles can also be applied and translated into VR and mental health contexts thus supporting the development of the VR analytics framework proposed in this paper.

The Evidence Centered Design (ECD) also provides a conceptual design framework, focusing on the production of educational assessments based on the principles of evidentiary reasoning and delivery (Mislevy, Almond, & Lukas, 2003). ECD in its simplest form can be represented by three fundamental models referred to as Competency, Evidence and Task models. The Competency Model defines the variables and attributes that need to be measured in relation to an educational construct. The Evidence Model, on the other hand, establishes how to update the competency model variables based on the observed actions as evidence. Finally, the Task Model describes the tasks or actions useful for collecting the necessary evidence. Drawing parallels from the ECD, we proposed the Behavioral Domains analogous to the Competency Model which defines the variables or attributes that need to be measured, pertinent to symptoms of mental disorders. Similarly, using VR tasks to measure the behavioral domains aligns with the Task Model of ECD.

Mandryk & Birk suggested the concept of using game data from commercial off-the-shelf games as digital biomarkers for assessing and modelling mental health (Mandryk & Birk, 2019). Because game data represents a rich data source that reflects the low-level cognitive and motor processing of the game players, it was argued that this kind of data has the potential to provide direct insights into mental health, and potentially even track possible decline. Five game-based digital biomarker categories of behavior, cognitive performance, motor performance, social behavior, and affect were proposed. The main issue with this approach is that the game data collection seems inefficient and arbitrary in the sense that it relies on collecting as much game data as possible from all available sources without proper direction or guidance from a psychological standpoint.

Despite the strengths of these existing frameworks, a fundamental aspect missing from each is the inclusion of measurement variables which is critical when developing digital applications such as games or VR for mental health. Measurement can be quite challenging especially for interdisciplinary studies because it represents the intersection of theory and technical implementation. Often, studies are well-planned yet pay little attention to the measurement variables (Ekkekakis, 2013) resulting in poor outcomes. However, measurement data informs both which psychological features should be quantified, and which VR task elements can be optimally used in order to successfully index. Designing VR tasks for mental health requires a comprehensive understanding of the underlying symptoms, their corresponding behavioral domains and their manifestations within the VR environment. Although there are a handful examples of mapping models available in the literature (Arnab et al., 2015; Lim et al., 2015), none were specifically designed for VR applications or mental health care. Thus, our proposed VRAM framework addresses an important gap with the help of behavioral domains (measurement variables) and the example VR tasks.

The idea behind harnessing VR analytics was established from behavior sensing or digital phenotyping for measuring, monitoring, and predicting mental disorders (Insel, 2017). There is a growing literature that shows how objective behavior sensing or digital phenotyping can be applied to a wide range of mental disorders including anxiety, depression, and schizophrenia (I. Barnett et al., 2018; Di Matteo et al., 2021; Jacobson, Weingarden, & Wilhelm, 2019; Meyer et al., 2018; Wang et al., 2018). The term digital phenotyping was first coined by (Torous, Kiang, Lorme, & Onnela, 2016) as "*Moment-by-moment quantification of the individual-level human phenotype in situ, using data from personal digital devices such as smartphones*". The core premise behind digital phenotyping is that the digital traces left behind on social media, smartphones, or other personal digital devices may be leveraged to gain fine-grained insights into behaviors pertinent to an individual's mental state. This data can then be used to identify novel digital biomarkers or behavioral correlates of mental disorders (Aung, Matthews, & Choudhury, 2017; Mohr, Zhang, & Schueller, 2017; Spinazze, Rykov, Bottle, & Car, 2019). For example, data from Global Positioning System (GPS) has been shown to be a potential behavioral correlate of depression (Saeb et al., 2015, 2017) which makes sense from a clinical perspective since depressed individuals often present with a loss of interest, become more socially withdrawn and are more likely to stay indoors (Soong, 2021) than the healthy population.

Nonetheless, such prior work establishes a precedent for the potential of VR analytics i.e., the data captured from VR headsets, sensors, and crucially the participant interactions within the virtual environment itself, for detecting or predicting mental disorders. The use of VR not only offers several benefits of immersion, increased engagement and more ecologically valid assessments in a controlled setting over these other mediums as discussed earlier, but also offers reduced reliance on sensitive information such as GPS, call logs or text messages. Such sensitive information is necessary for digital phenotyping approaches resulting in greater ethical concerns such as data insecurity and surveillance, lack of informed consent, or abuse of trust (Gooding, 2019; Kosinski, Stillwell, & Graepel, 2013; Martinez-Martin, Insel, Dagum, Greely, & Cho, 2018; Montag, Sindermann, & Baumeister, 2020).

In summary, the existing frameworks provide high-level guidance for the development of their respective systems but were unspecific regarding the measurement variables, nor were they developed specifically for applications in VR. Prior work has demonstrated the possibilities for digital analytics to be used in the detection and prediction of symptoms of mental disorders which could also be translated and applied to VR. Furthermore, VR appears to offer particular advantages concerning privacy, data security, ecological validity, and immersion over other platforms for using digital analytics to diagnose or predict mental disorders.

4. Proposed conceptual framework

Fig. 1 introduces the proposed Virtual Reality Analytics Map (VRAM), a six-step roadmap to support the design and development of VR applications integrated with VR analytics for detecting symptoms of mental disorders. The development of the VRAM framework was supported by a thorough review of literature in the fields of games, VR, education, digital phenotyping, and mental health with continuous refinements over time after expert consultations.

The following sections explain each of the six steps of the framework in detail.

4.1. Step 1 – research goals

The first step of the framework encompasses the psychological aspects and emphasizes identifying the three crucial research goals of selecting the mental disorder, its related symptoms, and the target audience prior to the development of the VR application. It is advised to identify the mental disorder of interest and its related symptoms adhering to standardized diagnostic criteria as defined in the DSM-5 or ICD-11. Identification of both the psychological and physiological symptoms will most often be necessary in order to understand whether these symptoms and their related feelings, cognition or behaviors can manifest within the virtual environment. For example, socially avoidant behaviors, a key symptom of social anxiety, can be translated and tracked in virtual environments through gaze behavior or eye-tracking data of participants in the presence of virtual avatars (M. Dechant,

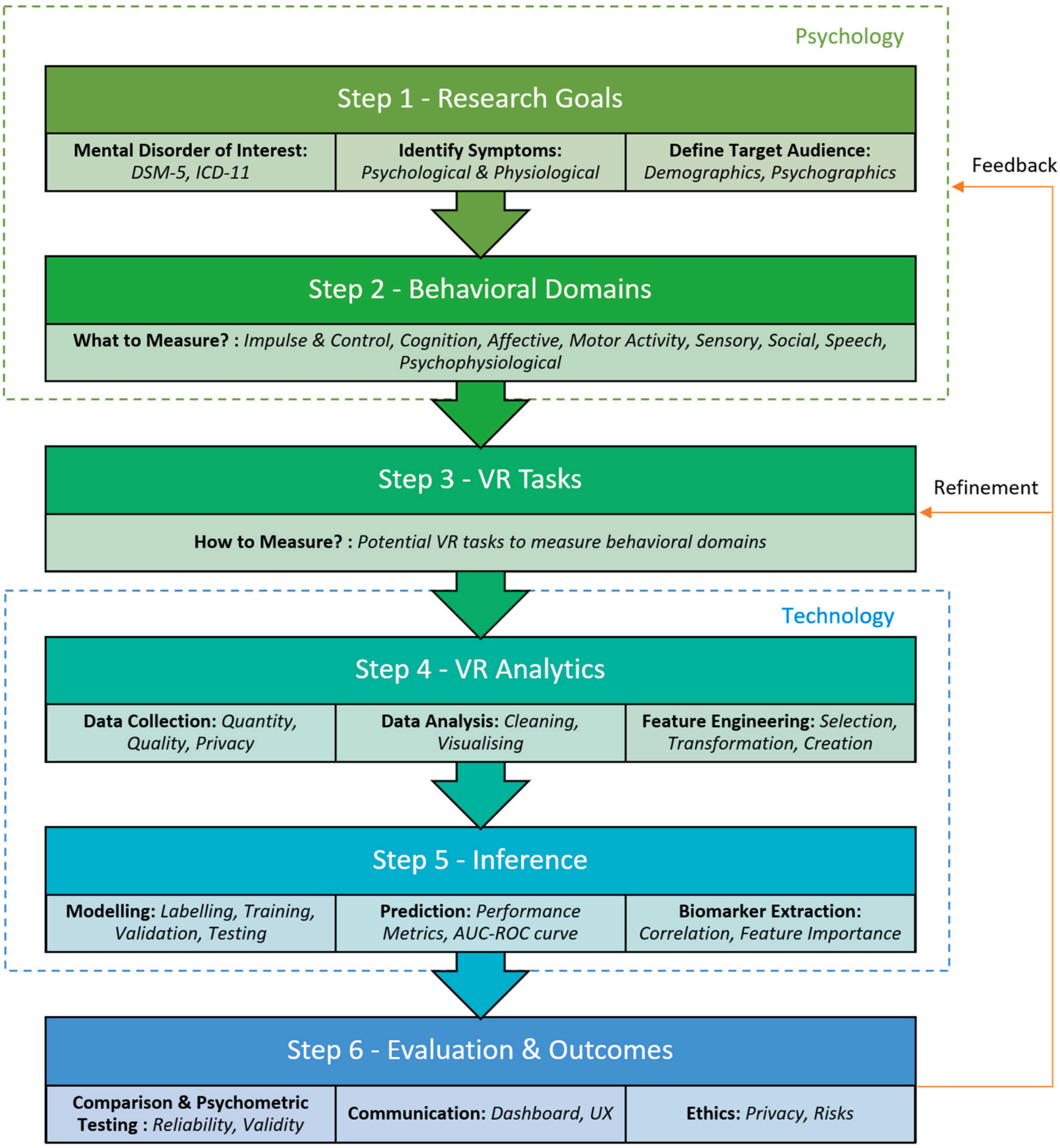

Fig. 1. Proposed virtual reality analytics Map (VRAM) conceptual framework.

Trimpl, Wolff, Mühlberger, & Shiban, 2017; Konovalova, Antolin, Bolderston, & Gregory, 2021; Rinck et al., 2010). However, it may not be possible to directly capture other symptoms such as rumination via this medium.

Similarly, determining the target audience is crucial to any design process. The target audience can be determined based on the general demographic or detailed psychographic information. Demographics information includes age, gender, ethnicity, and background while psychographic information includes traits, behaviors, actions, attitudes, and emotional responses. For the same mental illness, certain symptoms might be limited only to a specific target group. Therefore, it is recommended that a mental health professional be consulted from the outset to better understand the chosen mental disorder and the manifestations of its symptoms within different target groups. For example, cognitive difficulties may be more prominent in older people suffering from depression than in younger people (Murman, 2015). Hence, it is strongly advised against the development of a VR application without first determining these research goals. Modifying the application post-hoc to accommodate these research goals later might undermine its validity and reliability.

4.2. Step 2 – behavioral domains (what to measure?)

Building upon the research goals, the second step focuses on identifying and defining specific variables to be measured. Thus, the VRAM framework proposes several multidimensional variables or domains namely *Impulse & Control*, *Cognition*, *Affective*, *Motor Activity*, *Sensory, Social, Speech and Psychophysiological* as shown below in Fig. 2. Behavioral domains here refer to the various aspects that characterize the symptoms and behaviours pertinent to a mental disorder. These domains answer the question of "*What to measure?*" and are derived from the established DSM-5 and the Mental Status Examination (MSE) (Martin, 1990; Norris, Clark, & Shipley, 2016). The findings from the MSE, when integrated with demographic information and psychiatric history, form the basis for establishing a diagnosis according to DSM-5 criteria.

Accurately identifying and operationalizing these behavioral

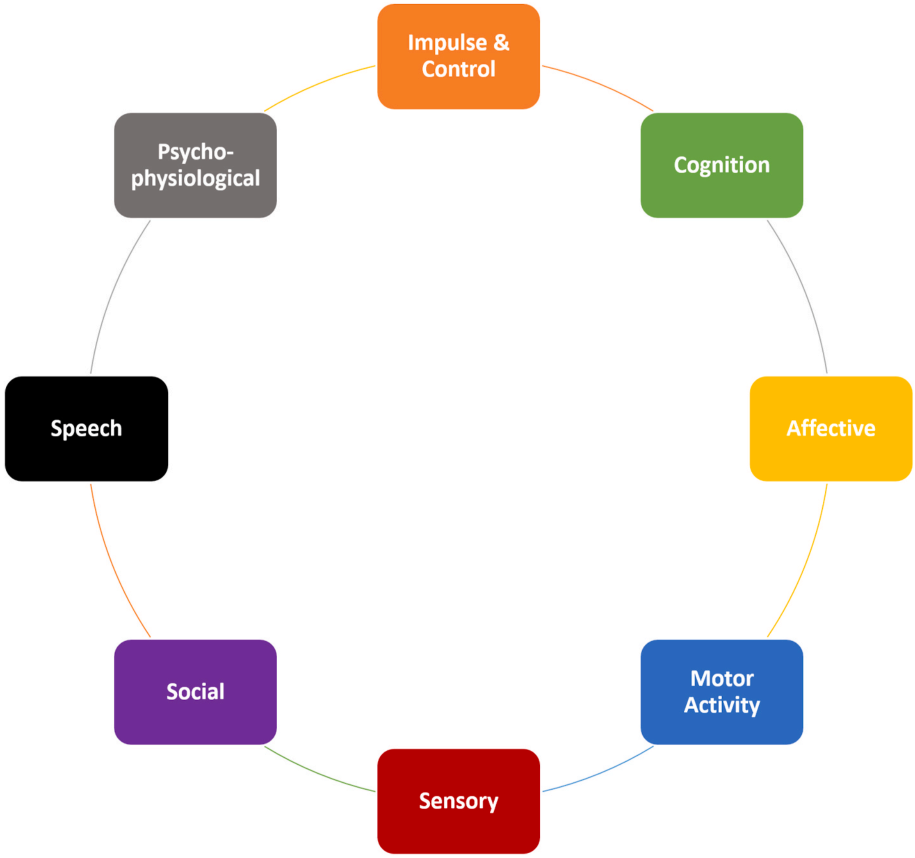

Fig. 2. Proposed behavioral domains.

domains in the context of VR is vital considering how each might manifest within the virtual environment. For instance, in the case of Attention-Deficit Hyperactivity Disorder (ADHD), the domain of Cognition, specifically Attention is crucial as it is associated with ADHD symptoms of sustained attention, difficulty concentrating or getting easily distracted. In the case of Major Depressive Disorder (MDD), the Affective domain is associated with the symptoms of mood changes or anhedonia. Likewise in the case of Autism Spectrum Disorder (ASD), communication difficulties or speech abnormalities are the main symptoms that can be associated with the Speech domain.

It is important to consider that many mental disorders have overlapping characteristics, while comorbid disturbances are also common, therefore symptoms can be associated with multiple domains. Table 1 presents a mapping matrix that maps the symptoms of common mental disorders to the proposed behavioral domains. It is important to highlight that even though this table was formulated based on DSM-5 and expert consultation, it can be subjective and is open to continuous refinement as new empirical findings emerge.

4.3. Step 3 – VR tasks

Once the symptoms and the associated behavioral domains have been identified, it is crucial to select and implement VR tasks that effectively measure these domains. This step in the framework essentially answers the question of *"How to Measure?"*. The selection of appropriate VR tasks requires careful design that balances ecological validity, domain fidelity and technological limitations while maintaining experimental control. For example, a task implemented to assess social anxiety might focus on detecting avoidant behaviours by placing users in virtual social scenarios such as interacting with virtual characters and tracking their physiological responses, eye-tracking, gaze behaviour information and other metrics (Binder & Spoormaker, 2020; Dechant et al., 2017; Kampmann, Emmelkamp, & Morina, 2018; Konovalova et al., 2021; Lange & Pauli, 2019; Yaremych & Persky, 2019). Here, the VR task should strive to create a sense of social presence and pressure that would mimic the real-world anxiety inducing situation. This virtual scenario provides a controlled setting mimicking a real-world environment thereby potentially enhancing the generalizability of findings. The technical considerations here would be regarding creating a realistic environment with virtual characters and environment that may be computationally expensive. The challenge here is to create immersive and engaging VR tasks but also to make sure that they do not interfere with the validity and reliability of the measurements (Levy, Lambeth, Solomon, & Gandy, 2018).

Table 2 lists some of the potential VR tasks through which the proposed behavioral domains can be measured within the virtual environment. This list was generated based on a thorough review of the literature, expert consultation, incremental updates, and the research team's domain expertise.

4.3.1. VR tasks to measure behavioral domains

Impulse and control domain refers to the observable actions, series of actions and reactions of users within the virtual environment. This domain encompasses numerous behaviours including impulsive, risk-taking, avoidant, inhibition, competitive, irritability/frustration and more that can manifest and quantified withing the virtual world. Example VR tasks to measure this domain include Go/No-go task (Gomez, Ratcliff, & Perea, 2007) based on Barkley's theory of ADHD (Barkley, 1997) which has been adapted to VR with promising findings (Kirkham et al., 2024; Liu et al., 2022; Mendez-Encinas, Sujar, Bayona, & Delgado-Gomez, 2023). As mentioned previously, avoidant behaviours can be measured through VR tasks that place the users in social setting with virtual characters and tracking relevant metrics such as gaze behaviour, distance from virtual characters, interaction choices

Table 1
Symptoms-to-domains mapping matrix.

Symptom	Cognition	Affective	Sensory	Social	Impulse & Control	Motor Activity	Speech & Psycho-Physiological
Generalized Anxiety Disorder							
Excessive Worry	✓ (Decision Making)						✓
Restlessness	✓ (Attention)				✓	✓	
Fatigue		✓				✓	✓
Difficulty Concentrating	✓ (Attention)						
Difficulty making decisions	✓ (Decision Making)						
Irritability		✓			✓		
Depressive Disorders							
Muscle tension						✓	✓
Mood changes		✓					
Anhedonia		✓					
Feelings of guilt or worthlessness		✓					
Social Anxiety Disorder							
Social withdrawal				✓			
Avoidant behaviour				✓	✓		
Fear of being judged				✓			
Sweating, Trembling						✓	✓
Schizophrenia							
Hallucinations (Visual, Auditory)			✓				
Delusions	✓ (Decision-Making)		✓				
Catatonia						✓	
Attention Deficit Hyperactivity Disorder							
Hyperactivity	✓ (Attention)				✓	✓	
Impulsiveness	✓ (Decision-making)				✓	✓	
Disorganization	✓ (Processing)						
Easy distractions	✓ (Attention)				✓		
Forgetfulness in daily activities	✓ (Memory)						
Autism Spectrum Disorder							
Communication or Interaction difficulty	✓ (Processing)			✓			✓
Repetitive behaviour					✓		
Sensory sensitivity			✓				
Speech abnormalities	✓ (Processing)			✓			✓
Obsessive Compulsive Disorders							
Compulsiveness	✓ (Decision-making)				✓		
Symmetry and Order	✓ (Processing)				✓		
Panic Disorders or Phobias							
Racing Heart							✓
Difficulty Breathing							✓
Trembling						✓	✓
Tingling, Numbness in fingers						✓	✓

(engaged or avoided) among a few which have been reported in several studies (Binder & Spoormaker, 2020; Dechant et al., 2017; Kampmann et al., 2018; Konovalova et al., 2021; Lange & Pauli, 2019; Yaremych & Persky, 2019).

Cognitive difficulties are a common feature of both neurological and psychiatric illnesses, although can vary considerably in both nature of the dysfunction and severity. There is substantial evidence in broader literature showing how VR can be used to assess or detect cognitive functioning (Kirkham et al., 2024; Neguţ et al., 2016; Rus-Calafell, Garety, Sason, Craig, & Valmaggia, 2018). VR adaptations of tasks like Balloon Analog Risk Task (BART) (Lejuez et al., 2002) and Iowa Gambling Task (IGT) (Buelow & Suhr, 2009) can be implemented to assess both the decision-making process and risk-taking behaviours (de-Juan-Ripoll, Soler-Domínguez, Chicchi Giglioli, Contero, & Alcaniz, 2020; Ju & Wallraven, 2023). Similarly, decision-making can also be potentially assessed using scenarios that present moral dilemmas possibly with the presence of time-pressure requiring uses to make quick and complex decisions (Arrambide et al., 2022; Francis et al., 2016; Zanon, Novembre, Zangrando, Chittaro, & Silani, 2014). Attention assessment has also benefited from VR adaptations of established measures (Climent et al., 2021; Gould et al., 2007; Parsons, 2015; Parsons, Courtney, Arizmendi, & Dawson, 2011; Rodríguez, Areces, García, Cueli, & González-Castro, 2018). For example, Parsons and Barnett validated a VR version of the Stroop test (Parsons & Barnett, 2018; Parsons et al., 2011) while Iriarte et al. (Iriarte et al., 2016) developed a Continuous Performance Test (CPT) which proved effective in measuring sustained attention. Similarly, other adaptations such as VR versions of Trail Making or Color Trail tests (Kirkham et al., 2024; Plotnik et al., 2021) or tasks involving virtual supermarket or kitchen settings (M. D. Barnett, Childers, & Parsons, 2021; Giovannetti et al., 2019; Porffy et al., 2022), virtual classrooms or spatial navigation tasks (Gould et al., 2007; A. A. Rizzo et al., 2000; Spieker, Astur, West, Griego, & Rowland, 2012; Yeh et al., 2020) have all reported promising results for the assessment of different domains of cognitive functioning.

The affective domain here relates to the emotion regulation, mood and engagement which can be measured by tracking engagement levels in tasks for completing certain activities. For example, anhedonia a core symptom of several mood disorders such as MDD refers to lack of pleasure from activities that bring joy, and such individuals often have difficulty engaging in and deriving pleasure from daily activities. Therefore, tracking how long user spends and engages with certain activities or tracking user's responses to rewarding stimuli (reward-based tasks that offer coins, points or rewards) and their motivation to persist (performance metrics) along with their physiological responses can potentially provide a way to measure the engagement levels (Lahiri, Bekele, Dohrmann, Warren, & Sarkar, 2013). Further, emotion recognition tasks, where users are asked to identify emotions from the expression of virtual characters or emojis can provide insights into how users perceive and interpret emotional cues (Geraets et al., 2021; Johannes Dechant, Frommel, & Mandryk, 2021; Riva et al., 2007; Tabbaa et al., 2021).

Unusual or abnormal motor patterns are also symptoms of some mental disorders. Motor skills can be easily translated into VR through some form of input-based tasks that can be tracked and measured continuously. Precision tasks in VR such as painting or writing provide a controlled approach to measure fine motor skills (Aldridge & Bethel, 2021; Christou, Michael-Grigoriou, Sokratous, & Tsiakoulia, 2018; Morales González, Freeman, & Georgiou, 2020). While larger body movement tasks such as exergames that simulate real-life activities can facilitate the measurement of gross motor skills, balance and overall coordination (Flores-Gallegos, Rodríguez-Leis, & Fernández, 2022; Pallavicini & Pepe, 2020; Xu et al., 2023; Yen & Chiu, 2021).

Many mental disorders often exhibit overlapping symptoms or are frequently comorbid. For instance, conditions such as anxiety disorders, PTSD or OCD share several overlapping symptoms including avoidance behaviours, hypervigilance (attentional biases) and cognitive difficulties. These shared symptoms allow for the design of VR tasks that can assess broad transdiagnostic symptoms across multiple disorders. However, many mental disorders also have core symptoms that are relatively unique, and that therefore require specific design and tuning of VR tasks. For example, a core symptom of depression is anhedonia, which is not commonly observed in anxiety disorder or PTSD.

The flexibility of VR systems is such that it allows for the implementation of tailored VR tasks for the general assessment of symptoms while also simultaneously allowing for the inclusion of disorder-specific elements. For example, a VR task that is trying to capture common symptoms shared by many anxiety disorders might involve participants asked to simply navigate a virtual environment that evaluates avoidance behaviours in general. However, this same task can also be modified to include stimuli that are specific to OCD, such as environments that trigger compulsive behaviours, whereas a trauma-related exposure could be appropriate for assessing PTSD symptoms. This adaptability allows VR tasks to potentially serve as a transdiagnostic tool applicable to common symptoms across multiple mental disorders while also being flexible enough to capture the core symptoms of specific mental disorders.

4.3.2. VR task implementation considerations

Implementing VR tasks can present a few challenges in terms of usability, accessibility and other technical limitations that must be considered. Usability and accessibility are of paramount importance for diverse participant populations. Usability refers to how effectively and efficiently users can interact with the VR system by designing user-friendly interfaces and providing clear instructions to help participants complete the tasks as required. This facilitates ease of navigation and task completion, reducing cognitive load and minimizing frustration which can be potential confounding factors. For instance, if a task is designed to measure irritability or frustration of the users, it is crucial that this measurement stems from the task rather than the technical difficulties. This results in enhanced validity as the observed measurements can be attributed to the constructs being measured rather than confounding factors such as poor interface design or other technical issues.

Training modules play a vital role when designing VR systems as they are responsible for familiarizing participants with the VR environment and reducing performance variability ensuring they understand the task requirements leading to more reliable and valid data collection. Motion sickness or cybersickness characterized by nausea or dizziness can commonly arise during VR experience. Rebenitsch and Owen (Rebenitsch & Owen, 2016) reviewed various factors contributing to cybersickness that includes a mismatch between the visual and vestibular systems, eye strain, and latency. Reducing motion sickness is crucial for ensuring participant safety as well as for maintaining data integrity. Therefore, optimizing frames per second (FPS) and minimizing latency can be effective in reducing motion sickness. A smooth and consistent FPS helps prevent motion blur, visual artifacts and reduces lag that can negatively affect the VR experience.

Sometimes technical limitations might pose a significant challenge when implementing a said VR task. For instance, implementing a VR task of emotion recognition from facial expressions of virtual characters might be big ask considering the rendering requirements of facial expressions that might require modelling blend shapes or motion capture data which can be computationally as well as resource intensive especially if premade assets are unavailable. In such cases, simpler yet effective implementations might ease the computational requirements while maintaining the task integrity. For example, Dechant et al. demonstrated a method whereby the emojis displayed on top of virtual characters represented their expressions/emotions simplifying the task implementation without the need for realistic avatar animations or realistic facial expressions (Johannes Dechant et al., 2021).

It is important to highlight that in some cases, there may be other simpler and inexpensive alternatives available to measure the behavioral domains that do not require expensive VR setups. These might include the traditional methods or basic computerized tasks that can provide adequate measurements without the need for extensive resources. However, the advantage of using VR is that it provides immersion, engagement, replication, and a controlled setting that increases the ecologically valid of assessments potentially yielding detailed and richer quality of data (Freeman et al., 2017; Pan & Hamilton, 2018). However, crucially these VR tasks need to be validated with extensive studies to establish their reliability, validity and generalizability and ideally should be an iterative process with continuous refinement based on the user feedback and empirical findings. Nonetheless, the example tasks demonstrate the potential of VR systems and serve as an excellent starting point for the design and development process.

4.4. Step 4 – virtual reality analytics

The fourth step in the VRAM framework involves the use of VR analytics to understand, visualize, and analyze the data generated from the VR tasks. Step 4 therefore focuses on the analytics part which includes the data collection, analysis, and feature engineering processes for building the inference models. Data collection encompasses considerations of quantity, quality and privacy. The quantity of data collected must be adequate to draw meaningful insights. It is imperative to preplan the data collection process as overcollection of data during the VR session can lead to problems such as frame rate drops, or lags known to cause motion sickness as mentioned earlier.

Data quality implies that the acquired data contains enough information to explain the context in which the event occurred, as well as enough detail to relate the data to a specific design element. This makes it easier to identify the key data metrics and design elements which are associated with the symptoms of interest. Data should be unambiguous, fine-grained, and as explicit as possible. For instance, collecting a data point such as a “button press” is not as helpful as an “exit button press”. Raw data collected often contains noise and artifacts that need to be handled before any analyses have been made. There are several ways for data to be corrupted, duplicated or incorrectly labelled when obtained from multiple sources. Data preprocessing involves cleaning or transforming data which includes filtering our outliers, correcting for sensor errors and normalizing the data. Data cleaning typically involves fixing corrupt data entries, removal of duplicates, or incorrectly formatted data points. Privacy and ethical issues are a critical concern during this step. Researchers are urged to use due diligence to ensure that no personally identifiable data is collected for analysis. Furthermore, even if basic data like demographics is accessible, the analyses should be conducted on anonymized data to protect user identities.

The quality of inference models produced depends on the quality of the features extracted from the collected data. Feature engineering is the process of selecting, transforming, or creating new features from the gathered data based on domain expertise. Since it is quite challenging and critical to determine which are the best features, oftentimes feature selection algorithms, exploratory data analysis, visualization using dimensionality reduction algorithms such as Principal Component Analysis (PCA) (Jolliffe & Cadima, 2016) or t-Stochastic Neighborhood Embedding (t-SNE) (van der Maaten & Hinton, 2008) and other techniques are used for better understanding and finding patterns within the data to assist in the feature engineering process.

4.5. Step 5 – inference

Inference involves applying statistical methods, training machine learning or even deep learning models depending on the quality and quantity of the collected data. Ground truth labels for training the inference models can be obtained through widely used self-report measures for the respective mental illnesses, or preferably, clinical interviews conducted by a qualified mental health professional. Several machine learning algorithms can be chosen based on the understanding of the core data, and the features or attributes of the dataset. Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting and other algorithms have been successfully used for the prediction of mental illnesses in the context of VR or games (Heller et al., 2013; Šalkevicius, Damaševičius, Maskeliunas, & Laukienė, 2019; Handouzi, Maaoui, Pruski, Moussaoui, & Bendiouis, 2013; Yeh et al., 2020).

The performance of these inference models can be evaluated using the performance metrics such as accuracy, precision, recall, F1-scores, and confusion matrix or by plotting the Area Under ROC curve (AUC - ROC) where ROC stands for receiver operating characteristic. The performance metrics will provide insights into the false positives and false negatives which represent important evaluation metrics for the assessment scores or detecting symptoms of mental disorders. An in-depth discussion on inference modelling is beyond the scope of this paper and researchers are urged to refer to the available resources (Japkowicz & Shah, 2014; Zheng, 2015).

Along with predictive power, it is important to discover the most important and influential features or data attributes that contribute towards the modelling of mental disorders. These features might be individual or aggregated traits derived from the gathered data. Since very little is known about such features due to the lack of studies in VR and mental health prediction, knowledge of such features presents a major milestone toward developing evidence-based customized VR applications for modelling, predicting, and evaluating the VR therapy solutions of mental disorders.

4.6. Step 6 – outcomes

The final Step 6 specifies how to evaluate and communicate the observed results to participants and clinicians. To achieve this objective, we advise consulting with trained mental health professionals for communication of outcomes that might have implications for participant safety. Outcomes can be presented to the clinicians in the form of dashboards or can be conveyed to the participants themselves in the form of visualizations, or assessment scores. If feasible, it will be beneficial to translate the outcomes of the inference models into the conventional output scores from assessments such as the Patient Health Questionnaire (PHQ-9) or GAD-7 (Spitzer et al., 1994) that are informative to the researchers, practitioners as well as the patient themselves.

Evaluation involves assessing the psychometric properties (validity, reliability) of the VR tasks as well as incorporating feedback from the users and other stakeholders to improve the usability and refine the framework. Scientific validity in research refers to the degree to which a method measures what it claims to measure whereas scientific reliability describes how consistently a method can measure something (Gidron, 2013).

Feedback from the participants is vital to understand the usability and the user experience. Usability evaluations can be conducted using measures such as System Usability Scale (SUS) (Bangor, Kortum, & Miller, 2008; Bangor et al., 2008) or the User Experience Questionnaire (UEQ) (Laugwitz, Held, & Schrepp, 2008). Engaging other stakeholders such as clinicians and professionals ensure that the VR tasks are aligned with the research goals. The feedback loop acknowledges the iterative nature of this process. Insights gained from evaluation and outcomes can inform the redesign of VR tasks, allowing for continuous improvement of the framework.

5. Example application

To highlight how the proposed VRAM conceptual framework might be useful to researchers and practitioners, the following table showcases how the framework can be applied to predict symptoms of various mental health disorders, including Generalized Anxiety Disorder (GAD), Major Depressive Disorder (MDD), Attention Deficit/Hyperactivity Disorder (ADHD) and Schizophrenia. The reason we decided to demonstrate the usefulness of the framework for these disorders and especially for GAD was because the literature around GAD detection using VR is scarce (Chitale, Baghaei, et al., 2022; Freeman et al., 2017) but is a common mental health condition, and therefore needs to be an important priority in clinical management or therapy applications.

Table 3 lists the various symptoms and their associated behavioral domains that need to be measured using the example VR tasks by referring to Tables 1 and 2 from step 2 and step 3 of the VRAM framework respectively. For each disorder, the framework systematically aligns relevant symptoms with corresponding VR tasks to measure specific psychological constructs. This instantly presents plenty of possible types of VR tasks, activities or scenarios that can be either implemented separately or together to achieve the intended research goals. These examples illustrate the generalizability of the VRAM framework across a wide range of mental disorders as well as demonstrating the value and utility of the proposed framework by ensuring a standardized design process in designing VR systems for detecting symptoms of mental disorders. The example tasks at this stage are still hypothetical and future research should look at empirical testing to support their effectiveness and reliability.

6. Limitations and future prospects

Although this paper proposes a novel conceptual framework using

Table 2
Examples of VR tasks to measure proposed behavioral domains.

Behavioral Domains	Examples of VR Tasks to measure domains
Impulse & Control	Go/No-go task in VR, social interactions or exposure to feared stimuli with options to engage, avoid or forced choice. Tasks similar to VR Balloon Analog Risk Tasks (BART) or VR Iowa gambling Tasks (IGT) or VR Symmetry and Ordering tasks.
Cognition – Decision Making	Decision-making tasks under time pressure (moral dilemmas, escape room puzzles) or VR BART and IGT.
Cognition - Attention	Tasks that require the players to focus on the objective while ignoring distractions (e.g., eye-tracking a moving orb), Trail making test or Color Trails Test in VR setting, VR Stroop Test, or VR Continuous Performance Test
Cognition – Processing Speed	Tasks requiring users to respond quickly to stimuli (e.g., VR games like Whack-A-Mole, Beat Sabre), VR cancellation tasks
Cognition – Memory	Spatial navigation tasks to remember and locate objects in a virtual environment or sequence recall tasks such as returning correct objects to NPC avatars
Affective	Can be measured through engagement levels in tasks (painting, playing musical instrument) or reward-based tasks (coins, points). Emotion recognition tasks such as identifying the emotions from facial expressions of avatars.
Motor Activity	Precision tasks using VR controllers (painting, writing) or larger body movement tasks (VR exergames)
Social	Interaction tasks with virtual characters such as greetings, social cues recognition, or maintaining eye contact
Sensory	Visual or sound recognition tasks. Responses to altered sensory inputs, interaction with virtual mirrors, reality vs illusion perception tests
Speech	Public speaking tasks, VR object naming tasks
Psychophysiological	Tasks involving induced stressor in virtual environment and measuring the physiological signal responses of users

Table 3
Application of VRAM framework to predict Symptoms of GAD, ADHD, MDD and Schizophrenia.

Disorder	Symptoms	Behavioral Domains	VR Task Examples
Generalized Anxiety Disorder (GAD)	Excessive Worry	Cognition (Decision Making), Psychophysiological	Decision-making tasks under time pressure, Tasks involving induced stressor
	Difficulty concentrating	Cognition (Attention)	VR Stroop Test, VR CPT, VR Trail Making Tests
	Restlessness and Fatigue	Cognition (Attention), Motor Activity	Precision tasks using VR controllers, Tasks that require the players to focus on the objective while ignoring distractions
	Irritability	Affective, Impulse & Control	Engagement level tracking in challenging tasks, VR BART
Attention Deficit Hyperactivity Disorder (ADHD)	Easy Distractions	Cognition (Attention)	VR Trail making or color trail test, VR Stroop, VR CPT
	Impulsivity	Impulse & Control	Go/No-go task in VR
	Disorganization & Slow Processing	Cognition (Processing Speed)	Tasks requiring users to respond quickly to stimuli (e.g., VR games like Whack-A-Mole)
	Forgetfulness	Cognition (Memory)	Spatial navigation tasks to remember and locate objects in a virtual environment
Major Depressive Disorder (MDD)	Anhedonia & Mood Changes	Affective	Tracking engagement levels in VR tasks like (painting, playing musical instrument) or reward-based tasks (coins, points)
	Muscle Tension	Motor Activity & Psychophysiological	Exergames to evaluate psychomotor activity, physiological measures during added stressors
Schizophrenia	Hallucinations	Sensory	Reality vs illusion perception tests in VR, auditory/visual stimuli recognition tasks.
	Catatonia	Motor Activity	Larger body movement tasks (VR exergames)

VR analytics for the detection or prediction of mental disorders, at this stage it is still theoretical. Future validation studies are needed to provide evidence of support for our proposed framework. Such studies need to establish strong convergence between VR applications and existing assessment tools as well as between the VR tasks and the symptoms of mental disorders to identify possible novel digital biomarkers. For the same reason, case studies have been planned which will evaluate the framework by analyzing existing research through the proposed framework's lens. These studies will critically assess the proposed VRAM framework's ability as an analysis tool. A series of experimental evaluations as a part of a larger project are planned which will employ the proposed VRAM framework for detecting anxiety and depression disorders.

An important aspect of inference modelling to consider is the concept of overdiagnosis. Overdiagnosis can be harmful and lead to psychological stress or unnecessary treatments (Gyuricza et al., 2018). Due to the nature of inference models, it is common to encounter the problem of overdiagnosis due to the presence of false positives. False positives in the prediction or diagnosis contexts indicate a false presence of a mental illness when in reality there is none. This happens due to the prediction error which can never be completely eliminated. Thus, it is advisable to exercise appropriate care by using other available resources (I. Barnett et al., 2018; Huckvale, Venkatesh, & Christensen, 2019) to minimize the potential for false positives. Another factor that is important to consider is the level of experience of the users interacting with VR applications. Inexperience or unfamiliarity can produce outliers or abnormalities in data patterns and compromise the predictions. Therefore, VR applications should include a testing or training module to increase the familiarity of users with the VR equipment, and the virtual environment to reduce the impact of inexperience as mentioned in earlier sections. Finally, we consider one of the most promising prospects of this approach to be that of personalized modelling. Personalized models compared to global models have the potential to produce more accurate predictions and identify more relevant risk factors (Ng, Sun, Hu, & Wang, 2015). If digital biomarkers and strong correlations of VR tasks with symptoms of mental disorders can be established, customized VR scenarios and applications can be developed according to the needs of each individual user.

7. Conclusion

This paper hypothesizes that VR analytics or tracking, visualizing, and analyzing VR data can lead to the identification of novel digital biomarkers indicative of symptoms of mental disorders. As a result, we propose a novel VRAM conceptual framework which provides a roadmap to researchers in designing and developing VR applications for detecting symptoms of mental disorders. The framework consists of six steps namely research goals, behavioral domains, VR tasks, analytics, inference, and outcomes. We believe that this approach of leveraging VR analytics holds great promise, but it must be directly tested with future validation studies. We hope that the insights from this paper will be valuable to future researchers and practitioners in designing VR scenarios and applications for detecting novel digital biomarkers of symptoms of mental disorders. It is important to emphasize that although this approach seems promising, VR analytics cannot be considered a direct replacement for traditional assessment tools. Instead, VR analytics can complement traditional methods with additional psychological as well as physiological insights. Combining the strengths of both approaches can assist clinicians obtain a more thorough understanding of an individual's mental state, enabling them to adapt the treatment plans accordingly. We believe the contributions of this paper will pave the way for large-scale efficacy testing, clinical application, and potentially cost-effective use of VR analytics for modelling, predicting, treating, and monitoring mental disorders in future.

FUNDING sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Vibhav Chitale: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Julie D. Henry:** Writing – review & editing, Supervision, Methodology, Formal analysis. **Hai-Ning Liang:** Writing – review & editing, Visualization, Methodology. **Ben Matthews:** Writing – review & editing, Supervision, Investigation. **Nilufar Baghaei:** Writing – review & editing, Supervision, Resources, Project administration, Methodology.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

Aldridge, A., & Bethel, C. L. (2021). A systematic review of the use of art in virtual reality. *Electronics, 10*(18), 2314. https://doi.org/10.3390/electronics10182314

Althubaiti, A. (2016). Information bias in health research: Definition, pitfalls, and adjustment methods. *Journal of Multidisciplinary Healthcare, 9*, 211–217. https://doi.org/10.2147/JMDH.S104807

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders*. American Psychiatric Association Publishing. https://doi.org/10.1176/appi.books.9780890425787

Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., de Freitas, S., Louchart, S., et al. (2015). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology: Journal of the Council for Educational Technology, 46*(2), 391–411. https://doi.org/10.1111/bjet.12113

Arrambide, K., Yoon, J., MacArthur, C., Rogers, K., Luz, A., & Nacke, L. E. (2022). "I don't want to shoot the android": Players translate real-life moral intuitions to in-game decisions in Detroit: become human. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3491102.3502019. Article Article 469.

Aung, M. H., Matthews, M., & Choudhury, T. (2017). Sensing behavioral symptoms of mental health and delivering personalized interventions using mobile technologies. *Depression and Anxiety, 34*(7), 603–609. https://doi.org/10.1002/da.22646

Baghaei, N., Chitale, V., Hlasnik, A., Stemmet, L., Liang, H. N., & Porter, R. (2021). Virtual reality for supporting the treatment of depression and anxiety: Scoping review. *JMIR mental health, 8*(9), Article e29681. https://doi.org/10.2196/29681

Balcombe, L., & De Leo, D. (2021). Digital mental health challenges and the horizon ahead for solutions. *JMIR Mental Health, 8*(3), Article e26811. https://doi.org/10.2196/26811

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction, 24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*(1), 65–94. https://doi.org/10.1037/0033-2909.121.1.65

Barnett, M. D., Childers, L. G., & Parsons, T. D. (2021). A virtual kitchen protocol to measure everyday memory functioning for meal preparation. *Brain Sciences, 11*(5). https://doi.org/10.3390/brainsci11050571

Barnett, I., Torous, J., Staples, P., Sandoval, L., Keshavan, M., & Onnela, J.-P. (2018). Relapse prediction in schizophrenia through digital phenotyping: A pilot study. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, 43*(8), 1660–1666. https://doi.org/10.1038/s41386-018-0030-z

Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in Clinical Neuroscience, 22*(2), 169–177. https://doi.org/10.31887/DCNS.2020.22.2/lvalmaggia

Binder, F. P., & Spoormaker, V. I. (2020). Quantifying human avoidance behavior in immersive virtual reality. *Frontiers in Behavioral Neuroscience, 14*, Article 569899. https://doi.org/10.3389/fnbeh.2020.569899

Buelow, M. T., & Suhr, J. A. (2009). Construct validity of the Iowa gambling task. *Neuropsychology review, 19*, 102–114.

Chitale, V., Baghaei, N., Playne, D., Liang, H. N., Zhao, Y., Erensoy, A., et al. (2022). The use of videogames and virtual reality for the assessment of anxiety and depression: A scoping review. *Games for Health Journal*. https://doi.org/10.1089/g4h.2021.0227

Chitale, V., Playne, D., Liang, H.-N., & Baghaei, N. (2022). Virtual reality data for predicting mental health conditions. *2022 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct)* (pp. 6–8). https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00011

Christou, C., Michael-Grigoriou, D., Sokratous, D., & Tsiakoulia, M. (2018). BuzzwireVR: An immersive game to supplement fine-motor movement therapy. *ICAT-EGVE 2018 - international conference on artificial reality and telexistence and eurographics symposium on virtual environments*. https://doi.org/10.2312/egve.20181327

Climent, G., Rodríguez, C., García, T., Areces, D., Mejías, M., Aierbe, A., et al. (2021). New virtual reality tool (nesplora aquarium) for assessing attention and working memory in adults: A normative study. *Applied Neuropsychology. Adult, 28*(4), 403–415. https://doi.org/10.1080/23279095.2019.1646745

Corcoran, C. M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., Javitt, D. C., et al. (2018). Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA), 17*(1), 67–75. https://doi.org/10.1002/wps.20491

D Alfonso, S. (2020). AI in mental health. *Current Opinion in Psychology, 36*, 112–117. https://doi.org/10.1016/j.copsyc.2020.04.005

Dechant, M., Trimpl, S., Wolff, C., Mühlberger, A., & Shiban, Y. (2017). Potential of virtual reality as a diagnostic tool for social anxiety: A pilot study. *Computers in Human Behavior, 76*, 128–134. https://doi.org/10.1016/j.chb.2017.07.005

de-Juan-Ripoll, C., Soler-Domínguez, J. L., Chicchi Giglioli, Contero, M., & Alcaniz, M. (2020). The Spheres & shield maze task: a virtual reality serious game for the assessment of risk taking in decision making. *Cyberpsychology, Behavior, and Social Networking, 23*(11), 773–781.

Di Matteo, D., Fotinos, K., Lokuge, S., Mason, G., Sternat, T., Katzman, M. A., et al. (2021). Automated screening for social anxiety, generalized anxiety, and depression from objective smartphone-collected data: Cross-sectional study. *Journal of Medical Internet Research, 23*(8), Article e28918. https://doi.org/10.2196/28918

Ekkekakis, P. (2013). *The measurement of affect, mood, and emotion: A guide for health-behavioral research*. Cambridge University Press. https://doi.org/10.1017/cbo9780511820724

Flores-Gallegos, R., Rodríguez-Leis, P., & Fernández, T. (2022). Effects of a virtual reality training program on visual attention and motor performance in children with reading learning disability. *International Journal of Child-Computer Interaction, 32*, Article 100394. https://doi.org/10.1016/j.ijcci.2021.100394

Francis, K. B., Howard, C., Howard, I. S., Gummerum, M., Ganis, G., Anderson, G., et al. (2016). Virtual morality: Transitioning from moral judgment to moral action? *PLoS One, 11*(10), Article e0164374. https://doi.org/10.1371/journal.pone.0164374

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., et al. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine, 47*(14), 2393–2400. https://doi.org/10.1017/S003329171700040X

Gavrilescu, M., & Vizireanu, N. (2019). Predicting depression, anxiety, and stress levels from videos using the facial action coding system. *Sensors, 19*(17). https://doi.org/10.3390/s19173693

Geraets, C. N. W., Klein Tuente, S., Lestestuiver, B. P., van Beilen, M., Nijman, S. A., Marsman, J. B. C., et al. (2021). Virtual reality facial emotion recognition in social environments: An eye-tracking study. *Internet Interventions, 25*, Article 100432. https://doi.org/10.1016/j.invent.2021.100432, 100432.

Gidron, Y. (2013). Reliability and validity. In M. D. Gellman, & J. R. Turner (Eds.), *Encyclopedia of behavioral medicine* (pp. 1643–1644). New York: Springer. https://doi.org/10.1007/978-1-4419-1005-9_1549.

Giovannetti, T., Yamaguchi, T., Roll, E., Harada, T., Rycroft, S. S., Divers, R., et al. (2019). The virtual kitchen challenge: Preliminary data from a novel virtual reality test of mild difficulties in everyday functioning. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition, 26*(6), 823–841. https://doi.org/10.1080/13825585.2018.1536774

Gkotsis, G., Oellrich, A., Velupillai, S., Liakata, M., Hubbard, T. J. P., Dobson, R. J. B., et al. (2017). Characterisation of mental health conditions in social media using Informed Deep Learning. *Scientific Reports, 7*(1), Article 45141. https://doi.org/10.1038/srep45141

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136*(3), 389–413. https://doi.org/10.1037/0096-3445.136.3.389

Gooding, P. (2019). Mapping the rise of digital mental health technologies: Emerging issues for law and society. *International Journal of Law and Psychiatry, 67*, Article 101498. https://doi.org/10.1016/j.ijlp.2019.101498

Gould, N. F., Holmes, M. K., Fantie, B. D., Luckenbaugh, D. A., Pine, D. S., Gould, T. D., et al. (2007). Performance on a virtual reality spatial memory navigation task in depressed patients. *American Journal of Psychiatry, 164*(3), 516–519. https://doi.org/10.1176/ajp.2007.164.3.516

Handouzi, W., Maaoui, C., Pruski, A., Moussaoui, A., & Bendiouis, Y. (2013). Short-term anxiety recognition induced by virtual reality exposure for phobic people. *IEEE International Conference on Systems, Man, and Cybernetics*, 3145–3150. https://doi.org/10.1109/SMC.2013.536

Hashim, N. W., Wilkes, M., Salomon, R., Meggs, J., & France, D. J. (2017). Evaluation of voice acoustics as predictors of clinical depression scores. *Journal of Voice: Official Journal of the Voice Foundation, 31*(2), 256.e1–256.e6. https://doi.org/10.1016/j.jvoice.2016.06.006

He, L., & Cao, C. (2018). Automated depression analysis using convolutional neural networks from speech. *Journal of Biomedical Informatics, 83*, 103–111. https://doi.org/10.1016/j.jbi.2018.05.007

Heller, M. D., Roots, K., Srivastava, S., Schumann, J., Srivastava, J., & Hale, T. S. (2013). A Machine Learning-Based Analysis of Game Data for Attention Deficit Hyperactivity Disorder Assessment. *Games for Health Journal, 2*(5), 291–298. https://doi.org/10.1089/g4h.2013.0058

Huckvale, K., Venkatesh, S., & Christensen, H. (2019). Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *NPJ Digital Medicine, 2*, 88. https://doi.org/10.1038/s41746-019-0166-1

Insel, T. R. (2017). Digital phenotyping: Technology for a new science of behavior. *JAMA, the Journal of the American Medical Association, 318*(13), 1215–1216. https://doi.org/10.1001/jama.2017.11295

Iriarte, Y., Diaz-Orueta, U., Cueto, E., Irazustabarrena, P., Banterla, F., & Climent, G. (2016). AULA-advanced virtual reality tool for the assessment of attention: Normative study in Spain: Normative study in Spain. *Journal of Attention Disorders, 20*(6), 542–568. https://doi.org/10.1177/1087054712465335

Jacobson, N. C., Weingarden, H., & Wilhelm, S. (2019). Using digital phenotyping to accurately detect depression severity. *The Journal of Nervous and Mental Disease, 207*(10), 893–896. https://doi.org/10.1097/NMD.0000000000001042

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences, 374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Japkowicz, N., & Shah, M. (2014). *Evaluating Learning Algorithms: A Classification Perspective* (1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511921803

Johannes Dechant, M., Frommel, J., & Mandryk, R. (2021). Assessing social anxiety through digital biomarkers embedded in a gaming task. *Proceedings of the 2021 CHI conference on human factors in computing systems*. https://doi.org/10.1145/3411764.3445238. Article Article 544.

Ju, U., & Wallraven, C. (2023). Relationship between virtual reality balloon analogue risk task and risky decision-making. *Plos one, 18*(2), Article e0282097.

Kampmann, I. L., Emmelkamp, P. M. G., & Morina, N. (2018). Self-report questionnaires, behavioral assessment tasks, and an implicit behavior measure: Do they predict social anxiety in everyday life? *PeerJ, 6*, Article e5441. https://doi.org/10.7717/peerj.5441

Kilbourne, A. M., Beck, K., Spaeth-Rublee, B., Ramanuj, P., O'Brien, R. W., Tomoyasu, N., et al. (2018). Measuring and improving the quality of mental health care: A global perspective. *World Psychiatry: Official Journal of the World Psychiatric Association, 17*(1), 30–38. https://doi.org/10.1002/wps.20482

Kim, J., Lee, J., Park, E., & Han, J. (2020). A deep learning model for detecting mental illness from user content on social media. *Scientific Reports, 10*(1), Article 11846. https://doi.org/10.1038/s41598-020-68764-y

Kirkham, R., Kooijman, L., Albertella, L., Myles, D., Yücel, M., & Rotaru, K. (2024). Immersive virtual reality-based methods for assessing executive functioning: Systematic review. *JMIR Serious Games, 12*(1), Article e50282. https://doi.org/10.2196/50282

Konovalova, I., Antolin, J. V., Bolderston, H., & Gregory, N. J. (2021). Adults with higher social anxiety show avoidant gaze behaviour in a real-world social setting: A mobile eye tracking study. *PLoS One, 16*(10), Article e0259007. https://doi.org/10.1371/journal.pone.0259007

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America, 110*(15), 5802–5805. https://doi.org/10.1073/pnas.1218772110

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality and Quantity, 47*(4), 2025–2047. https://doi.org/10.1007/s11135-011-9640-9

Kumar, A., & Phookun, H. R. (2016). Barriers in the treatment of psychiatric disorders. *Open Journal of Psychiatry & Allied Sciences, 7*(2), 99. https://doi.org/10.5958/2394-2061.2016.00015.x

Lahiri, U., Bekele, E., Dohrmann, E., Warren, Z., & Sarkar, N. (2013). Design of a virtual reality based adaptive response technology for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society, 21*(1), 55–64. https://doi.org/10.1109/TNSRE.2012.2218618

Lange, B., & Pauli, P. (2019). Social anxiety changes the way we move — a social approach-avoidance task in a virtual reality CAVE system. *PLoS One, 14*(12), 1–19. https://doi.org/10.1371/journal.pone.0226805

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience Questionnaire. In *HCI and usability for education and work, 4th symposium of the workgroup human-computer interaction and usability engineering of the Austrian computer society, USAB 2008, graz, Austria, november 20-21, 2008. Proceedings* (Vol. 5298, pp. 63–76). https://doi.org/10.1007/978-3-540-89350-9_6

Lejuez, C. W., Read, J. P., Kahler, C. W., Richards, J. B., Ramsey, S. E., Stuart, G. L., … Brown, R. A. (2002). Evaluation of a behavioral measure of risk taking: the Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied, 8*(2), 75.

Levy, L., Lambeth, A., Solomon, R., & Gandy, M. (2018). Method in the madness: The design of games as valid and reliable scientific tools. *Proceedings of the 13th international conference on the foundations of digital games*. https://doi.org/10.1145/3235765.3235793. Article Article 9.

Lim, T., Carvalho, M. B., Bellotti, F., Arnab, S., Sara, D. F., Louchart, S., et al. (2015). *The LM-GM framework for serious games analysis*. Pittsburgh: University of Pittsburgh.

Liu, F., Qin, J., Zhang, Z., Tang, K., Hu, J., & Zhou, Y. (2022). Eye tracking emotion evoking interaction based on go/no-go paradigm VR demonstration. *Proceedings of the 18th ACM SIGGRAPH international Conference on virtual-reality Continuum and its Applications in industry. VRCAI '22: The 18th ACM SIGGRAPH international conference on virtual-reality continuum and its applications in industry*. https://doi.org/10.1145/3574131.3574465. Guangzhou China.

Zheng, A. (2015). *Evaluating Machine Learning Models: A Beginner's Guide to Key Concepts and Pitfalls*. O'Reilly Media.

Lowood, H.E. (2012). Living in virtual worlds. Encyclopedia Britannica. Retrieved from https://www.britannica.com/technology/virtual-reality. Accessed September 8, 2024.

Mandryk, R. L., & Birk, M. V. (2019). The potential of game-based digital biomarkers for modeling mental health. *JMIR Mental Health, 6*(4), Article e13485. https://doi.org/10.2196/13485

Martin, D. C. (1990). The mental Status examination. In H. K. Walker, W. D. Hall, & J. W. Hurst (Eds.), *Clinical methods: The history, physical, and laboratory examinations*. Butterworths. https://www.ncbi.nlm.nih.gov/pubmed/21250162.

Martinez-Martin, N., Insel, T. R., Dagum, P., Greely, H. T., & Cho, M. K. (2018). Data mining for health: Staking out the ethical territory of digital phenotyping. *NPJ Digital Medicine, 1*. https://doi.org/10.1038/s41746-018-0075-8

Mendez-Encinas, D., Sujar, A., Bayona, S., & Delgado-Gomez, D. (2023). Attention and impulsivity assessment using virtual reality games. *Scientific Reports, 13*(1), Article 13689. https://doi.org/10.1038/s41598-023-40455-4

Meyer, N., Kerz, M., Folarin, A., Joyce, D. W., Jackson, R., Karr, C., et al. (2018). Capturing rest-activity profiles in schizophrenia using wearable and mobile technologies: Development, implementation, feasibility, and acceptability of a remote monitoring platform. *JMIR MHealth and UHealth, 6*(10), e188. https://doi.org/10.2196/mhealth.8292

Meyerbröker, K., & Morina, N. (2021). The use of virtual reality in assessment and treatment of anxiety and related disorders. *Clinical Psychology & Psychotherapy, 28*(3), 466–476. https://doi.org/10.1002/cpp.2623

Miranda, D., Calderón, M., & Favela, J. (2014). Anxiety detection using wearable monitoring. *Proceedings of the 5th Mexican conference on human-computer interaction*. https://doi.org/10.1145/2676690.2676694

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence-centered design. *ETS Research Report Series, 2003*(1), i–29. https://doi.org/10.1002/j.2333-8504.2003.tb01908.x

Mohr, D. C., Zhang, M., & Schueller, S. M. (2017). Personal sensing: Understanding mental health using ubiquitous sensors and machine learning. *Annual Review of Clinical Psychology, 13*, 23–47. https://doi.org/10.1146/annurev-clinpsy-032816-044949

Montag, C., Sindermann, C., & Baumeister, H. (2020). Digital phenotyping in psychological and medical sciences: A reflection about necessary prerequisites to reduce harm and increase benefits. *Current Opinion in Psychology, 36*, 19–24. https://doi.org/10.1016/j.copsyc.2020.03.013

Morales González, R., Freeman, E., & Georgiou, O. (2020). Levi-loop: A mid-air gesture controlled levitating particle game. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–4. https://doi.org/10.1145/3334480.3383152

Murman, D. L. (2015). The Impact of Age on Cognition. *Seminars in Hearing, 36*(3), 111–121. https://doi.org/10.1055/s-0035-1555115

Neguţ, A., Matu, S.-A., Sava, F. A., & David, D. (2016). Virtual reality measures in neuropsychological assessment: A meta-analytic review. *The Clinical Neuropsychologist, 30*(2), 165–184. https://doi.org/10.1080/13854046.2016.1144793

Ng, K., Sun, J., Hu, J., & Wang, F. (2015). Personalized predictive modeling and risk factor identification using patient similarity. *AMIA Joint Summits on Translational Science, 2015*, 132–136.

Norris, D., Clark, M. S., & Shipley, S. (2016). The mental Status examination. *American Family Physician, 94*(8), 635–641. https://www.ncbi.nlm.nih.gov/pubmed/27929229.

O'Brien, J. T., Gallagher, P., Stow, D., Hammerla, N., Ploetz, T., Firbank, M., et al. (2017). A study of wrist-worn activity measurement as a potential real-world biomarker for late-life depression. *Psychological Medicine, 47*(1), 93–102. https://doi.org/10.1017/S0033291716002166

Ozkanca, Y., Öztürk, M. G., Ekmekci, M. N., Atkins, D. C., Demiroglu, C., & Ghomi, R. H. (2019). Depression screening from voice samples of patients affected by Parkinson's disease. *Digital Biomarkers, 3*(2), 72–82. https://doi.org/10.1159/000500354

Pallavicini, F., & Pepe, A. (2020). Virtual reality games and the role of body involvement in enhancing positive emotions and decreasing anxiety: Within-subjects pilot study. *JMIR Serious Games, 8*(2), Article e15635. https://doi.org/10.2196/15635

Pampouchidou, A., Pediaditis, M., Kazantzaki, E., Sfakianakis, S., Apostolaki, I. A., Argyraki, K., et al. (2020). Automated facial video-based recognition of depression and anxiety symptom severity: Cross-corpus validation. *Machine Vision and Applications, 31*(4), 30. https://doi.org/10.1007/s00138-020-01080-7

Pan, X., & Hamilton, A. F. de C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology, 109*(3), 395–417. https://doi.org/10.1111/bjop.12290

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience, 9*, 660. https://doi.org/10.3389/fnhum.2015.00660

Parsons, T. D., & Barnett, M. D. (2018). Virtual apartment stroop task: Comparison with computerized and traditional stroop tasks. *Journal of Neuroscience Methods, 309*, 35–40. https://doi.org/10.1016/j.jneumeth.2018.08.022

Parsons, T. D., Courtney, C. G., Arizmendi, B., & Dawson, M. (2011). Virtual reality stroop task for neurocognitive assessment. *Studies in Health Technology and Informatics, 163*, 433–439. https://www.ncbi.nlm.nih.gov/pubmed/21335835.

Plotnik, M., Ben-Gal, O., Doniger, G. M., Gottlieb, A., Bahat, Y., Cohen, M., et al. (2021). Multimodal immersive trail making-virtual reality paradigm to study cognitive-motor interactions. *Journal of NeuroEngineering and Rehabilitation, 18*(1), 82. https://doi.org/10.1186/s12984-021-00849-9

Porffy, L. A., Mehta, M. A., Patchitt, J., Boussebaa, C., Brett, J., D'Oliveira, T., et al. (2022). A novel virtual reality assessment of functional cognition: Validation study. *Journal of Medical Internet Research, 24*(1), Article e27641. https://doi.org/10.2196/27641

Rebenitsch, L., & Owen, C. (2016). Review on cybersickness in applications and visual displays. *Virtual Reality, 20*(2), 101–125. https://doi.org/10.1007/s10055-016-0285-9

Rinck, M., Rörtgen, T., Lange, W.-G., Dotsch, R., Wigboldus, D. H. J., & Becker, E. S. (2010). Social anxiety predicts avoidance behaviour in virtual encounters. *Cognition and Emotion, 24*(7), 1269–1276. https://doi.org/10.1080/02699930903309268

Riva, G., Mantovani, F., Capideville, C. S., Preziosa, A., Morganti, F., Villani, D., et al. (2007). Affective interactions using virtual reality: The link between presence and emotions. *CyberPsychology and Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 10*(1), 45–56. https://doi.org/10.1089/cpb.2006.9993

Rizzo, A. A., Buckwalter, J. G., Bowerly, T., van der Zaag, C., Humphrey, L., Neumann, U., et al. (2000). The virtual classroom: A virtual reality environment for the assessment and rehabilitation of attention deficits. *CyberPsychology and Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 3*(3), 483–499. https://doi.org/10.1089/10949310050078940

Rizzo, A. S., & Koenig, S. T. (2017). Is clinical virtual reality ready for primetime? *Neuropsychology, 31*(8), 877–899. https://doi.org/10.1037/neu0000405

Rodríguez, C., Areces, D., García, T., Cueli, M., & González-Castro, P. (2018). Comparison between two continuous performance tests for identifying ADHD: Traditional vs. virtual reality. *International Journal of Clinical and Health Psychology: IJCHP, 18*(3), 254–263. https://doi.org/10.1016/j.ijchp.2018.06.003

Rus-Calafell, M., Garety, P., Sason, E., Craig, T. J. K., & Valmaggia, L. R. (2018). Virtual reality in the assessment and treatment of psychosis: A systematic review of its utility, acceptability and effectiveness. *Psychological Medicine, 48*(3), 362–391. https://doi.org/10.1017/S0033291717001945

Saeb, S., Lattie, E. G., Kording, K. P., & Mohr, D. C. (2017). Mobile phone detection of semantic location and its relationship to depression and anxiety. *JMIR MHealth and UHealth, 5*(8), Article e112. https://doi.org/10.2196/mhealth.7297

Saeb, S., Zhang, M., Karr, C. J., Schueller, S. M., Corden, M. E., Kording, K. P., et al. (2015). Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *Journal of Medical Internet Research, 17*(7), Article e175. https://doi.org/10.2196/jmir.4273

Šalkevicius, J., Damaševičius, R., Maskeliunas, R., & Laukienė, I. (2019). Anxiety Level Recognition for Virtual Reality Therapy System Using Physiological Signals. *Electronics, 8*(9), 1039. https://doi.org/10.3390/electronics8091039

Soong, J. (2021). *Depression traps: Social withdrawal, rumination, and more*. WebMD. https://www.webmd.com/depression/features/depression-traps-and-pitfalls.

Spieker, E. A., Astur, R. S., West, J. T., Griego, J. A., & Rowland, L. M. (2012). Spatial memory deficits in a virtual reality eight-arm radial maze in schizophrenia. *Schizophrenia Research, 135*(1–3), 84–89. https://doi.org/10.1016/j.schres.2011.11.014

Spinazze, P., Rykov, Y., Bottle, A., & Car, J. (2019). Digital phenotyping for assessment and prediction of mental health outcomes: A scoping review protocol. *BMJ Open, 9*(12), Article e032255. https://doi.org/10.1136/bmjopen-2019-032255

Spitzer, R. L., Williams, J. B., Kroenke, K., Linzer, M., deGruy, F. V., Hahn, S. R., 3rd, et al. (1994). Utility of a new procedure for diagnosing mental disorders in primary care. The PRIME-MD 1000 study. *JAMA, the Journal of the American Medical Association, 272*(22), 1749–1756. https://www.ncbi.nlm.nih.gov/pubmed/7966923.

Tabbaa, L., Searle, R., Bafti, S. M., Hossain, M. M., Intarasisrisawat, J., Glancy, M., et al. (2021). Vreed: Virtual Reality emotion recognition dataset using eye tracking & physiological measures. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 5*(4), 1–20. https://doi.org/10.1145/3495002

Torous, J., Kiang, M. V., Lorme, J., & Onnela, J.-P. (2016). New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research. *JMIR Mental Health, 3*(2), e16. https://doi.org/10.2196/mental.5165

van der Maaten, L., & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research: JMLR, 9*(86), 2579–2605.

Verschueren, S., Buffel, C., & Vander Stichele, G. (2019). Developing theory-driven, evidence-based serious games for health: Framework based on research community insights. *JMIR Serious Games, 7*(2), Article e11565. https://doi.org/10.2196/11565

Voinescu, A., Petrini, K., Stanton Fraser, D., Lazarovicz, R.-A., Papavă, I., Fodor, L. A., et al. (2021). The effectiveness of a virtual reality attention task to predict depression and anxiety in comparison with current clinical measures. *Virtual Reality*. https://doi.org/10.1007/s10055-021-00520-7

Wainberg, M. L., Scorza, P., Shultz, J. M., Helpman, L., Mootz, J. J., Johnson, K. A., et al. (2017). Challenges and opportunities in global mental health: A research-to-practice perspective. *Current Psychiatry Reports, 19*(5), 28. https://doi.org/10.1007/s11920-017-0780-z

Wang, R., Wang, W., daSilva, A., Huckins, J. F., Kelley, W. M., Heatherton, T. F., et al. (2018). Tracking depression dynamics in college students using mobile phone and wearable sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., 2*(1), 1–26. https://doi.org/10.1145/3191775

Wongkoblap, A., Vadillo, M. A., & Curcin, V. (2021). Deep learning with anaphora resolution for the detection of tweeters with depression: Algorithm development and validation study. *JMIR Mental Health, 8*(8), Article e19824. https://doi.org/10.2196/19824

World Health Organization. (2019). International classification of diseases. Eleventh Revision (ICD-11) https://icd.who.int/en.

World Health Organization. (2022). Mental disorders. https://www.who.int/news-room/fact-sheets/detail/mental-disorders.

World Health Organization. (2023). Depressive disorder (depression). https://www.who.int/news-room/fact-sheets/detail/depression.

Xu, W., Liang, H.-N., Yu, K., Wen, S., Baghaei, N., & Tu, H. (2023). Acceptance of virtual reality exergames among Chinese older adults. *International Journal of Human-Computer Interaction, 39*(5), 1134–1148. https://doi.org/10.1080/10447318.2022.2098559

Yaremych, H. E., & Persky, S. (2019). Tracing physical behavior in virtual reality: A narrative review of applications to social psychology. *Journal of Experimental Social Psychology, 85*. https://doi.org/10.1016/j.jesp.2019.103845

Yeh, S.-C., Lin, S.-Y., Wu, E. H.-K., Zhang, K.-F., Xiu, X., Rizzo, A., et al. (2020). A virtual-reality system integrated with neuro-behavior sensing for attention-deficit/hyperactivity disorder intelligent assessment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society, 28*(9), 1899–1907. https://doi.org/10.1109/TNSRE.2020.3004545

Yen, H.-Y., & Chiu, H.-L. (2021). Virtual reality exergames for improving older adults' cognition and depression: A systematic review and meta-analysis of randomized control trials. *Journal of the American Medical Directors Association, 22*(5), 995–1002. https://doi.org/10.1016/j.jamda.2021.03.009

Zanon, M., Novembre, G., Zangrando, N., Chittaro, L., & Silani, G. (2014). Brain activity and prosocial behavior in a simulated life-threatening situation. *NeuroImage, 98*, 134–146. https://doi.org/10.1016/j.neuroimage.2014.04.053