CAEVR: Biosignals-Driven Context-Aware Empathy in Virtual Reality

Kunal Gupta, Yuewei Zhang, Tamil Selvan Gunasekaran, Nanditha Krishna, Yun Suen Pai, Mark Billinghurst

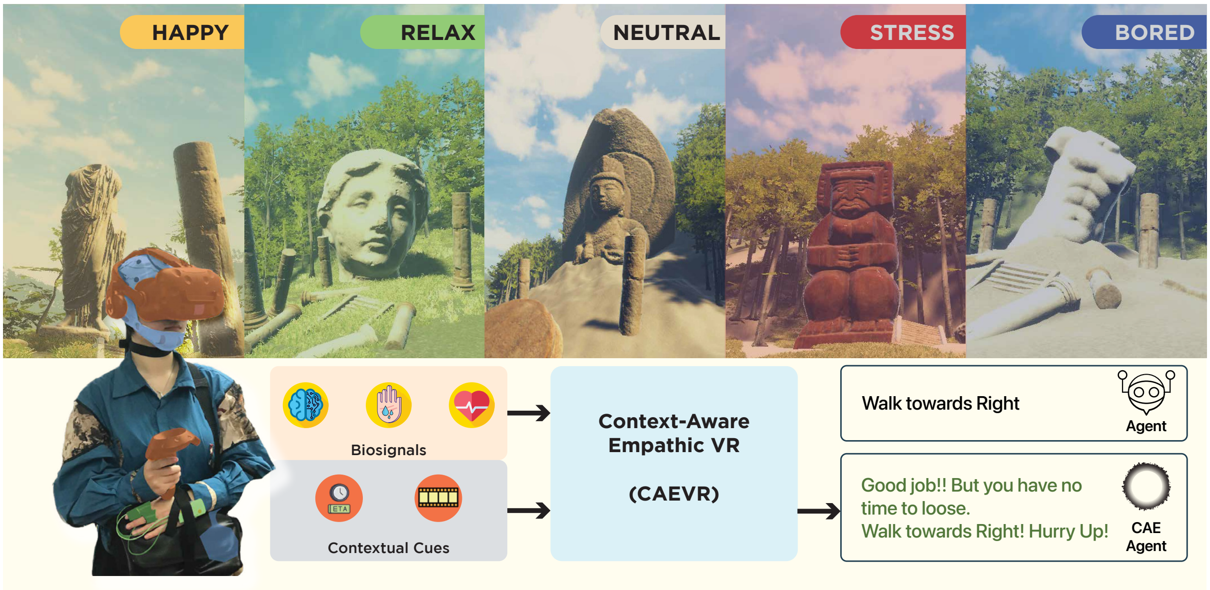

Fig. 1: CAEVR: integrating biosignals and context to adapt VR environment and agent feedback based on user emotions.

Abstract— There is little research on how Virtual Reality (VR) applications can identify and respond meaningfully to users' emotional changes. In this paper, we investigate the impact of Context-Aware Empathic VR (CAEVR) on the emotional and cognitive aspects of user experience in VR. We developed a real-time emotion prediction model using electroencephalography (EEG), electrodermal activity (EDA), and heart rate variability (HRV) and used this in personalized and generalized models for emotion recognition. We then explored the application of this model in a context-aware empathic (CAE) virtual agent and an emotion-adaptive (EA) VR environment. We found a significant increase in positive emotions, cognitive load, and empathy toward the CAE agent, suggesting the potential of CAEVR environments to refine user-agent interactions. We identify lessons learned from this study and directions for future work.

Index Terms—empathy, VR, metaverse, physiology, emotion, context-aware, virtual agents

1 Introduction

This paper describes a Context-Aware Empathic Virtual Reality (CAEVR) system that recognizes emotion from biosignal data to create enhanced immersive experiences. The significance of empathy in shaping human interactions, extending even to AI entities, is well-documented [37,61]. However, computers' limited emotional responsiveness presents a challenge in replicating the depth of human empathy. Empathic Computing (EmpComp) seeks to bridge this gap, leveraging AI and biosensing to create more human-centric interactions [14,65].

VR's sensory-rich environment and virtual agents offer enhanced user-agent connections, particularly relevant in sectors like healthcare and education. This facilitates research in affective computing [55], cognitive psychology [32], context-aware systems [58], and multimodal adaptation [70] in VR. A comprehensive exploration of how combining real-time physiology-based emotion recognition, emotion-adaptive VR, and context-aware empathic responses could improve users' VR experience.

This motivated us to ask how context-aware empathic interactions (CAEIxs) in VR could enhance the emotional and cognitive aspects of the experience. We are interested in the research questions (RQs):

RQ1 How can physiological signals be used to predict emotions and facilitate CAEIxs in VR environments?

RQ2 What are the effects of CAEIxs on elicited emotions, cognitive load, and empathy towards a virtual agent in VR?

RQ3 How can the impact of CAEIxs on users' emotional and cognitive load during VR experiences be evaluated?

To answer these RQs, we first developed a four-class machine learning (ML) model to recognize emotions. This model used biosignals such as electroencephalography (EEG, brainwaves), electrodermal activity (EDA: skin sweat response), and photoplethysmography (PPG: cardiac activity). Next, we used this model in the CAEVR system and evaluated this system using a within-subject user study with CAEIxs and Emotion-Adaptive VR as independent variables. The contributions of this work are: **1. Empirical**: We identified correlations between valence ratings and Emotion-Adaptive (EA) and Context-Aware Empathy (CAE) systems, informing how VR can adapt to user emotions; **2. Methodological**: We designed a system for real-time emotion recognition in VR using EEG, EDA, and HRV biosignals; **3. Methodological**:

- *Kunal Gupta, Yuewei Zhang, Tamil Selvan Gunasekaran, Yun Suen Pai, and Mark Billinghurst are with the University of Auckland. E-mail: kunal.gupta@auckland.ac.nz*
- *Nanditha Krishna is with Amrita Vishwa Vidyapeetham*

Manuscript received 4 October 2023; revised 17 January 2024; accepted 24 January 2024. Date of publication 4 March 2024; date of current version 15 April 2024. This article has supplementary downloadable material available at https://doi.org/10.1109/TVCG.2024.3372130, provided by the authors. Digital Object Identifier no. 10.1109/TVCG.2024.3372130

We created a VR application demonstrating use of CAEIxs for enhanced user engagement and integrating empathic virtual agents into this VR application; and **4. Technological**: We developed an integrated VR system that merges real-time emotion detection via EEG, EDA, and HRV with CAEIxs, streamlining the incorporation of empathic virtual agents to enhance user engagement.

2 Related Work

This section overviews empathy's theoretical and computational background, emotional responses, and empathy in virtual agents, emotion-adaptive virtual reality, and context-aware virtual reality.

2.1 Theoretical and Computational Empathy

Defining empathy is still elusive [19]. Researchers such as Hoffman [38], Davis [22], and Preston [69] have delved into multidimensional approaches, considering both cognitive and affective components of empathy. Preston introduced the Perception/Action Model (PAM) of empathy, emphasizing shared emotional experiences, and further explored the role of "cognitive empathy" in situations where subjects imagine the object's state [68]. Building on these principles, Rodrigues et al. [73] developed a computational model that aligns with cognitive theories of emotion, focusing on appraisals that assess the impact of environmental events on individual emotions and incorporate advanced cognitive processes such as the theory of mind and self-projection [61].

Two major empathic response strategies include emotion contagion [35,36] and emotion regulation [29,84]. Emotion contagion strengthens relationships and enhances shared experiences but can also be a source of stress and burnout. On the other hand, emotion regulation helps maintain emotional balance when empathizing with those experiencing negative emotions [52].

Our system intricately integrates these concepts of empathic appraisal and response along with the strategies of contagion and regulation in our CAEVR system as described in section 3.

2.2 Emotional Responses and Empathy in Virtual Agents

In exploring the interplay between emotional responses and human-agent interactions, a key focus has been on how agent tone and dialogue adaptations influence user experience. Affective states have been expressed through diverse means, emphasizing agent nonverbal and verbal cues [61]. The potential for misinterpretation of virtual agents' facial expressions was identified, emphasizing the need for accurate emotional representation. Notably, verbal acknowledgments of a user's emotions outweighed the influence of facial expressions alone in generating perceived empathy [23].

Voice tone, an intrinsic aspect of communication, was recognized as a conveyor of diverse emotions and intentions. Emphasis on the critical nature of nonverbal communication was evident, especially with findings suggesting a dominant role of such cues, accounting for approximately 60-65% of interpretations in human communication [44]. The alignment of nonverbal vocal cues with user emotions is pivotal for establishing empathy in voice-only communication [11]. Adjustments in affective voice tone during sensitive conversations enhanced perceived empathy [74] and during learning activity, reduced cognitive load [50].

However, there is a gap in applying these principles in VR settings, particularly in aligning TTS tones with specific emotional states. This gap underlines the need for more research into how empathic virtual agents can influence cognitive load. Cognitive Load Theory suggests that an intuitive and user-friendly design of virtual agent interactions could reduce cognitive load by streamlining communication and reducing unnecessary cognitive effort [13]. The "Empathic Companion" exemplifies the use of real-time emotional feedback in virtual agents [67]. This suggests a promising direction for VR systems that adapt their interactions based on real-time emotional states, potentially impacting cognitive load. However, direct empirical evidence linking empathic, context-aware VR systems to cognitive load reduction remains limited.

2.3 Emotion-Adaptive Virtual Reality

Integrating emotion-adaptiveness in VR, focusing on dynamic responses to user emotions as biofeedback, is a key area in enhancing user experience. Using technologies to interpret emotions and adapt the VR environment, facilitates awareness and voluntary modification of bodily reactions. Studies like those by Bouchard et al. [7] have demonstrated the effectiveness of visual and auditory biofeedback in stress reduction using physiological measures such as salivary cortisol and heart rate.

Badia et al. [40] employed physiological signals to adjust VR content. Their system used biofeedback signals to change visual elements like images, and auditory elements like music, to represent different emotions. However, their method did not recognize emotional states or adapt to the environment in real-time. Liang et al. [49] further developed this concept by introducing specific characters in a VR environment to influence emotions based on EEG features related to relaxation and attentiveness. Yet, their study primarily focused on a limited range of emotional states and did not integrate psychological principles beyond biofeedback.

Research has also linked specific colors to emotional states [53], with studies adapting VR environment colors based on real-time biofeedback for therapeutic purposes [42,57]. For instance, green has been associated with relaxation and calmness [45], and yellow with empathy and joy [39]. A comprehensive review of affective visualization in VR indicates that combining audio-visual cues like color and sound can effectively represent users' emotions [64]. These findings underpin the hypothesis that changing environmental colors as biofeedback could aid emotion regulation and improve emotional well-being.

2.4 Context-Aware Virtual Reality

Within ubiquitous computing, various context definitions have emerged [1, 3, 9, 75, 76]. Dey et al.'s [26] state, "Context is any information characterizing the situation of an entity, which might be a person, place, or object relevant to the interaction between a user and an application." Given this, most VR systems are context-aware, adapting to player movements and tracking interactions. While investigations into context-aware systems cover AR [30], mobile [31], and IoT [62], VR remains underexplored.

Lee et al. [48] introduced vr-UCAM, an architecture facilitating context-aware VR with user-specific attributes to enhance realism in a virtual heritage tour. Later, Dennemont et al. [24] devised a semantic context-aware engine for enhancing 3D interaction within Virtual Environments, employing concept graphs to inform decisions based on historical events and new information. In a First Aid VR Training scenario, Yigitbas et al. [88] incorporated contextual information, including user, platform, and environment. However, an extensive exploration of incorporating emotional states for empathic interaction within VR remains absent in these approaches.

2.5 Hypotheses

We've developed two main hypotheses based on the research questions and reviewed the literature. The first, **[H1]**, is that emotion-adaptive (EA) and context-aware empathic (CAE) interactions will enhance emotional well-being in VR by improving users' emotional states **[H1a]**, their sense of presence **[H1b]**, and raising their empathy levels **[H1c]**. The second hypothesis, **[H2]**, suggests that these interactions will significantly affect cognitive aspects in VR, influencing users' cognitive load **[H2a]**, and flow state **[H2b]**. To evaluate these hypotheses, we created the Context-Aware Empathic VR system described next.

3 CAEVR: Context-Aware Empathic VR

The CAEVR system consists of five modules: (1) 'Empathic Virtual Experience' (EVE), which uses VR audio and visual cues to stimulate the user's psycho-physiological responses, (2) 'Data Collection', enabling the system to acquire biosignals from the user, (3) 'Empathic Appraisal', enabling recognition of the user's emotion from their biosignals and contextual cues, and (4) 'Empathic Response', generate CAEIxs. In this section, we discuss these modules in turn.

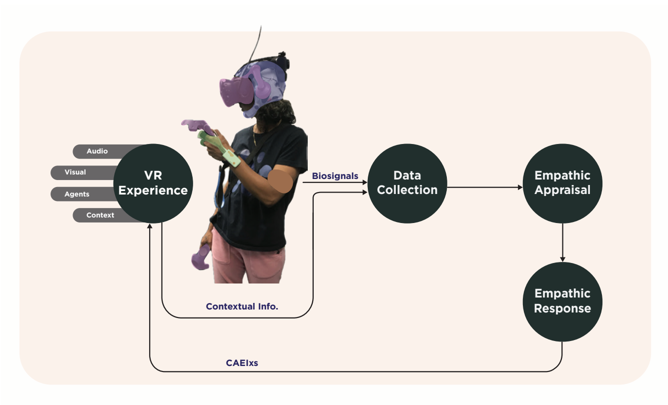

Fig. 2: The CAEVR System

3.1 Empathic Virtual Experience

We focused on designing an experience that could induce flow, a mental state characterized by complete immersion in an activity [18]. Flow has been shown to positively impact well-being, such as feelings of worth, mastery, achievement, and satisfaction [82]. To enable the flow state, we designed the experience to be challenging but with achievable tasks, to be immersive, and to provoke users with a sense of control and agency, following the guidelines of [17].

Previously, empathic virtual agents have been explored extensively in improving engagement, immersion, presence, and flow state [81]. Therefore, the VR experience was designed to naturally accommodate virtual agents' interactions that could exploit the predicted empathic emotion from the empathic appraisal process.

We designed a VR application in the context of a virtual photography competition, where users aim to take the best photographs of two monuments on an island. This creates a clear goal and motivation for users to explore the island, search for the monuments, and take photos. Users had to search without a map over a 7-minute journey, and they were encouraged to take as many photos as they wanted, but they could only save eight pictures. Limiting the number of photos creates a sense of urgency and encourages users to make careful decisions about what to capture. In addition, they have a virtual companion, shown as an orb-shaped character, which communicates in natural language and provides on-demand assistance to help them with navigation, timing, and photo-taking. This helps to increase engagement and interaction.

3.2 Data Collection

This module was designed to capture biosignals and user and activity-related contextual information. We used the HTC Vive Pro VR HMD and controllers, interfaced via SteamVR, to record user head movement and activity performance at 90 Hz. This data collection included specific variables such as Estimated Time of Arrival (ETA) and the count of photographs taken, allowing for contextual understanding as discussed in section 3.3.2. To gain comprehensive insights into the physiological and emotional responses of the users, the OpenBCI EEG Cap and the Shimmer GSR+ sensor were used, streaming EEG, EDA, and PPG signals at sampling frequencies of 125 Hz and 128 Hz, respectively. The Biosignal Manager subsection describes their integration within the CAEVR system.

3.3 Empathic Appraisal

This module understands the physiological and situational emotions to facilitate empathetic interactions in VR. Three methods were implemented: 1) BioEmoVR, a generalized emotion recognition model; 2) Self-Projected Appraisal to assess context and user emotional state; and 3) Empathic Emotion Features, synthesizing contextual and emotional information. These are described in more detail next.

3.3.1 BioEmoVR

BioEmoVR is a generalized emotion recognition model for VR developed using a dataset of EEG, EDA, and HRV biosignals collected from 9 participants following the method mentioned in our previous work [33]. The Dataset underwent rigorous pre-processing methods and feature engineering to improve BioEmoVR's performance.

Data Preprocessing: The raw data from EEG, PPG, and EDA signals was cleaned using eyeballing, bandpass filtering, and moving average filters. For EEG data, preprocessing involved noise reduction, artifact removal, and bandpass filtering ($1-50$ Hz). EDA data underwent preprocessing for noise elimination, outlier handling, bandpass ($0.05-5$ Hz), and baseline correction. HRV data preprocessing encompassed moving average filter, RR interval extraction, outlier detection, and bandpass filter ($0.04-1$ Hz). The data was then segmented into 30-second epochs with 50% overlap.

Feature Extraction: The Neurokit2 package [54] extracted 219 total features from EEG (194), HRV (18), and EDA (7), including time-domain, frequency-domain, and statistical features. Normalization techniques, such as z-score scaling, were applied to ensure consistent data scaling. The SAM questionnaire responses were ground truth for labeling emotional states into four classes.

Feature Scaling, Selection, and Balancing: Features were standardized to zero mean and unit variance using the standard scaler. The Yeo-Johnson transformation normalized feature distribution. Redundant features were reduced using backward feature elimination, and class distribution was balanced with SMOTE.

Machine Learning Algorithms, Tuning, and Validation: The gradient boosting classifier was trained to predict four emotional states: 'happy' for Positive Valence - Positive Arousal (PVPA), 'stress' for Negative Valence - Positive Arousal (NVPA), 'bored' for Negative Valence - Negative Arousal (NVNA), and 'relaxed' for Positive Valence - Negative Arousal (PVNA). These classes were specifically chosen to align with the emotional experiences anticipated in the VR photo-walk experience designed for the study. A generalized modeling approach was applied, with an 80/20 split for training and testing/ validation. This model achieved a 10-fold cross-validation accuracy of 91.67%, indicating high reliability and effectiveness in emotion classification.

3.3.2 Self-Projected Appraisal (SPA)

The Self-Projected Appraisal (SPA) technique is designed for VR systems to empathetically understand users by considering their perspectives, needs, desires, and motivations. SPA's key function is eliciting situational emotions (SE) based on user activities and the context of the VR environment. For instance, in our VR photography scenario, SPA assesses activities like navigation, time spent, and photo count. For example, a stress SE is triggered if a user does not reach a monument in time. Conversely, a sad SE is triggered if not enough photos are taken. If neither scenario applies, the user feels happiness. In navigation, the SPA system aims to motivate users to explore as many monuments as possible. This increases their chances of finding the best photo opportunity. The system gauges the remaining journey time (RJT), sets it against the estimated time (ET) to monuments, and uses algorithm 1 (Appendix) to decide the SE. For photography, SPA encourages users to capture multiple angles of each monument. This enhances their chances of having at least one good photo. Based on the remaining photos (RP) and monuments (ML) left to visit, algorithm 2 (Appendix) determines the SE.

While the SPA method in this research is specialized for virtual photography, the underlying principle of SPA — deriving situational emotions based on user actions and context — has broader potential applicability. For instance, in VR relaxation applications, SPA might evaluate interactions with serene environments or completion of relaxation exercises to infer emotions such as calmness or stress. For training scenarios, SPA's focus could shift to assessing task completions and response times, helping to predict emotions like satisfaction or frustration. Similarly, in metaverse applications, SPA could analyze social interactions and achievements within the virtual world, providing insights into emotions like happiness or disappointment.

Similarly, for remote communication applications, SPA could evaluate the quality of virtual interactions, task collaboration, and overall engagement within virtual meeting spaces. This would be key in inferring teamwork-related emotions, such as feelings of collaboration or isolation. In healthcare and rehabilitation VR, SPA could monitor patient interactions with exercises or healthcare simulations, offering insights into emotions tied to recovery, motivation, or discomfort during rehabilitation processes. Lastly, in VR experiences focusing on cultural and historical preservation, SPA could analyze user exploration patterns and engagement with exhibits or virtual tours, predicting emotions such as awe, nostalgia, or cultural appreciation. This wide-ranging applicability of SPA demonstrates its potential as a versatile tool in enhancing user experiences across various VR applications.

3.3.3 Empathic Emotion Features

The Empathic Emotion Features (EEFs) in our research are designed to provide emotional and contextual information to adapt interactions within a VR environment. These EEFs consist of five categories: Empathic Emotion (EE), Comfort and Reassurance (CR), Motivation and Support (MS), Positive Reinforcement and Encouragement (PRE), and Urgency and Pressure (UP). Each category uniquely tailors the VR experience to the user's emotional and situational needs.

Empathic Emotions (EE): The EE category is central to the EEFs, representing the user's primary emotional state. For this study, the EE category is selected based on the valence of physiological emotion (PE) and situational emotion (SE), focusing on negative valence. Negative emotions often indicate distress, boredom, discomfort, and dissatisfaction. The system targets states where empathic responses are most needed by focusing on these. This aligns with studies showing that negative emotions in social interactions benefit from more empathic responses [78]. While positive emotions are generally beneficial, they don't always necessitate an empathic response and can sometimes lead to decreased empathic performance. This suggests the importance of carefully interpreting positive emotions in empathic systems [25]. Based on this rationale, EE selection criteria were determined as 1) When both PE and SE are Positive: EE is selected as positive, aligning with the user's overall positive emotional state; 2) When both PE and SE are Negative: EE is negative, reflecting a compounded need for empathic support; 3) When PE in Negative, and SE Positive: EE is negative, prioritizing physiological indicators of distress over situational positivity; and 4) When PE Positive, and SE Negative: EE is negative, focusing on situational challenges that require attention.

Comfort and Reassurance (CR): CR focuses on providing emotional support and validation, which is important for alleviating distress in users [16]. Enhancing the user experience in scenarios with prevalent negative emotions is crucial. **Motivation and Support (MS):** MS encourages users to pursue their goals within the VR environment. Empathic responses that include guidance and support can significantly boost user motivation and engagement [59]. **Positive Reinforcement and Encouragement (PRE):** PRE aims to bolster self-esteem and confidence by acknowledging the user's achievements and positive qualities [80]. This category is important for maintaining user engagement and satisfaction, particularly when users successfully overcome challenges or reach milestones. **Urgency and Pressure (UP):** The UP category is designed to create a sense of urgency and prompt immediate action when necessary. This can effectively motivate quick decision-making and responses in time-sensitive situations [83].

Implementation in Interaction Design: The EEF categories are implemented as Boolean variables in the interaction design, activated ('1') or deactivated ('0') based on the PE and SE combination. This theoretical framework activates CR when physiological stress is detected, indicating that the user might benefit from emotional comfort and reassurance. However, it is not activated in certain stress situations (like when PE is relaxed and SE is stressed) where the user's relaxed state suggests that they are managing the stress without needing additional comfort. MS is activated when there's an opportunity for positive engagement or goal achievement, even in negative emotions like stress or boredom. PRE is generally activated in positive or mixed emotional states to reinforce positive aspects. It's used in scenarios where at least one of the emotions (PE or SE) is positive, such as "Happy & Stress" or "Relax & Stress", suggesting that the user is in a state where positive reinforcement can be beneficial. The UP category is selectively activated to motivate quick decision-making or action in specific contexts, and deactivated when PE and SE are stressed, as adding pressure in high-stress situations could be counterproductive. Instead, the system opts for less pressurized empathic responses in these scenarios, focusing on managing and alleviating stress rather than exacerbating it. This approach is informed by stress management principles, where reducing additional pressure is crucial in intense stress conditions. This approach ensures the VR experience dynamically aligns with the user's emotional and situational state. Table **??** illustrates these combinations and their impact on EEF selection, demonstrating a responsive and contextually aware empathic interaction system.

3.4 Empathic Response

This module provides guidelines about the selected interaction modality. This includes using various modalities to express affective states, including colors and lighting, the virtual agent's nonverbal cues, such as tone, and verbal cues, such as dialogues. The rest of this section discusses aspects of this in more detail.

3.4.1 Emotion-Adaptive Responses

Environmental colors were changed in response to the user's emotional state, to enhance understanding and regulation of the users' emotions [51]. Based on the insights on colour-emotion associations [5], we assign PVPA (Happy) with Yellow, NVPA (Stress) with Red, NVNA (Sad) with Blue, and PVNA (Relax) with Green. The goal was to increase users' awareness of their emotional states and empower them to regulate their emotions to achieve a more positive emotional experience.

3.4.2 Empathic Tone

The VR empathic virtual agent needed to provide speech feedback with an empathic tone. No previous research has indicated which text-to-speech (TTS) tone style would appropriately correspond to which emotional state. Addressing a research gap concerning the correlation between TTS tone styles and distinct emotional states, a pilot study was conducted with six participants using the Empathic Dialogues dataset [71]. Participants evaluated the perceived empathy of the TTS on a visual analogue scale. The results identified the most empathic tones as 'excited' for PVPA (Happy), 'empathetic' for NVPA (Stress), 'hopeful' for NVNA (Sad), and 'whispering' for PVNA (Relaxed).

3.4.3 Context-Aware Empathic Assistance

An empathic dialogue-capable virtual agent was designed to deliver emotionally aligned responses tailored to user-specific contexts. A Belief-Desire-Intention (BDI) framework was integrated to achieve this harmonization [8]. Within this framework, Beliefs were informed by users' emotions and situational contexts. Desires focused on dialogues that resonate with users' emotional states. Intentions outlined specific actions, prompting users to explore more monuments or capture various photographic perspectives.

A rule-based approach was employed, using predefined statements to ensure the generated Empathic Dialogue aligned seamlessly with the Empathic Emotion Features (EEFs) and the given context [15]. With the guidance of the EEFs presented in table **??**, dialogues were tailored for specific scenarios. For instance, when a user engages in a photo-taking activity displaying happiness (EE = Happy), the system, informed by the EEFs, might suggest an Empathic Dialogue such as *"Splendid shot!"* (PRE=1), complemented by relevant Assistance like *"You have 7 photos left. Would you like to save this photo?"*. Refer to the Empathic Dialogues List in the appendix for a comprehensive overview of dialogues.

4 User Study

4.1 Experimental Design

We conducted a 2x2 within-subject study to evaluate CAEVR, focusing on Emotion-Adaptive (EA) and Context-Aware Empathic Interactions

(CAEIxs). The conditions in Table 1 include NoEA, EA, NoCAE, and CAE. In the EA setup, a color filter adjusted to the user's real-time physiological emotion was introduced. The virtual companion modified its speech tone for CAE, employing empathic dialogues infused with contextual information. In the NoEA-NoCAE baseline, the environment remained without a color filter, and the virtual companion used a standard 'customer service' tone without empathic dialogue.

Table 1: CAEVR: Experimental conditions

	No CAE	CAE
No EA	A: NoEA-NoCAE	C: NoEA-CAE
EA	B: EA-NoCAE	D: EA-CAE

To mitigate potential ordering effects, we randomized the condition order for each participant using a Latin Square design. Our evaluation of the independent variables combined subjective, physiological, and behavioral measures, addressing emotions, presence, empathy, flow state, and cognitive load.

Participants began with the Positive and Negative Affect Schedule (PANAS) [86] to establish a balanced affective state. Subsequent measures included the Self-Assessment Manikin (SAM) [47] for emotional perception, the IGroup Presence Questionnaire (IPQ) [77] for presence, and select dimensions from the Game Experience Questionnaire (GEQ) [41] to evaluate in-game emotions and empathy. We deployed the Flow Short State Questionnaire (FSSQ) [72] for flow state assessment and the NASA-TLX [34] captured workload.

Our study collected EEG signals using a 16-channel OpenBCI EEG Cap with electrode placement [43], targeting the frontal, prefrontal cortex (FP1, FP2, F3, F4, F7, and F8), and occipital (O1 and O2) areas in response to emotional stimuli [63] (see Figure 3). Additionally, we recorded EDA and PPG signals using a Shimmer GSR+ sensor positioned on the non-dominant wrist and electrodes on the index and middle fingers, sampled at 128 Hz. We transmitted the physiological data to the Unity application via the LSL protocol [1], using the OpenBCI GUI configured with a 50Hz notch filter for the EEG cap, and a custom Java application for the Shimmer.

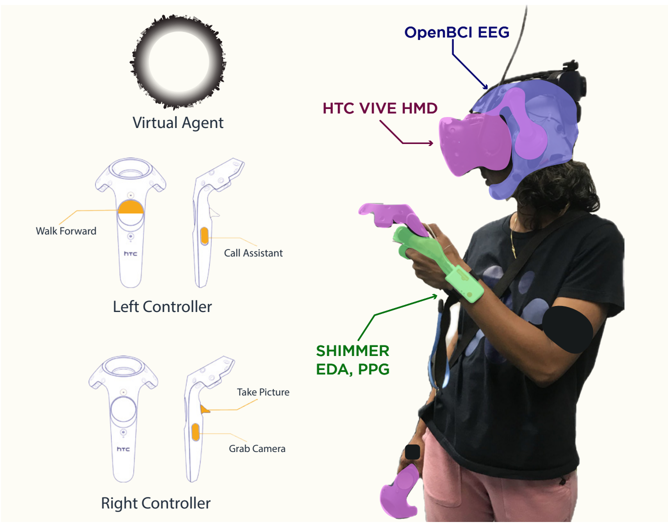

Fig. 3: System Setup on the participant

4.2 System Design

Our system is designed to create an immersive VR environment for a photography contest scenario. We have developed it using Unity3D version 2020.3.28f1 and Python 3.7, combining various components to deliver a seamless and engaging user experience while considering emotional states, providing assistance, and managing data effectively. This section provides an architectural overview (figure 10 in the appendix) and describes the information flow between key components.

4.2.1 VR Environment

We created an immersive forest terrain on a computer powered by an Intel Core i7 8700 CPU and an Nvidia RTX 2070 GPU. Within this environment, we strategically placed eight monuments to serve as subjects for photography, allocating two per journey. We designated four starting points for the journeys. We set the VR experience on a sunny day and enriched it with the ambient wind and bird sounds.

4.2.2 Activity

The VR application focuses on two main activities: Navigation and Photo Capturing. Navigation uses an invisible node map to guide users to the next monument using audible instructions, provided in four primary directions, such as *"Walk Straight"*, *"Turn back and walk straight"*, *"Walk towards Left"*, and *"Walk towards Right"*. When the user requested directions, the system computed the direction concerning the user's head direction. For the photo-capturing activity, we integrated a DSLR camera model. We established a secondary Unity 'Main Camera' at the model's lens to simulate a real photography experience, with camera flash and click sound effects. Users could capture, display, save, and discard photos through a developed photo management script.

Users navigate using head movements and teleportation, activated by the Left controller's touchpad and 'Grab' button. The camera becomes available when the user holds the right controller's 'Grab' button, allowing users to direct it towards a point of interest and capture images using the trigger. After capturing, a photo review window appears, showing the status of the remaining photos and save options (see figure 4).

Fig. 4: Camera View (Left), photo reviewing in VR (Right)

4.2.3 Virtual Agent

In this study, we developed a virtual agent to be a companion to the user. The agent, represented as a colorful floating orb, assists in navigation, photo, and time management. It appears upon user requests and communicates through natural languages. The interaction with the agent is facilitated through a nudge feature, allowing users to seek information or assistance on tasks like inquiring about nearby monuments or managing photos. The agent supports users in saving or discarding photos and provides timely reminders at critical intervals, such as half-time and the last minute, to ensure efficient time and photo management.

4.2.4 Biosignal Manager

The Biosignal Manager system collects and streams biosignals during the VR experience. This system efficiently manages data flow by leveraging the LabStreaming Layer (LSL) protocol and the LSL4Unity API. It gathers diverse biosignal data, including a 16-channel EEG stream from the OpenBCI, EDA, and PPG signals from the Shimmer3 sensor. The EEG was sampled at 125 Hz, while the EDA and PPG were recorded at 128 Hz. We employed event markers within the Unity environment to mitigate the challenge of asynchronous data collection, including start, stop, call for assistance (CFA), and photo capture. These markers served as reference points for data synchronization. The event markers and the biosignal data were captured simultaneously and streamed to the Python system through LSL. The Biosignal Manager

[1] https://github.com/labstreaminglayer/LSL4Unity

system collected timestamps and real-time predicted emotional states from AffectivelyVR. This contextual information is invaluable for subsequent analyses. We stored this data in CSV files, facilitating offline analysis and enhancing the comprehensiveness of our dataset.

4.2.5 Companion Manager

The Companion Manager (CM) is designed to oversee companion-related interactions during navigation tasks. The system managed contextual data collection, streaming, voice assistance, photo management, and time management, activating only upon a user's request for assistance. When navigation assistance was requested, the CM system calculated and relayed directions, journey durations, and time estimations to reach monuments 1 and 2. The system also recorded details regarding remaining photos and monuments when a photo was captured. This contextual information was subsequently streamed to the Companion Adaptor system using the LabStreaming Layer (LSL).

The Companion was shown as an orb-shaped avatar, appearing while providing voice assistance. Speech functionality was enabled by the Microsoft Azure Text-to-Speech (TTS) SDK package integrated into Unity. The Speech Synthesis Markup Language (SSML) file format was utilized to encompass voice name, style, and content, generating speech as an audio file, played back using Unity's AudioSource component upon receiving confirmation of an SSML file update from the Companion Adaptor system via LSL. Beyond responding to Call for Assistance (CFA), the Companion was activated during photo captures, at the midway point, and in the final minute and ten seconds, maintaining a photo count, reducing it as users saved photos.

A texture was applied to the User's 'MainCamera', to change colors and create the perception of wearing colored sunglasses. The texture's color was set based on the physiological emotion (PE) obtained from the BioEmoVR (section 3.3.1).

4.3 Participants and Procedure

We recruited 15 participants (8 females, 7 males) aged between 22 and 47 (mean=31.1, SD=6.1). Each had normal or corrected vision and previous VR experience. After screening for neurological and psychological issues, participants provided informed consent and completed a pre-experiment questionnaire about demographics and VR experiences. They then wore a Shimmer sensor on their non-dominant hand, the OpenBCI EEG cap, and the HTC Vive Pro Eye VR HMD (figure 3).

We provided a brief practice session involving photography interactions to familiarise them with the VR tasks. The main experimental sequence began with a 60-second dark-room Rest phase, which served as a "Baseline" for biosignal analysis. During this phase, the moderator instructed participants to relax, clear their minds, and not think about anything. This phase allowed their emotional state to return to a neutral baseline. This was followed by a 7-minute Condition phase, where they photographed VR monuments. A virtual companion greeted them with: *"Welcome to the Virtual Reality photography contest! I will be your companion throughout this seven-minute journey. I will help you navigate to the monuments, and don't forget to have fun!"*. Post-trial, participants filled out the SAM scale, IPQ, GEQ, FSSQ, and NASA-TLX questionnaires. This sequence was repeated for all conditions, with a 5-minute break provided after the second trial.

4.4 Analysis

This section outlines the methodology to analyze the subjective, physiological, and behavioral measures acquired during the experiment. The scales employed include the Self-Assessment Manikin (SAM), IGroup Presence Questionnaire (IPQ), Game Experience Questionnaire (GEQ), NASA-TLX, and Flow Short State Questionnaire (FSSQ). We coded SAM responses on a 9-point scale, with 1 and 9 indicating the lowest and highest levels, respectively. For the IPQ, we used a 7-point Likert scale from -3 to +3, reverse-coded negatively worded items, and derived factor scores as Spatial Presence, Involvement, and Experienced Realism. GEQ responses were recorded on a 5-point Likert scale. Factor scores, like Positive Affect and Empathy, were averaged from pertinent items. The NASA-TLX data, computed using the Task Load Index (TLX) methodology and the FSSQ scale, was measured on a 7-point Likert scale ranging from 1 to 7. To compute the factor scores for each participant, we averaged the scores of the relevant items for 'Absorption by Activity' and 'Performance Fluency' flow characteristics.

The study examined the impact of motion artifacts on EEG, EDA, and PPG data. All data traces were inspected, and segments with pronounced and irregular fluctuations indicative of motion artifacts were excluded. Noise peaks in EEG signals were removed, and a bandpass filter between 1-50 Hz was applied. The EDA signals were processed by discarding noise peaks, and a bandpass filter spanning 0.05-5 Hz was utilized. A 10-second moving average filter was used on the EDA data to retain phasic components, then downsampled to 4 Hz. For PPG signals, a filter from 0.04-1 Hz was applied, baseline removal was conducted, and a moving average filter was employed for data smoothing. Utilizing the Neurokit2 python package [54], we extracted EEG, EDA, and HRV features. The EEG features accounted for emotional and cognitive aspects, including Frontal Alpha Asymmetry (FAA) [21], various theta and beta powers in select regions [6]. The engagement was quantified by the ratio of combined beta power to the combined sum of alpha and theta power [66].

5 Results

This section presents our analysis of the experiment data, focusing on emotion-adaptive (EA) and context-aware empathic (CAE) interactions in three parts; (1) findings from the subjective questionnaires, (2) results from physiological measurements, illuminating the participants' emotional and cognitive processes, and (3) behavioral measurements, to evaluate the companion's effectiveness. We only report on significant results. Overall, the key findings are:

1. A notable correlation was observed between subjective valence ratings and the EA and CAE constructs, with TLX ratings showcasing significant effects from CAE.
2. Physiological data, including EDA and HRV metrics like RMSSD, SDNN, and the LF/HF ratio, demonstrated pronounced influences from both EA and CAEIxs.
3. EEG analysis indicated significant effects of CAE on FA-Theta and FP-Beta, while F-Theta exhibited interaction between EA and CAE.

5.1 Subjective Measures

We recorded 60 post-trial subjective responses from four trials with 15 participants each. We used the SAM for valence and arousal, the IPQ for presence, the GEQ for affect and empathy, the FSSQ for flow, and the NASA-TLX for task load. We tested the normality of subjective ratings across conditions, namely A (NoEA-NoCAE), B (EA-NoCAE), C (NoEA-CAE), and D (EA-CAE) using the Shapiro-Wilk test [79]. Parametric scores were analyzed with a Two-way Repeated Measure ANOVA, followed by an Estimated Means Margin post hoc analysis with Tukey HSD correction. In contrast, non-parametric scores were analyzed using the Aligned Ranked Transform (ART) [87].

SAM: To analyze the effects of EA, CAE, and their interaction (EA*CAE) on SAM Valence and Arousal ratings across all conditions, we conducted an ART test. This indicated significant main effects of EA ($F(1,42)$=10.18, $p < 0.001$) and CAE ($F(1,42)$=27.44, $p < 0.001$) on SAM Valence, as well as a significant interaction effect between EA and CAE ($F(1,42)$=11.36, p=0.001). Post-hoc pairwise comparisons indicated significant differences between Condition A and C ($t(1,42)$=−5.31, $p < 0.0001$), Condition A and B ($t(1,42)$=−3.85, p=0.002), and Condition A and D ($t(1,42)$=−5.01, p=0.0001). Regarding SAM arousal, ART showed no significance.

IPQ: The IGroup Presence (IPQ) questionnaire assessed subjective ratings for overall presence (OP), spatial presence (SP), involvement (INV), and experienced realism (REAL). The Shapiro-Wilk test indicated normal distribution for ratings on OP, SP, INV, and REAL ($p > 0.05$). General presence (GP) did not show a normal distribution and was evaluated using the ART. Both the Two-way Repeated Measure ANOVA and ART revealed no significant main or interaction effects of EA and CAE on the respective dimensions.

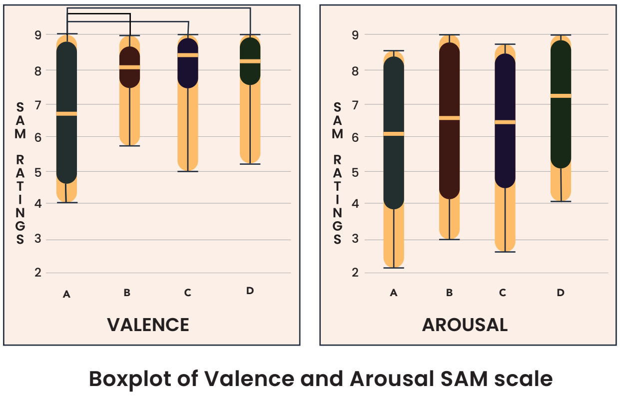

Fig. 5: Box plot of Valence and Arousal SAM Scale

GEQ: The effects of Emotion-Adaptive (EA) and Context-Aware Emotion (CAE) on the GEQ scores, specifically Empathy, Positive Affect, and Negative Affect, were investigated. Based on a Shapiro-Wilk normality test, these scores followed a non-parametric distribution ($p < 0.05$). The ART test was used for the evaluation.

For Empathy, a significant main effect was observed for CAE ($F(1,42)$=8.99, p=0.004). Post-hoc comparisons demonstrated significant differences between conditions A and D ($t(1,42)$=-3.86, p=0.002) as well as conditions B and D ($t(1,42)$=-2.73, p=0.04). Regarding Positive Affect and Negative Affect, there were also no significant differences between conditions.

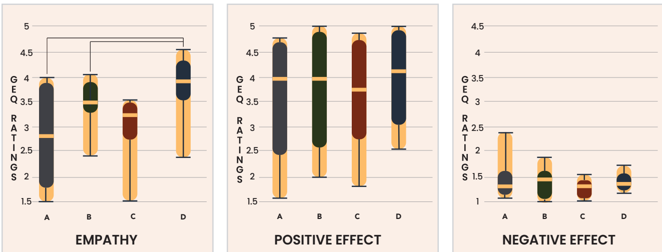

Fig. 6: Empathy, Positive and Negative Affect in GEQ Scale

FSSQ: The Shapiro-Wilk test confirmed that the factors of Absorption by Activity, Performance Fluency, and the overall Flow State Score from the FSSQ followed a normal distribution ($p > 0.05$). A Two-Way Repeated Measures ANOVA was employed to evaluate the effects of Emotion-Adaptive (EA), Context-Aware Emotion (CAE), and their interaction. For Activity Absorption, there were no significant effects for EA or CAE.

NASA-TLX: Based on the NASA-TLX Task Load Index (TLI), the Shapiro-Wilk normality test showed that the data was non-parametric ($p < 0.05$). Using Aligned Rank Transformation (ART), we evaluated the effects of Emotion-Adaptive (EA) and context-aware Emotion (CAE) and their combined influence. The effect of EA was non-significant, but a significant influence of CAE was found ($F(1,42)$=7.9, p=0.01). However, post-hoc comparisons between the conditions revealed no differences in cognitive load.

5.2 Physiological Measures

This section presents the findings from the Electrodermal Activity (EDA) features, Heart Rate Variability (HRV) metrics, and EEG features. We began our analysis with the Shapiro-Wilk test to verify data normality. If the data adhered to a normal distribution ($p > 0.5$), we proceeded with the Two-way Repeated Measures ANOVA. For non-normally distributed data ($p < 0.5$), the Aligned Rank Transformation (ART) was employed for statistical evaluation.

EDA: For the normalized EDA Peak Number (EDA-PN), the Shapiro-Wilk test confirmed its parametric nature ($p > 0.05$). Subsequently, the Two-way repeated measures ANOVA revealed a significant interaction effect between EA and CAE ($F(1,42)$=19.66, $p < 0.0001$). Tukey HSD post-hoc EMM comparison detected significant differences between conditions C-D ($t(1,42)$=-3.424, p=0.007) and B-D ($t(1,42)$=-3.26, p=0.01).

For the EDA Peak Amplitude (EDA-PA), the Shapiro-Wilk test indicated a non-parametric distribution ($p < 0.05$). The subsequent ART test showed a significant main effect of EA ($F(1,42)$=10.19, p=0.002). Post-hoc pairwise comparisons highlighted significant differences between conditions A-B ($t(1,42)$=-2.83, p=0.034) and A-D ($t(1,42)$=-3.73, p=0.003).

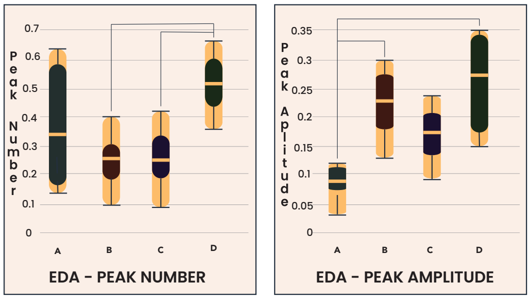

Fig. 7: Box plot of EDA peak number and amplitude features

HRV: RMSSD data was reported as non-parametric ($p < 0.05$). The ART test revealed both EA ($F(1,42)$=6.39, p=0.01) and CAE ($F(1,42)$=7.69, p=0.008) had a significant main effect on it. Post-hoc pairwise comparison revealed a significant difference between conditions A-B ($t(1,42)$=-2.84, p=0.03), B-C ($t(1,42)$=-3.56, p=0.005), and B-D ($t(1,42)$=-2.9, p=0.03). SDNN data was reported as non-parametric ($p < 0.05$). The ART test found both EA ($F(1,42)$=11.35, p=0.001) and CAE ($F(1,42)$=8.21, p=0.006) had a significant main effect. Post-hoc pairwise comparison found a significant difference between conditions A-B ($t(1,42)$=3.09, p=0.01), and B-C ($t(1,42)$=-4.1, p=0.001).

Low Frequency (LF) data were normally distributed ($p < 0.05$). A two-way repeated measures ANOVA reported no significant main effect of EA and CAE. No significant interaction between EA and CAE was also reported. High Frequency (HF) data were normally distributed ($p < 0.05$). Two-way repeated measures ANOVA reported no significant main effect of EA and CAE. No significant interaction between EA and CAE was also reported.

The ratio of LF and HF (LF/HF) powers were normally distributed ($p < 0.05$). A two-way repeated measures ANOVA reported a significant main effect of EA ($F(1,42)$=16.38, p=< 0.001) and CAE ($F(1,42)$=9.97, p=0.002). Post-hoc EMM comparison with Tukey HSD revealed a significant difference between conditions A-B ($t(1,42)$=3.5, p=0.005), B-C ($t(1,42)$=-4.48, p=0.0001), and B-D ($t(1,42)$=-2.78, p=0.03).

EEG: A Shapiro-Wilk test on the EEG data reported FAA, FA-Theta, PF-Alpha, FP-Beta, OP-Alpha, Alpha/Thera Ratio (ATR), F-Theta, FCP-Alpha, and Engagement as non-parametric, whereas M-Theta was parametric. For the FAA data, neither EA ($F(1,42)$=0.93, p=0.33) nor CAE ($F(1,42)$=1.62, p=0.20) showed significant effects. The Normalized FAA mean (SD) for Conditions A, B, C, and D were 0.50 (0.28), 0.46 (0.32), 0.52 (0.28), and 0.29 (0.22), respectively. An ART test reported a significant main effect of CAE ($F(1,42)$=11.11, p=0.001) on FA-Theta, but not for EA ($F(1,42)$=1.98, p=0.16). The FA-Theta mean (SD) for Conditions A, B, C, and D were 0.02 (0.02), 0.03 (0.04), 0.08 (0.09), and 0.14 (0.16), respectively. There was a non-significant main effect of EA ($F(1,42)$=1.79, p=0.18) and CAE ($F(1,42)$=0.54, p=0.46) on PF-Alpha. However, a significant interaction effect was found between EA-CAE ($F(1,42)$=6.22, p=0.01). The

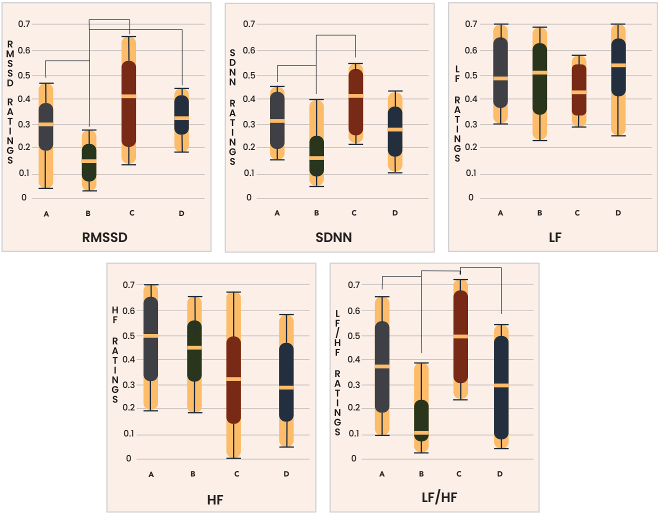

Fig. 8: HRV RMSSD, SDNN, HF, LF and LF/HF features

PF-Alpha mean (SD) for Conditions A, B, C, and D were 0.16 (0.27), 0.11 (0.24), 0.13 (0.27), and 0.20 (0.29), respectively. An ART test reported a significant main effect of CAE ($F(1,42)$=8.19, p=0.006) on FP-Beta, whereas EA showed no main effect and no interaction effect between EA and CAE. An ART test reported a significant main effect of CAE ($F(1,42)$=19.38, $p < 0.001$) on F-Theta, whereas EA ($F(1,42)$=2.46, p=0.12) showed no main effect. EA and CAE had a significant interaction effect ($F(1,42)$=8.51, p=0.005). Post-hoc pairwise comparison revealed a significant difference between conditions A-D ($t(1,42)$=3.33, p=0.009), B-C ($t(1,42)$=2.74, p=0.04), and B-D ($t(1,42)$=4.92, p=0.0001). Other measures, such as M-Theta, OP-Alpha, ATR, FCP-Alpha, and Engagement, showed no significant main or interaction effects for either EA or CAE.

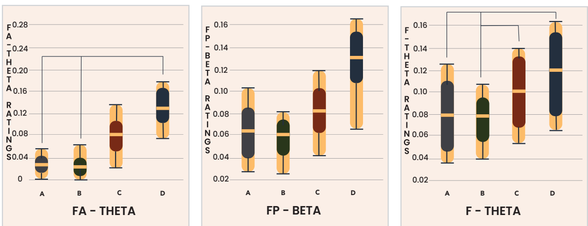

Fig. 9: Box plot of EEG FA Theta, FP Beta and F Theta features

6 Discussion

6.1 Enhancing Emotional Well-Being

The study aimed to investigate the first hypothesis that providing emotion-adaptive (EA) and context-aware empathic (CAE) interactions improves emotional well-being in VR (H1). We analyzed self-reported subjective responses and bio-responses to understand the emotional experience, sense of presence, and empathy towards virtual agents contributing to emotional well-being.

6.1.1 Emotional Experience

The study found that EA and CAE interventions positively influenced subjective emotional valence both individually and when they were combined, as demonstrated by significantly higher SAM Valence ratings in conditions involving these interventions (B, C, and D) compared to the control condition (A), which did not involve any interventions. However, the study did not report any significant effects of EA and CAE interventions on SAM Arousal or the GEQ positive and negative affect ratings, which aligns with previous research reporting mixed findings regarding the impact of environmental factors and empathic virtual agents on arousal and affect [10].

Nevertheless, the study found that EA and CAE interventions significantly impacted several physiological measures associated with emotional arousal and regulation, such as RMSSD, SDNN, LF/HF ratio, EDA-PA, and FA-Theta, which aligns with previous literature [2, 4]. The significant main effect of EA on EDA Peak Amplitude (EDA-PA) suggests that manipulating environmental color alone can increase emotional arousal. This supports the idea that integrating environmental color manipulation and CAE virtual agents independently and, in some cases, interactively influenced can lead to more positive emotional experiences and better regulate emotions.

The significant main effect of CAE on FA-Theta data suggests that the context-aware empathic virtual agent increased emotional arousal when combined with environmental color manipulation (Condition D). Similarly, the significant main effect of CAE on FP-Beta data indicates that the empathic virtual agent increased emotional arousal during attention and emotional processing tasks.

Overall, the study provides preliminary evidence supporting the potential benefits of integrating environmental color manipulation and empathic virtual agents in virtual reality environments to enhance emotional experiences, supporting Hypothesis **H**1**(a)**.

6.1.2 Sense of Presence

The study uncovered no significant impact of EA or CAE on subjective presence measures, such as overall presence, spatial presence, involvement, and realism. The non-significant results observed in this study regarding Emotion-Adaptation (EA) and Context-Aware Empathy (CAE) interventions on presence measures align with previous research [20] that has reported mixed or inconclusive findings concerning the impact of such interventions on presence in highly immersive and engaging VR environments.

Future research should consider introducing environmental quality and engagement level variations to assess whether these factors interact with or modulate EA and CAE's effects. While all conditions exhibited increased presence, the absence of significance contradicts hypothesis **H**1**(b)**, which states that EA and CAEIxs will improve the sense of presence in VR. To better understand these effects, future studies should examine the temporal dynamics of EA and CAE and more carefully account for the potential interference of high-quality environments and engaging activities when evaluating their impact on presence.

6.1.3 Empathic Connection

The GEQ analysis revealed a significant effect of context-aware empathic (CAE) interventions on players' perceived empathy towards the virtual agent. At the same time, emotion adaptation (EA) did not significantly impact affective responses. Furthermore, post hoc comparisons indicated that combining EA and CAE interventions (Condition D) elicited a stronger empathic response towards game characters than the control and EA alone (Conditions A and B). This finding aligns with Tielman et al.'s research [85], which explored the influence of verbal and textual presentation on adherence and user engagement, demonstrating the potential benefits of empathic virtual agents in therapeutic settings.

Overall, The study findings partially supported the first hypothesis, **H**1, that manipulating environmental color and using context-aware empathic interactions independently or, in some cases, interactively can enhance users' emotional experience and strengthen empathy with the virtual agent. By demonstrating the effectiveness of CAE interventions in promoting empathic responses and considering the non-significant impact on presence, this research emphasizes the need for a deeper understanding of how different strategies can be employed in designing empathic virtual agents.

6.2 Cognitive Load and Flow

We aimed to address H2 by investigating the effect of EA and CAE on cognitive load and flow state using subjective, physiological, and behavioral measures.

6.2.1 Cognitive Load

The present study emphasizes that context-aware empathic (CAE) interventions notably decrease cognitive load, but emotion-adaptive (EA) interventions exhibit an insignificant effect. In support of this, the research uncovered a significant influence on cognitive load when CAE was present, as evidenced by the perceived Task Load Index (NASA-TLX) and EEG Frontal Theta Activity. This finding is congruent with previous research that suggests a positive correlation between theta power and self-reported cognitive load [12]. Additionally, the results align with Lampen et al.'s [46] study, positing that context-aware assistance could effectively impact cognitive load by providing real-time personalized and relevant information. However, in this study, the mean TLX rating and F-Theta correlated to higher cognitive load, indicating increased cognitive demand in the tasks when the agent was context-aware empathic. There could be a possibility that the CAE responses using only audio cues were used as extra information that could lead to cognitive overload [28], whereas allocating responses via two channels (visual and auditory) could help avoid this overload [56].

The study partially validates hypothesis **H**2(**a**) that integrating emotion-adaptive (EA) and context-aware empathic (CAE) interventions can impact cognitive load. Despite identifying a significant primary effect of CAE interventions on cognitive load, the impact of EA interventions was statistically insignificant. Future research should delve deeper into the potential advantages of merging these interventions within more intricate and emotionally demanding tasks, refine the design of EA interventions, and explore an embodied virtual agent to better cater to users' emotional requirements.

6.2.2 Flow State

The findings indicated that either EA or CAE interventions did not significantly influence the self-reported flow state measures of Activity Absorption, Performance Fluency, or the overall Flow State Score. This aligns with prior research [27], where no significant self-reported flow states were observed when controlled for mindfulness and attention. Such results might be attributed to individual differences and the complex nature of the flow state.

The study's results partially supported hypothesis **H**2(**b**), suggesting combining EA and CAE interventions would enhance the flow state. Despite the absence of significant effects on self-reported flow state measures, the substantial impact on HRV-SDNN indicates that the integration of EA and CAE may still influence the physiological aspects of flow [60]. Further research is warranted to comprehend the intricate interplay between emotion-adaptive and context-aware empathic interventions and the multifaceted nature of the flow state.

7 Summary of Contributions

The study aimed to understand the impact of context-aware empathic (CAE) interactions on the user experience, with a focus on emotional and cognitive aspects of user experience within emotion-adaptive (EA) virtual reality (VR) environments. To investigate this, we asked three questions and sum the answers here:

RQ1: *How can physiological signals be used to predict emotions and facilitate CAEIxs in VR environments?*

A1: The BioEmoVR system, described in the Empathic Appraisal Module, utilized physiological signals to predict emotions and enable CAEIxs in VR settings. The research created a generalized emotion model with an accuracy of up to 91.67% using EEG, EDA, and HRV biosignals. The CAEVR system was designed to predict emotional states every 30 seconds, incorporating empathic virtual agents in VR tasks such as navigation and time management.

RQ2: *What are the effects of CAEIxs on elicited emotions, cognitive load and empathy towards a virtual agent in VR?*

A2: The results revealed that context-aware empathic (CAE) virtual agents, which adapt based on emotional and contextual signals, can enrich user experiences. Enhanced positive emotions and increased empathy towards the agent and environment independently or interactively were reported by participants. Additionally, combining CAE agents can influence cognitive burdens on users but boost empathy.

RQ3: *How can the impact of CAEIxs on users' emotional and cognitive load during VR experiences be evaluated?*

A3: The study suggests a multi-faceted approach to evaluating user experiences, merging subjective, physiological, and behavioral metrics. Tools such as the Self-Assessment Manikin (SAM) questionnaire, EDA features, HRV, and EEG metrics captured users' emotional responses. No definitive measure was pinpointed for assessing the presence, but the empathy component of the Game Experience Questionnaire's (GEQ) Social Presence module assessed users' psychological empathy. Metrics such as the NASA-TLX questionnaire, EDA measurements, and EEG responses gauged cognitive load. The HRV SDNN metric provided insights into flow states in CAE-based EA VR environments.

8 Limitations and Future Work

The effectiveness of the interventions faced several limitations that warrant future attention. Initially, an environment that was visually appealing and realistic might have drawn user attention, as evidenced by the significant pleasurable perceived emotion score. This might have contributed to an immersive experience, potentially overshadowing the effects of EA or CAE. Furthermore, participants' high engagement levels might have intensified the sense of presence, obscuring the impacts of EA or CAE. To counter these challenges, temporal analysis will examine the effects of EA and CAE during distinct VR experience periods like onboarding, the first half, and the latter half. Future studies will vary the environmental quality and engagement levels to ascertain if these factors influence the effects of EA and CAE interventions on presence. Employing more refined tools such as pupil and neuroscientific fNIRS measures will be considered to broaden understanding.

Another notable limitation was the study's small sample size, possibly hindering the generalizability of results. It has been proposed that subsequent research encompass a broader participant range of 30-50 individuals. Despite the interventions, participants did not mark significant emotional intensity shifts during the study. The empathic virtual agent's design, represented by a non-expressive floating orb, might not have been ideally suited to trigger notable emotional changes. The rule-based appraisal and response mechanism could also have contributed by offering unwarranted or irrelevant support. Upcoming studies will evaluate other virtual agent features and LLM-based intelligent agents that might amplify arousal effects noticeable in physiological responses. Additionally, the motion required in the VR experience could have produced motion artifacts, particularly in EEG readings, influencing real-time emotion identification. In response, cutting-edge artifact removal algorithms will be utilized, and system performance will be assessed. Lastly, future endeavors will evaluate the CAEVR system's impact, aiming to ascertain the influence of emotional states, flow state, presence, empathy, cognitive load, and trust in virtual agents and to derive an impact score to enhance the CAEVR system's efficacy.

9 Conclusion

In exploring Context-Aware Empathic VR (CAEVR), this study has paved the way for enhancing user experiences by adapting to real-time emotional feedback. Through integrating EEG, EDA, and HRV biosignals, a novel system was designed to recognize emotions in real-time within VR environments. Employing the CAEVR application showcased the potential for heightened user engagement, significantly enhancing positive emotions and empathy toward empathic virtual agents. Furthermore, the research highlighted the transformative capability of VR applications when underpinned by contextual and emotional intelligence. As the digital landscape continuously evolves, the findings of this study illuminate the profound promise that empathic interactions hold for the future of immersive VR experiences. Future work should refine these interactions and expand the application areas, ensuring a holistic and emotionally attuned VR environment.

Acknowledgments

This work is supported by the Empathic Computing grant under the Entrepreneurial Universities initiative of the TEC, New Zealand.

References

[1] G. D. Abowd and E. D. Mynatt. Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7(1):29–58, 2000. 2

[2] L. I. Aftanas and S. A. Golocheikine. Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution eeg investigation of meditation. *Neuroscience letters*, 310(1):57–60, 2001. 8

[3] C. B. Almazan. *Rover: architectural support for exposing and using context*. PhD thesis, 2010. 2

[4] B. M. Appelhans and L. J. Luecken. Heart rate variability as an index of regulated emotional responding. *Review of general psychology*, 10(3):229–240, 2006. 8

[5] L. Bartram, A. Patra, and M. Stone. Affective color in visualization. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pp. 1364–1374, 2017. 4

[6] E. Başar, C. Başar-Eroglu, S. Karakaş, and M. Schürmann. Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International journal of psychophysiology*, 39(2-3):241–248, 2001. 6

[7] S. Bouchard, F. Bernier, É. Boivin, B. Morin, and G. Robillard. Using biofeedback while immersed in a stressful videogame increases the effectiveness of stress management skills in soldiers. *PloS one*, 7(4):e36169, 2012. 2

[8] M. Bratman. Intention, plans, and practical reason. 1987. 4

[9] O. Brdiczka, M. Langet, J. Maisonnasse, and J. L. Crowley. Detecting human behavior models from multimodal observation in a smart home. *IEEE Transactions on automation science and engineering*, 6(4):588–597, 2008. 2

[10] T. Brosch, K. Scherer, D. Grandjean, and D. Sander. The impact of emotion on perception, attention, memory, and decision-making. *Swiss medical weekly*, 143(1920):w13786–w13786, 2013. 8

[11] A. Bunglowala and A. Bunglowala. Nonverbal communication: An integral part of teaching learning process. *International Journal of Research in Advent Technology*, 1:371–375, 2015. 2

[12] L. J. Castro-Meneses, J.-L. Kruger, and S. Doherty. Validating theta power as an objective measure of cognitive load in educational video. *Educational Technology Research and Development*, 68:181–202, 2020. 9

[13] Z. Chang, H. Bai, L. Zhang, K. Gupta, W. He, and M. Billinghurst. The impact of virtual agents' multimodal communication on brain activity and cognitive load in virtual reality. *Frontiers in Virtual Reality*, 3:179, 2022. 2

[14] T. L. Chartrand, W. W. Maddux, and J. L. Lakin. Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. *The new unconscious*, pp. 334–361, 2005. 1

[15] H. Chen, X. Liu, D. Yin, and J. Tang. A survey on dialogue systems: Recent advances and new frontiers. *Acm Sigkdd Explorations Newsletter*, 19(2):25–35, 2017. 4

[16] J. A. Coan, H. S. Schaefer, and R. J. Davidson. Lending a hand: Social regulation of the neural response to threat. *Psychological science*, 17(12):1032–1039, 2006. 4

[17] M. Csikszentmihalyi and I. S. Csikszentmihalyi. *Optimal experience: Psychological studies of flow in consciousness*. Cambridge university press, 1992. 3

[18] M. Csikszentmihalyi, R. Larson, et al. *Flow and the foundations of positive psychology*, vol. 10. Springer, 2014. 3

[19] B. M. Cuff, S. J. Brown, L. Taylor, and D. J. Howat. Empathy: A review of the concept. *Emotion review*, 8(2):144–153, 2016. 2

[20] J. J. Cummings and J. N. Bailenson. How immersive is enough? a meta-analysis of the effect of immersive technology on user presence. *Media psychology*, 19(2):272–309, 2016. 8

[21] R. J. Davidson. Affective style and affective disorders: Perspectives from affective neuroscience. *Cognition & emotion*, 12(3):307–330, 1998. 6

[22] M. H. Davis. *Empathy: A social psychological approach*. Routledge, 2018. 2

[23] M. M. De Graaf and S. B. Allouch. Exploring influencing variables for the acceptance of social robots. *Robotics and autonomous systems*, 61(12):1476–1486, 2013. 2

[24] Y. Dennemont, G. Bouyer, S. Otmane, and M. Mallem. 3d interaction assistance in virtual reality: a semantic reasoning engine for context-awareness. In *International Conference on Context-Aware Systems and Applications*, pp. 30–40. Springer, 2012. 2

[25] H. C. Devlin, J. Zaki, D. C. Ong, and J. Gruber. Not as good as you think? trait positive emotion is associated with increased self-reported empathy but decreased empathic performance. *PloS one*, 9(10):e110470, 2014. 4

[26] A. K. Dey. Understanding and using context. *Personal and ubiquitous computing*, 5:4–7, 2001. 2

[27] F. M. Diaz. Mindfulness, attention, and flow during music listening: An empirical investigation. *Psychology of Music*, 41(1):42–58, 2013. 9

[28] K. Fraser, I. Ma, E. Teteris, H. Baxter, B. Wright, and K. McLaughlin. Emotion, cognitive load and learning outcomes during simulation training. *Medical education*, 46(11):1055–1062, 2012. 9

[29] J. J. Gross. Emotion regulation: Current status and future prospects. *Psychological inquiry*, 26(1):1–26, 2015. 2

[30] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE transactions on visualization and computer graphics*, 23(6):1706–1724, 2016. 2

[31] C. P. Gumbheer, K. K. Khedo, and A. Bungaleea. Personalized and adaptive context-aware mobile learning: review, challenges and future directions. *Education and Information Technologies*, 27(6):7491–7517, 2022. 2

[32] K. Gupta, R. Hajika, Y. S. Pai, A. Duenser, M. Lochner, and M. Billinghurst. Measuring human trust in a virtual assistant using physiological sensing in virtual reality. In *2020 IEEE Conference on virtual reality and 3D user interfaces (VR)*, pp. 756–765. IEEE, 2020. 1

[33] K. Gupta, J. Lazarevic, Y. S. Pai, and M. Billinghurst. AffectivelyVR: Towards VR Personalized Emotion Recognition. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*. Association for Computing Machinery, 11 2020. doi: 10.1145/3385956.3422122 3

[34] S. G. Hart. Nasa task load index (tlx). volume 1.0; paper and pencil package. 1986. 5

[35] E. Hatfield, J. T. Cacioppo, and R. L. Rapson. Emotional contagion. *Current directions in psychological science*, 2(3):96–100, 1993. 2

[36] U. Hess and A. Fischer. Emotional mimicry: Why and when we mimic emotions. *Social and personality psychology compass*, 8(2):45–57, 2014. 2

[37] M. L. Hoffman. Sex differences in empathy and related behaviors. *Psychological Bulletin*, 84(4):712–722, 1977. doi: 10.1037/0033-2909.84.4.712 1

[38] M. L. Hoffman. *Empathy and moral development: Implications for caring and justice*. Cambridge University Press, 2001. 2

[39] A. C. Hurlbert and Y. Ling. Biological components of sex differences in color preference. *Current biology*, 17(16):R623–R625, 2007. 2

[40] S. B. i Badia, L. V. Quintero, M. S. Cameirao, A. Chirico, S. Triberti, P. Cipresso, and A. Gaggioli. Toward emotionally adaptive virtual reality for mental health applications. *IEEE journal of biomedical and health informatics*, 23(5):1877–1887, 2018. 2

[41] W. A. IJsselsteijn, Y. A. De Kort, and K. Poels. The game experience questionnaire. 2013. 5

[42] S. Järvelä, B. Cowley, M. Salminen, G. Jacucci, J. Hamari, and N. Ravaja. Augmented virtual reality meditation: Shared dyadic biofeedback increases social presence via respiratory synchrony. *ACM Transactions on Social Computing*, 4(2):1–19, 2021. 2

[43] V. Jurcak, D. Tsuzuki, and I. Dan. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. *Neuroimage*, 34(4):1600–1611, 2007. 5

[44] M. L. Knapp and J. A. Daly. *Handbook of interpersonal communication*. Sage, 2002. 2

[45] N. Kwallek, C. M. Lewis, and A. S. Robbins. Effects of office interior color on workers' mood and productivity. *Perceptual and Motor Skills*, 66(1):123–128, 1988. 2

[46] E. Lampen, J. Lehwald, and T. Pfeiffer. A context-aware assistance framework for implicit interaction with an augmented human. In *Virtual, Augmented and Mixed Reality. Industrial and Everyday Life Applications: 12th International Conference, VAMR 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22*, pp. 91–110. Springer, 2020. 9

[47] P. Lang. Behavioral treatment and bio-behavioral assessment: Computer applications. *Technology in mental health care delivery systems*, pp. 119–137, 1980. 5

[48] S. Lee, Y. Lee, S. Jang, and W. Woo. vr-ucam: Unified context-aware application module for virtual reality. In *Conference on Artificial Reality*, 2004. 2

[49] H. Liang, S. Liu, Y. Wang, J. Pan, J. Fu, and Y. Yuan. Eeg-based vr

scene adaptive generation system for regulating emotion. In *2023 9th International Conference on Virtual Reality (ICVR)*, pp. 361–368. IEEE, 2023. 2

[50] T. W. Liew, S.-M. Tan, T. M. Tan, and S. N. Kew. Does speaker's voice enthusiasm affect social cue, cognitive load and transfer in multimedia learning? *Information and Learning Sciences*, 121(3/4):117–135, 2020. 2

[51] M. M. Linehan. *Skills training manual for treating borderline personality disorder.* Guilford press, 1993. 4

[52] P. L. Lockwood, A. Seara-Cardoso, and E. Viding. Emotion regulation moderates the association between empathy and prosocial behavior. *PloS one*, 9(5):e96555, 2014. 2

[53] V. Lorenzetti, B. Melo, R. Basílio, C. Suo, M. Yücel, C. J. Tierra-Criollo, and J. Moll. Emotion regulation using virtual environments and real-time fmri neurofeedback. *Frontiers in neurology*, p. 390, 2018. 2

[54] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. H. A. Chen. NeuroKit2: A python toolbox for neurophysiological signal processing. *Behavior Research Methods*, 53(4):1689–1696, feb 2021. doi: 10.3758/s13428-020-01516-y 3, 6

[55] J. Marín-Morales, C. Llinares, J. Guixeres, and M. Alcañiz. Emotion recognition in immersive virtual reality: From statistics to affective computing. *Sensors (Switzerland)*, 20(18):1–26, 2020. doi: 10.3390/s20185163 1

[56] E. K. Miller and T. J. Buschman. Working memory capacity: Limits on the bandwidth of cognition. *Daedalus*, 144(1):112–122, 2015. 9

[57] A. Moldoveanu, O. Mitruț, N. Jinga, C. Petrescu, F. Moldoveanu, V. Asavei, A. M. Anghel, and L. Petrescu. Immersive phobia therapy through adaptive virtual reality and biofeedback. *Applied Sciences*, 13(18):10365, 2023. 2

[58] J. Moon, M. Jeong, S. Oh, T. H. Laine, and J. Seo. Data collection framework for context-aware virtual reality application development in unity: Case of avatar embodiment. *Sensors*, 22(12):4623, 2022. 1

[59] K. Oatley and P. N. Johnson-Laird. Cognitive approaches to emotions. *Trends in cognitive sciences*, 18(3):134–140, 2014. 4

[60] E. Ortega and C. Wang. Pre-performance physiological state: Heart rate variability as a predictor of shooting performance. *Applied psychophysiology and biofeedback*, 43:75–85, 2018. 9

[61] A. Paiva, I. Leite, H. Boukricha, and I. Wachsmuth. Empathy in virtual agents and robots: A survey. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 7(3):1–40, 2017. 1, 2

[62] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos. Context aware computing for the internet of things: A survey. *IEEE communications surveys & tutorials*, 16(1):414–454, 2013. 2

[63] P. C. Petrantonakis and L. J. Hadjileontiadis. Emotion recognition from eeg using higher order crossings. *IEEE Transactions on information Technology in Biomedicine*, 14(2):186–197, 2009. 5

[64] A. Pinilla, J. Garcia, W. Raffe, J.-N. Voigt-Antons, R. P. Spang, and S. Möller. Affective visualization in virtual reality: An integrative review. *Frontiers in Virtual Reality*, 2:630731, 2021. 2

[65] T. Piumsomboon, Y. Lee, G. A. Lee, A. Dey, and M. Billinghurst. Empathic Mixed Reality: Sharing What You Feel and Interacting with What You See. In *2017 International Symposium on Ubiquitous Virtual Reality (ISUVR)*, pp. 38–41. IEEE, 6 2017. doi: 10.1109/ISUVR.2017.20 1

[66] A. T. Pope, E. H. Bogart, and D. S. Bartolome. Biocybernetic system evaluates indices of operator engagement in automated task. *Biological psychology*, 40(1-2):187–195, 1995. 6

[67] H. Prendinger, H. Dohi, H. Wang, S. Mayer, and M. Ishizuka. Empathic embodied interfaces: Addressing users' affective state. In *Tutorial and Research Workshop on Affective Dialogue Systems*, pp. 53–64. Springer, 2004. 2

[68] S. D. Preston. A perception-action model for empathy. *Empathy in mental illness*, 1:428–447, 2007. 2

[69] S. D. Preston and F. B. De Waal. Empathy: Its ultimate and proximate bases. *Behavioral and brain sciences*, 25(1):1–20, 2002. 2

[70] A. Raja, E. Niforatos, and C. Schneegass. An emotion-adaptive vr experience for recreational use in eldercare. In *Proceedings of Mensch und Computer 2023*, pp. 354–358. 2023. 1

[71] H. Rashkin, E. M. Smith, M. Li, and Y.-L. Boureau. Towards empathetic open-domain conversation models: A new benchmark and dataset. *arXiv preprint arXiv:1811.00207*, 2018. 4

[72] F. Rheinberg, R. Vollmeyer, and S. Engeser. Die erfassung des flowerlebens. 2003. 5

[73] S. H. Rodrigues, S. Mascarenhas, J. Dias, and A. Paiva. A process model of empathy for virtual agents. *Interacting with Computers*, 27(4):371–391, 2015. 2

[74] M. Q. Rosenzweig. Breaking bad news: a guide for effective and empathetic communication. *The Nurse Practitioner*, 37(2):1, 2012. 2

[75] N. S. Ryan, J. Pascoe, and D. R. Morse. Enhanced reality fieldwork: the context-aware archaeological assistant. In *Computer applications in archaeology*. Tempus Reparatum, 1998. 2

[76] B. N. Schilit and M. M. Theimer. Disseminating active map information to mobile hosts. *IEEE network*, 8(5):22–32, 1994. 2

[77] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001. 5

[78] M. Seehausen, P. Kazzer, M. Bajbouj, H. R. Heekeren, A. M. Jacobs, G. Klann-Delius, W. Menninghaus, and K. Prehn. Effects of empathic social responses on the emotions of the recipient. *Brain and cognition*, 103:50–61, 2016. 4

[79] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611, 1965. 6

[80] K. M. Sheldon, B. Garton, R. Orr, and A. Smith. The advisor quality survey: Good college advisors are available, knowledgeable, and autonomy supportive. *Journal of College Student Development*, 56(3):261–273, 2015. 4

[81] D. Shin. Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in human behavior*, 78:64–73, 2018. 3

[82] N. A. Stavrou, S. A. Jackson, Y. Zervas, and K. Karteroliotis. Flow experience and athletes' performance with reference to the orthogonal model of flow. *Sport Psychologist*, 21(4):438–457, 2007. doi: 10.1123/tsp.21.4.438 3

[83] S. Stürmer and M. Snyder. The psychological study of group processes and intergroup relations in prosocial behavior: Past, present, future. *The psychology of prosocial behavior: Group processes, intergroup relations, and helping*, pp. 1–10, 2009. 4

[84] N. M. Thompson, A. Uusberg, J. J. Gross, and B. Chakrabarti. Empathy and emotion regulation: An integrative account. *Progress in brain research*, 247:273–304, 2019. 2

[85] M. L. Tielman, M. A. Neerincx, M. Van Meggelen, I. Franken, and W.-P. Brinkman. How should a virtual agent present psychoeducation? influence of verbal and textual presentation on adherence. *Technology and Health Care*, 25(6):1081–1096, 2017. 8

[86] D. Watson, L. A. Clark, and A. Tellegen. Development and validation of brief measures of positive and negative affect: the panas scales. *Journal of personality and social psychology*, 54(6):1063, 1988. 5

[87] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 143–146, 2011. 6

[88] E. Yigitbas, J. Heindörfer, and G. Engels. A context-aware virtual reality first aid training application. In *Proceedings of Mensch und Computer 2019*, pp. 885–888. 2019. 2